

Optimal control of convective FitzHugh–Nagumo equation



Murat Uzunca^{a,*}, Tuğba Küçükseyhan^d, Hamdullah Yücel^b,
Bülent Karasözen^{b,c}

^a Department of Industrial Engineering, University of Turkish Aeronautical Association, 06790 Ankara, Turkey

^b Institute of Applied Mathematics, Middle East Technical University, 06800 Ankara, Turkey

^c Department of Mathematics, Middle East Technical University, 06800 Ankara, Turkey

^d Department of Mathematics, Balıkesir University, 10145 Balıkesir, Turkey

ARTICLE INFO

Article history:

Received 24 July 2016

Received in revised form 1 December 2016

Accepted 21 February 2017

Available online 17 March 2017

Keywords:

FitzHugh–Nagumo equation

Traveling waves

Sparse controls

Second order optimality conditions

Discontinuous Galerkin method

ABSTRACT

We investigate smooth and sparse optimal control problems for convective FitzHugh–Nagumo equation with traveling wave solutions in moving excitable media. The cost function includes distributed space–time and terminal observations or targets. The state and adjoint equations are discretized in space by symmetric interior point Galerkin (SIPG) method and by backward Euler method in time. Several numerical results are presented for the control of the traveling waves. We also show numerically the validity of the second order optimality conditions for the local solutions of the sparse optimal control problem for vanishing Tikhonov regularization parameter. Further, we estimate the distance between the discrete control and associated local optima numerically by the help of the perturbation method and the smallest eigenvalue of the reduced Hessian.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Spatially and temporally varying structures occur in the form of Turing patterns, traveling waves, fronts, periodic pulses in many physical, chemical, and biological systems. They are described mathematically in the form of coupled semi-linear partial differential equations (PDEs) [1]. The FitzHugh–Nagumo (FHN) equation is one of the most known generic models in physiology, describing complex wave phenomena in excitable or oscillatory media. The most known type of the FHN equation in the literature consists of a PDE with a non-monotone nonlinear term, known as activator equation, and an ordinary differential equation (ODE), known as inhibitor equation. We call such kind of activator–inhibitor system as classical FHN equation. Another type of the activator–inhibitor system is diffusive FHN equation consisting of an activator PDE, and one or two inhibitor PDEs [2]. Recently, the convective FHN equation has been proposed as a model for wave propagation in blood coagulation and bioreactor systems [3–5]. The presence of convective field leads to complex wave phenomena, like triggering and autonomous waves in a moving excitable media [4].

The classical FHN equation and its PDE part, the Schlögl or Nagumo equations were investigated theoretically and numerically for the wave-type optimal control solutions [6–9]. Optimal control of semi-linear parabolic equation is an active research field with many applications in controlling pattern formation [10] and feedback control of the monodomain equations in cardiac electrophysiology [11] to give a few examples.

In this paper, we investigate the numerical treatment of optimal control problems governed by the convective FHN equation. The uncontrolled solutions of the convective FHN equation behave like traveling waves [4]. To control such a

* Corresponding author.

E-mail addresses: muzunca@thk.edu.tr (M. Uzunca), kucukseyhan@balikesir.edu.tr (T. Küçükseyhan), yucelh@metu.edu.tr (H. Yücel), bulent@metu.edu.tr (B. Karasözen).

<http://dx.doi.org/10.1016/j.camwa.2017.02.028>

0898-1221/© 2017 Elsevier Ltd. All rights reserved.

traveling wave, we use sparse control, which is a non-smooth L^1 -control cost in addition to the L^2 -control cost. When the control signals are localized in some regions of the space–time cylinder, sparse control provides solutions without any a priori knowledge of these sub-areas. Sparse optimal control was first investigated in [12] for linear elliptic equations and later studied in [13,8,14] for semi-linear elliptic and parabolic equations.

Here, we use symmetric interior penalty Galerkin (SIPG) method for space discretization. The discontinuous Galerkin (dG) methods are more stable for convection dominated problems than the continuous finite element methods and they do not require the stabilization terms like the streamline upwind/Petrov–Galerkin method (SUPG). The dG methods have several advantages compared to other numerical techniques such as finite volume and finite element methods; the trial and test spaces can be easily constructed, inhomogeneous boundary conditions and curved boundaries can be handled easily. They are also flexible in handling non-matching grids and in designing hp-adaptive mesh refinement. The dG methods were successfully applied for the steady state [15–17], the time dependent linear convection–diffusion–reaction [18,19], and the semi-linear steady state [20,21] optimal control problems. To solve the optimization problem, we here apply the *optimize–discretize* approach. The first order optimality conditions are derived and then they are discretized by using the dG method. We employ the projected nonlinear conjugate gradient (CG) method as an optimization algorithm [22]. We show the controllability of the traveling waves of the FHN equation with target functions in the whole space–time domain and at the final time with and without sparse controls. In addition, the performance of the method is demonstrated for convection dominated problems by increasing the wave velocity, for sparse controls with different sparse parameters.

Due to the semi-linearity of the FHN equation, the control problem is non-convex. Therefore, the fulfillment of the first necessary conditions does not imply the optimality. In order to guarantee the optimality, the second-order sufficient optimality conditions (SSCs) have to be checked. In the recent years, the fulfillment of the SSCs for infinite dimensional and finite dimensional semi-linear PDE constrained optimal control problems has been investigated extensively (see, e.g., [23] for a recent overview). Except few examples with analytical solutions, it is not possible to prove the SSCs for the infinite dimensional problems since the unknown optimal solution is required. Therefore the finite dimensional approximations of the infinite dimensional problem are considered. Provided that the local minima satisfies the SSC, one can check the SSC numerically by finding a bound for the distance between the local minima and discrete solution [24,25]. For this purpose, the associated coercivity constant of the reduced Hessian operator is estimated numerically by computing its smallest eigenvalue. Similar techniques were applied to measure how far the control obtained by a reduced order optimization model is away from a local full control solution [26,27]. Moreover, the Tikhonov regularization parameter in the cost function expresses the cost of the control, and it increases the numerical stability of the optimal solution. Recently the second optimality conditions have been investigated for semi-linear parabolic control problems with the objective function, not including the Tikhonov regularization term [23]. We test the discrete optimization problem for vanishing Tikhonov regularization parameter as in [8]. The numerical results of the control of two dimensional waves confirm the convergence of the optimal solutions for vanishing Tikhonov regularization parameter as it was demonstrated for the one dimensional wave solutions of the classical FHN equation in [8,9].

The paper is organized as follows: In the next section the optimal control problem governed by the convective FHN equation is described as a model problem. We first prove the existence and uniqueness of a convective FHN equation, called as a state equation, by transforming into the one with monotone nonlinearity. Then, we introduce the first and second order optimality conditions. In Section 3 we give a full discretization of the optimality system using the SIPG method in space and the backward Euler discretization in time. In Section 4, we discuss some benchmark examples with and without sparse controls. We investigate the effect of the Tikhonov parameter as it goes to zero. Further, with the help of the perturbation method and the smallest eigenvalue of the reduced Hessian, we find a bound for the distance between the local minima and discrete solution. The paper ends with some conclusions.

2. Optimal control of the convective FHN system

In this paper, we consider optimal control problems governed by the following convective FHN system:

$$y_t(x, t) - d_y \Delta y(x, t) + \mathbf{V} \cdot \nabla y(x, t) + g(y(x, t)) + z(x, t) = u(x, t) \quad \text{in } Q_T, \quad (2.1a)$$

$$z_t(x, t) - d_z \Delta z(x, t) + \mathbf{V} \cdot \nabla z(x, t) + \varepsilon(z(x, t) - c_3 y(x, t)) = 0 \quad \text{in } Q_T, \quad (2.1b)$$

with homogeneous Neumann boundary conditions

$$\partial_n y(x, t) = 0, \quad \partial_n z(x, t) = 0 \quad \text{on } \sum_T^N, \quad (2.1c)$$

Dirichlet boundary conditions

$$y(x, t) = y_D(x, t), \quad z(x, t) = z_D(x, t) \quad \text{on } \sum_T^D, \quad (2.1d)$$

and initial conditions

$$y(x, 0) = y_0(x), \quad z(x, 0) = z_0(x) \quad \text{in } \Omega. \quad (2.1e)$$

In this setting, let $T > 0$ be a fixed end time. We denote Q_T the time space cylinder $Q_T := \Omega \times (0, T)$, where $\Omega = (0, L) \times (0, H)$ is a bounded, Lipschitz domain. The lateral surface is denoted by $\Sigma = \Gamma \times (0, T)$. We use the notation $\sum_T^D := \Gamma_D \times (0, T)$ and $\sum_T^N := \Gamma_N \times (0, T)$, where Dirichlet Γ_D and Neumann Γ_N boundaries, where the Dirichlet $y_D, z_D \in H^{3/2}(\Gamma_D)$ and the Neumann boundary conditions are prescribed. Moreover, the initial functions are given as $y_0, z_0 \in L^\infty(\Omega)$. We denote the outward unit normal vector and the associated outward normal derivative on $\partial\Omega$ by $\mathbf{n}$ and ∂_n , respectively. The diffusion coefficients are denoted by d_y and d_z . The parameters c_3 and ε are real constants. Further, the function $g(y)$ denotes the cubic polynomial nonlinearity

$$g(y) = c_1 y(y - c_2)(y - 1) \quad (2.2)$$

with the nonnegative real numbers $c_i, i = 1, 2$, which is monostable for $0 < c_1 < 20$ and $c_2 = 0.02$ [4] in contrast to the bistable cubic nonlinearity for the Schlögl equation [6], the classical FHN equation [7], and the diffusive FHN equation [2]. The velocity field denoted by $\mathbf{V} = (V_{x_1}, V_{x_2})$ is given along the x_1 -direction with a parabolic profile

$$V_{x_1}(x_2) = ax_2(H - x_2), \quad V_{\max} = \frac{1}{4}aH^2, \quad a > 0, \quad V_{x_2} = 0, \quad (2.3)$$

where $V_{\max}$ denotes the maximum wave speed of the velocity field. Moreover, the velocity field is divergence free, i.e., $\operatorname{div} \mathbf{V} = 0$.

We also make the following assumption for the solutions y, z

$$0 =: y_0 \leq y \leq y_1, \quad 0 =: z_0 \leq z \leq z_1 \quad \text{a.e. } Q_T, \quad (2.4)$$

which is admissible for the sake of physical realism. We note that the bounds are constant.

Here, we want to minimize an objective function of misfit type, i.e., the function is designed to penalize deviations of the function values from the observed or measured data. We formulate our minimization functional such as

$$(P_\mu) \quad \left\{ \min_{u \in \mathcal{U}_{ad}} J(u) := I(u) + \mu j(u), \right. \quad (2.5)$$

with

$$\begin{aligned} I(u) &= \frac{\omega_Q^y}{2} \int_0^T \int_\Omega (y(x, t) - y_Q(x, t))^2 dx dt + \frac{\omega_Q^z}{2} \int_0^T \int_\Omega (z(x, t) - z_Q(x, t))^2 dx dt \\ &\quad + \frac{\omega_T^y}{2} \int_\Omega (y(x, T) - y_T(x, T))^2 dx + \frac{\omega_T^z}{2} \int_\Omega (z(x, T) - z_T(x, T))^2 dx \\ &\quad + \frac{\omega_u}{2} \int_0^T \int_\Omega (u(x, t))^2 dx dt, \\ j(u) &= \int_0^T \int_\Omega |u(x, t)| dx dt, \end{aligned}$$

where the pair (y, z) denotes the solution of (2.1) associated to control u . In (2.1), the partial differential equation for y is said *activator* equation, while the one for z is called the *inhibitor* equation. The functions $y_T, z_T \in L^\infty(\Omega)$ and $y_Q, z_Q \in L^\infty(Q_T)$ are the given targets or desired states. We have given constants $\omega_Q^y, \omega_Q^z, \omega_T^y, \omega_T^z$, Tikhonov regularization parameter $\omega_u \geq 0$, and sparse parameter $\mu \geq 0$.

We consider the optimization problem (2.5) with pointwise box constraints

$$u \in \mathcal{U}_{ad} := \{u \in L^\infty(Q_T) : u_a \leq u(x, t) \leq u_b \text{ for a.e. } (x, t) \in Q_T\} \quad (2.6)$$

with the real numbers $u_a \leq u_b$.

The aim of the optimal control is to ensure that the state variables y and z are as close as possible to the desired or observed states by minimizing the objective functional in the L^2 or L^1 -norms. For well-defined optimal solutions, one has to show that there exists a unique solution (y, z) of (2.1) for each $u \in \mathcal{U}_{ad}$. The FHN equation (2.1) belongs to the class of semi-linear parabolic equations with a non-monotone nonlinearity. The theory of the existence and uniqueness of solutions (y, z) of the state equation (2.1) is more delicate than the monotone nonlinearities [6,7]. In the next section, we show the existence and uniqueness of the state equation (2.1) by transforming the state equation to one with the monotone nonlinearity for the convective FHN equation (2.1).

2.1. Well-posedness of the state equation

The existence and uniqueness of a weak solution for the Schlögl and FHN equation was shown in [6] by transforming (2.1) into the one with monotone nonlinearity using the transformation $y = e^{\eta t} v$ with sufficiently large parameter η [6]. Here we apply the same technique for the FHN equation with the convective term. Next, we construct the upper and lower solutions for the transformed equation, which yield pointwise bounds for the desired solution following [28]. These bounds

are then used as an initial iterate to construct two monotonically convergent sequences. Finally, we show that their common limit is the unique solution of the transformed equation with the monotone nonlinearity.

Let us first perform the transformation of the state equation (2.1) by substituting $y = e^{\eta t} v$. Then, we obtain the following system:

$$v_t - d_y \Delta v + \mathbf{V} \cdot \nabla v + e^{-\eta t} g(e^{\eta t} v) + \eta v = e^{-\eta t} (u - z) \quad \text{in } Q_T, \quad (2.7a)$$

$$z_t - d_z \Delta z + \mathbf{V} \cdot \nabla z + \varepsilon(z - c_3 e^{\eta t} v) = 0 \quad \text{in } Q_T \quad (2.7b)$$

with the boundary conditions

$$\partial_n v = 0, \quad \partial_n z = 0 \quad \text{on } \sum_T^N, \quad \text{and} \quad e^{\eta t} v = y_D, \quad z = z_D \quad \text{on } \sum_T^D, \quad (2.7c)$$

and the initial conditions

$$v(x, 0) = y_0(x), \quad z(x, 0) = z_0(x) \quad \text{in } \Omega. \quad (2.7d)$$

Here, the nonlinear term

$$\tilde{g} : v \mapsto e^{-\eta t} g(e^{\eta t} v) + \eta v$$

is a monotone non-decreasing function with respect to v for all $(x, t) \in Q_T$. Moreover, it satisfies the following properties [29, Sec. 4.3]:

- (i) For every fixed $v \in \mathbb{R}$ is Lebesgue measurable in Q_T .
- (ii) For almost all $(x, t) \in Q_T$, $\tilde{g}$ is twice continuously differentiable with respect to v and locally Lipschitz continuous of order 2 with respect to v , i.e., there exists $L(\rho) = 6c_1 e^{2\eta t} > 0$ such that

$$|\tilde{g}_{vv}(x, t, v_1) - \tilde{g}_{vv}(x, t, v_2)| \leq L(\rho) |v_1 - v_2|$$

holds for all $v_1, v_2 \in \mathbb{R}$ with $|v_i| \leq \rho, i = 1, 2$.

The nonlinearity is uniformly bounded and monotone increasing in the following sense:

- (iii) There exists a constant $C = c_1 c_2 + \eta + 2e^{\eta t} (c_1 c_2 + c_1) > 0$ such that

$$|\tilde{g}(x, t, 0)| + |\tilde{g}_v(x, t, 0)| + |\tilde{g}_{vv}(x, t, 0)| \leq C.$$

- (iv) It holds $0 \leq \tilde{g}_v(x, t, v)$ for almost all $(x, t) \in Q_T$, all $v \in \mathbb{R}$.

Before defining a weak formulation of the system (2.7), we need to define the following Hilbert space

$$W(0, T) := \{w \in L^2(0, T; V); w' \in L^2(0, T; V^*)\}$$

equipped with the norm

$$\|w\|_{W(0, T)} = \left(\int_0^T (\|w(t)\|_V^2 + \|v_t(t)\|_{V^*}^2) dt \right)^{\frac{1}{2}},$$

where $V = H^1(\Omega)$ and V^* is the dual space of V . Now, we can define a weak solution of the system (2.7).

Definition 2.1. A pair of functions $(v, z) \in (W(0, T) \cap L^\infty(Q_T))^2$ is called weak solution of the system (2.7), if the equations

$$\int_0^T (v_t, \varphi)_{V^*, V} dt + \iint_{Q_T} (d_y \nabla v \cdot \nabla \varphi + \mathbf{V} \cdot \nabla v \varphi + \tilde{g} \varphi - e^{\eta t} (u - z) \varphi) dx dt = 0, \quad (2.8a)$$

$$\int_0^T (z_t, \varphi)_{V^*, V} dt + \iint_{Q_T} (d_z \nabla z \cdot \nabla \varphi + \mathbf{V} \cdot \nabla z \varphi + \varepsilon(z - c_3 e^{\eta t} v) \varphi) dx dt = 0, \quad (2.8b)$$

and

$$v(\cdot, 0) = y_0, \quad z(\cdot, 0) = z_0 \quad (2.8c)$$

are satisfied for all $\varphi \in L^2(0, T; H^1(\Omega))$. It is noted that ∇ denotes the gradient with respect to x .

Next, we give the definition of the ordered upper and lower solutions as done in [30,28]. The pair functions $(\tilde{v}, \tilde{z})$ and $(\hat{v}, \hat{z})$ in $(W(0, T) \cap L^\infty(Q_T))^2$ are said to be ordered upper and lower solutions of (2.7), respectively, if they satisfy

$$(\hat{v}, \hat{z}) \leq (\tilde{v}, \tilde{z})$$

and

$$\begin{aligned} \widehat{v}_t - d_y \Delta \widehat{v} + \mathbf{V} \cdot \nabla \widehat{v} + \widetilde{g}(x, t, \widehat{v}) - e^{-\eta t}(u - \widetilde{z}) &\leq 0 & \leq \widetilde{v}_t - d_y \Delta \widetilde{v} + \mathbf{V} \cdot \nabla \widetilde{v} + \widetilde{g}(x, t, \widetilde{v}) - e^{-\eta t}(u - \widehat{z}), \\ \widehat{z}_t - d_z \Delta \widehat{z} + \mathbf{V} \cdot \nabla \widehat{z} + \varepsilon(z - c_3 e^{\eta t} \widehat{v}) &\leq 0 & \leq \widetilde{z}_t - d_z \Delta \widetilde{z} + \mathbf{V} \cdot \nabla \widetilde{z} + \varepsilon(z - c_3 e^{\eta t} \widetilde{v}), \\ \partial_n \widehat{v} &\leq 0 & \leq \partial_n \widetilde{v}, \\ \partial_n \widehat{z} &\leq 0 & \leq \partial_n \widetilde{z}, \\ \widehat{v} &\leq e^{-\eta t} y_D & \leq \widetilde{v}, \\ \widehat{z} &\leq z_D & \leq \widetilde{z}, \\ \widehat{v}(x, 0) &\leq y_0(x) & \leq \widetilde{v}(x, 0), \\ \widehat{z}(x, 0) &\leq z_0(x) & \leq \widetilde{z}(x, 0). \end{aligned}$$

By taking

$$\widetilde{v}(x, t) = \widetilde{z}(x, t) = M, \quad \widehat{v}(x, t) = \widehat{z}(x, t) = 0$$

for some $M > 0$, we can rewrite our system (2.7) as

$$v_t - d_y \Delta v + \mathbf{V} \cdot \nabla v + \widetilde{g}(e^{\eta t} v) - e^{-\eta t}(u - M) = e^{-\eta t}(M - z), \quad (2.9a)$$

$$z_t - d_z \Delta z + \mathbf{V} \cdot \nabla z + \varepsilon(z - c_3 e^{\eta t} M) = \varepsilon c_3 e^{\eta t}(M - v), \quad (2.9b)$$

$$\partial_n v = 0, \quad \partial_n z = 0, \quad (2.9c)$$

$$v = e^{-\eta t} y_D, \quad z = z_D, \quad (2.9d)$$

$$v(x, 0) = y_0(x), \quad z(x, 0) = z_0(x). \quad (2.9e)$$

Here, we have

$$\begin{aligned} \frac{\partial(e^{-\eta t}(M - z))}{\partial v}, \quad \frac{\partial(\varepsilon c_3 e^{\eta t}(M - v))}{\partial z} &\geq 0, \\ \frac{\partial(e^{-\eta t}(M - z))}{\partial z}, \quad \frac{\partial(\varepsilon c_3 e^{\eta t}(M - v))}{\partial v} &\leq 0 \end{aligned}$$

for all $v, z \in [0, M]$.

Now, we can state the existence and uniqueness of the system (2.7) for each control variable u .

Theorem 2.2. Assume that the initial conditions y_0 and z_0 are nonnegative functions, and (2.4) holds. Then, the system (2.7) admits a unique solution $(v, z) \in (W(0, T) \cap C(Q_T))^2$ for each control $u \in \mathcal{U}_{ad}$.

Proof. We adopt the iteration technique introduced in [28] and construct sequences $\{(\bar{v}^k, \bar{z}^k)\}_{k=0}^\infty, \{(\underline{v}^k, \underline{z}^k)\}_{k=0}^\infty$ with initial elements

$$\begin{aligned} \bar{v}^0 &= \widetilde{v} = M, & \bar{z}^0 &= \widetilde{z} = M, \\ \underline{v}^0 &= \widehat{v} = 0, & \underline{z}^0 &= \widehat{z} = 0. \end{aligned}$$

Initiating from $(\bar{v}^k, \bar{z}^k)$ and $(\underline{v}^k, \underline{z}^k)$, $(\bar{v}^{k+1}, \bar{z}^{k+1})$ and $(\underline{v}^{k+1}, \underline{z}^{k+1})$ are found by solving

$$\begin{aligned} \bar{v}_t^{k+1} - d_y \Delta \bar{v}^{k+1} + \mathbf{V} \cdot \nabla \bar{v}^{k+1} + \widetilde{g}(e^{\eta t} \bar{v}^{k+1}) - e^{-\eta t}(u - M) &= e^{-\eta t}(M - \bar{z}^k), \\ \bar{z}_t^{k+1} - d_z \Delta \bar{z}^{k+1} + \mathbf{V} \cdot \nabla \bar{z}^{k+1} + \varepsilon(\bar{z}^{k+1} - c_3 e^{\eta t} M) &= \varepsilon c_3 e^{\eta t}(M - \underline{v}^k), \\ \partial_n \bar{v}^{k+1} &= 0, & \partial_n \bar{z}^{k+1} &= 0, \\ \bar{v}^{k+1} &= e^{-\eta t} y_D, & \bar{z}^{k+1} &= z_D, \\ \bar{v}^{k+1}(x, 0) &= y_0(x), & \bar{z}^{k+1}(x, 0) &= z_0(x) \end{aligned}$$

and

$$\begin{aligned} \underline{v}_t^{k+1} - d_y \Delta \underline{v}^{k+1} + \mathbf{V} \cdot \nabla \underline{v}^{k+1} + \widetilde{g}(e^{\eta t} \underline{v}^{k+1}) - e^{-\eta t}(u - M) &= e^{-\eta t}(M - \bar{z}^k), \\ \underline{z}_t^{k+1} - d_z \Delta \underline{z}^{k+1} + \mathbf{V} \cdot \nabla \underline{z}^{k+1} + \varepsilon(\underline{z}^{k+1} - c_3 e^{\eta t} M) &= \varepsilon c_3 e^{\eta t}(M - \bar{v}^k), \\ \partial_n \underline{v}^{k+1} &= 0, & \partial_n \underline{z}^{k+1} &= 0, \\ \underline{v}^{k+1} &= e^{-\eta t} y_D, & \underline{z}^{k+1} &= z_D, \\ \underline{v}^{k+1}(x, 0) &= y_0(x), & \underline{z}^{k+1}(x, 0) &= z_0(x), \end{aligned}$$

respectively. The constructed sequence $\{(\bar{v}^k, \bar{z}^k)\}_{k=0}^\infty$ is monotone non-increasing and upper solution for all k . Conversely $\{(\underline{v}^k, \underline{z}^k)\}_{k=0}^\infty$ is monotone non-decreasing and lower solution for all k . Further, we have

$$\underline{u}^k(x, t) \leq \bar{u}^k(x, t) \quad \text{and} \quad \underline{v}^k(x, t) \leq \bar{v}^k(x, t)$$

for all $k \in \mathbb{N}$ and $(x, t) \in Q_T$.

By induction, we can verify the monotonicity of the sequence $\{\bar{v}^k\}_{k=0}^\infty$: For $k = 0$,

$$\begin{aligned}\bar{v}_t^1 - d_y \Delta \bar{v}^1 + \mathbf{V} \cdot \nabla \bar{v}^1 + \tilde{g}(e^{\eta t} \bar{v}^1) - e^{-\eta t}(u - M) &= e^{-\eta t}(M - \underline{z}^0), \\ \bar{z}_t^1 - d_z \Delta \bar{z}^1 + \mathbf{V} \cdot \nabla \bar{z}^1 + \varepsilon(\bar{z}^1 - c_3 e^{\eta t} M) &= \varepsilon c_3 e^{\eta t}(M - \underline{v}^0), \\ \partial_n \bar{v}^1 &= 0, \quad \partial_n \bar{z}^1 = 0, \\ \bar{v}^1 &= e^{-\eta t} y_D, \quad \bar{z}^1 = z_D, \\ \bar{v}^1(x, 0) &= y_0(x), \quad \bar{z}^1(x, 0) = z_0(x).\end{aligned}$$

The property that $\bar{v}^0$ is an upper solution gives us

$$\begin{aligned}\bar{v}_t^0 - d_y \Delta \bar{v}^0 + \mathbf{V} \cdot \nabla \bar{v}^0 + \tilde{g}(e^{\eta t} \bar{v}^0) - e^{-\eta t}(u - M) &\geq e^{-\eta t}(M - \underline{z}^0), \\ \bar{z}_t^0 - d_z \Delta \bar{z}^0 + \mathbf{V} \cdot \nabla \bar{z}^0 + \varepsilon(\bar{z}^0 - c_3 e^{\eta t} M) &\geq \varepsilon c_3 e^{\eta t}(M - \underline{v}^0), \\ \partial_n \bar{v}^0 &\geq 0, \quad \partial_n \bar{z}^0 \geq 0, \\ \bar{v}^0 &\geq e^{-\eta t} y_D, \quad \bar{z}^0 \geq z_D, \\ \bar{v}^0(x, 0) &\geq y_0(x), \quad \bar{z}^0(x, 0) \geq z_0(x).\end{aligned}$$

Hence, we obtain

$$\begin{aligned}\bar{v}_t^0 - \bar{v}_t^1 - d_y \Delta (\bar{v}^0 - \bar{v}^1) + \mathbf{V} \cdot \nabla (\bar{v}^0 - \bar{v}^1) + \tilde{g}(e^{\eta t} \bar{v}^0) - \tilde{g}(e^{\eta t} \bar{v}^1) &\geq 0, \\ \bar{z}_t^0 - \bar{z}_t^1 - d_z \Delta (\bar{z}^0 - \bar{z}^1) + \mathbf{V} \cdot \nabla (\bar{z}^0 - \bar{z}^1) + \varepsilon(\bar{z}^0 - \bar{z}^1) &\geq 0, \\ \partial_n (\bar{v}^0 - \bar{v}^1) &\geq 0, \quad \partial_n (\bar{z}^0 - \bar{z}^1) \geq 0, \\ (\bar{v}^0 - \bar{v}^1) &\geq 0, \quad (\bar{z}^0 - \bar{z}^1) \geq 0, \\ (\bar{v}^0 - \bar{v}^1)(x, 0) &\geq 0, \quad (\bar{z}^0 - \bar{z}^1)(x, 0) \geq 0.\end{aligned}$$

So it follows from the comparison principle for nonlinear parabolic equations $\bar{v}^0 - \bar{v}^1, \bar{z}^0 - \bar{z}^1 \geq 0$. Now if we assume that $\bar{v}^{k-1} - \bar{v}^k, \bar{z}^{k-1} - \bar{z}^k \geq 0$, one can easily show that $\bar{v}^k - \bar{v}^{k+1}, \bar{z}^k - \bar{z}^{k+1} \geq 0$. Analogously, the monotonicity of $(\underline{v}^k, \underline{z}^k)$ can be proved.

Now, we show the convergence of the sequence $\{\bar{v}^k, \bar{z}^k\}$ to a solution of (2.7). The sequence $\{\bar{v}^k\}$ is monotone non-increasing and bounded from below by $\hat{u} = 0$. Hence by Lebesgue dominated convergence theorem [30,28], it converges to v in space $L^p(Q)$, $p < \infty$. On the other hand, the sequence $\{\bar{z}^k\}$ is monotone non-decreasing and bounded from above by $\tilde{z} = M$. It converges to z in a similar way.

Finally, we prove the uniqueness of the solution of (2.7). Suppose that $(v_1, z_1), (v_2, z_2)$ are solutions of (2.8) and set $v := v_1 - v_2, z := z_1 - z_2$. Then v and z satisfy the initial conditions obviously. Moreover, the following equations

$$\int_0^T (v_t, \varphi)_{V^*, V} dt + \iint_{Q_T} (d_y \nabla v \cdot \nabla \varphi + \mathbf{V} \cdot \nabla v \varphi + \tilde{g}(x, t, v_1) \varphi - \tilde{g}(x, t, v_2) \varphi - e^{\eta t}(u - z) \varphi) dx dt = 0, \quad (2.10a)$$

$$\int_0^T (z_t, \varphi)_{V^*, V} dt + \iint_{Q_T} (d_z \nabla z \cdot \nabla \varphi + \mathbf{V} \cdot \nabla z \varphi + \varepsilon(z - c_3 e^{\eta t} v) \varphi) dx dt = 0 \quad (2.10b)$$

hold for all $\varphi \in W(0, T)$. Then, following the steps in [31] by taking $\varphi = v$ in (2.10a) and $\varphi = z$ in (2.10b) we obtain the desired result $v = 0$ and $z = 0$. $\square$

Hence, we can give the existence of an optimal control u for the optimal control problem (2.5).

Theorem 2.3. *The optimal control problem (2.5) has at least one optimal solution u with associated optimal state y .*

Proof. Here we just sketch the key ideas of the proofs in [29,31, Sec. 5.3, Thm. 7.4]. Since $\mathcal{U}_{ad}$ is non-empty and bounded in $L^\infty(Q_T)$, it is bounded in any space $L^p(Q_T)$ and it follows from the existence and uniqueness of the state variables that they are also bounded. Hence, the cost functional is bounded below, which allows the existence of an infimum. Therefore one can find a weakly convergent minimizing sequence due to the boundedness of this sequence. Then, the compact embedding results give us the strong convergence of the state in a weaker norm. Hence, there exists a feasible limit point, and convergence of the objective function can be shown using the continuity argument. $\square$

Remark 2.4. In this paper, we show the well-posedness of the state equation by introducing upper and lower solutions for the state with monotone nonlinearity, obtained after a suitable transformation. However, there exist other possible reformulations in the literature, for instance, Schauder fixed point theorem applied in [32] for FitzHugh–Nagumo equation, Brouwer fixed point theorem applied in [33] for a class of quasi-linear elliptic problems which are of nonmonotone type, Leray–Schauder fixed-point theorem applied in [34] for a coupled system of semi-linear parabolic reaction–diffusion equations, Faedo–Galerkin method applied in [35] for a hyperbolic quasi-linear hemivariational inequalities.

We continue this section by introducing the necessary and sufficient optimality conditions of the optimal control problem (2.5).

2.2. First order optimality conditions

The OCP

$$\min_{u \in \mathcal{U}_{ad}} J(u) := f(y_u, z_u, u) = I(u) + \mu j(u) \quad (2.11)$$

subject to the convective FHN equation (2.1) is a non-convex programming problem so that different local minima might occur. Since any global solution is one of these local solutions, we set up the first-order optimality conditions satisfied by the local minima.

The cost functional $J(u)$ in (2.11) consists of two terms with different smoothness. While the first part $I(u)$ is smooth, the second part $j(u) : L^1(Q) \rightarrow \mathbb{R}$ is not differentiable, but it is subdifferentiable and the directional derivative is given by

$$j'(u, v - u) = \max_{\lambda \in \partial j(u)} \langle \lambda, v - u \rangle \quad (2.12)$$

with

$$\partial j(u) = \left\{ \lambda \in L^\infty(Q_T) : j(v) \geq j(u) + \int_0^T \int_\Omega \lambda(v - u) \, dx \, dt \quad \forall v \in L^\infty(Q_T) \right\},$$

where

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

We note that the relations above are required only almost everywhere. With the help of the Lagrangian functional, we obtain the following variational inequality:

$$I'(u)(v - u) + \mu j'(u, v - u) \geq 0 \quad \forall v \in \mathcal{U}_{ad}, \quad (2.13)$$

that is,

$$\int_0^T \int_\Omega (p(x, t) + \omega_u u(x, t) + \mu \lambda(x, t))(v(x, t) - u(x, t)) \, dx \, dt \geq 0 \quad \forall v \in \mathcal{U}_{ad},$$

where $p(x, t)$ with $q(x, t)$ are called the adjoint variables as the solution of the following system

$$-p_t - d_y \Delta p - \mathbf{V} \cdot \nabla p + g_y(y)p - \varepsilon c_3 q = \omega_Q^y (y - y_Q) \quad \text{in } Q_T, \quad (2.14a)$$

$$-q_t - d_z \Delta q - \mathbf{V} \cdot \nabla q + \varepsilon q + p = \omega_Q^z (z - z_Q) \quad \text{in } Q_T, \quad (2.14b)$$

with the mixed boundary conditions

$$d_y \partial_n p(x, t) + (\mathbf{V} \cdot \mathbf{n})p(x, t) = 0, \quad d_z \partial_n q(x, t) + (\mathbf{V} \cdot \mathbf{n})q(x, t) = 0 \quad \text{on } \sum_T^N, \quad (2.14c)$$

the Dirichlet boundary conditions

$$p(x, t) = 0, \quad q(x, t) = 0 \quad \text{on } \sum_T^D, \quad (2.14d)$$

and final time conditions

$$p(x, T) = \omega_T^y (y(x, T) - y_T(x)), \quad q(x, T) = \omega_T^z (z(x, T) - z_T(x)) \quad \text{in } \Omega. \quad (2.14e)$$

The convection term in the adjoint system (2.14) is the negative of the one in the FHN equation (2.1). As a consequence, errors in the solution can potentially propagate in both directions. Therefore, the numerical treatment of the state and adjoint systems together is more delicate.

For $\omega_u > 0$ and $\mu > 0$, from the variational inequality (2.13) the following projection formulas are obtained [8,9]

$$u(x, t) = \mathbb{P}_{[u_a, u_b]} \left\{ -\frac{1}{\omega_u} (p(x, t) + \mu \lambda(x, t)) \right\} \quad \text{for a.a. } (x, t) \in Q_T, \quad (2.15)$$

$$\lambda(x, t) = \mathbb{P}_{[-1, 1]} \left\{ -\frac{1}{\mu} p(x, t) \right\} \quad \text{for a.a. } (x, t) \in Q_T, \quad (2.16)$$

where the projection operator $\mathbb{P}_{[a, b]} : \mathbb{R} \rightarrow [a, b]$ is defined by

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

Further, the following relation holds for almost all $(x, t) \in Q_T$

$$u(x, t) = 0 \Leftrightarrow \begin{cases} |p(x, t)| \leq \mu, & \text{if } u_a < 0, \\ p(x, t) \geq -\mu, & \text{if } u_a = 0. \end{cases} \quad (2.17)$$

We refer to [13,7,29] for a further discussion on the projection operator and a derivation of (2.16) and (2.17).

2.3. Second order optimality conditions

Due to the nonlinearity of the state equation (2.1), the optimization problem is non-convex. Therefore, the fulfillment of the first necessary conditions does not imply optimality. In order to guarantee the optimality, the second-order sufficient optimality conditions (SSCs) have to be satisfied. The SSCs are related to certain critical cones that must be chosen as small as possible.

Now, we give the critical cone related to our optimization problem for $\omega_u > 0$ derived in [8]:

$$C_u = \left\{ v \in L^2(Q_T) : v \text{ satisfies the sign condition and } I'(u)v + \mu j'(u) = 0 \right\}$$

with the sign condition:

$$\begin{cases} v(x, t) \geq 0, & \text{if } u(x, t) = u_a, \\ v(x, t) \leq 0, & \text{if } u(x, t) = u_b. \end{cases}$$

The set C_u is a convex and closed cone in $L^2(Q_T)$. Moreover, if u is a local minima for (P_μ) , then the following inequalities hold [8, Theorem 3.3],

$$I''(u)v^2 \geq 0 \quad \forall v \in C_u \setminus \{0\}, \quad \text{equivalently } I''(u)v^2 \geq \delta \|v\|_{L^2(Q_T)}^2 \quad \forall v \in C_u, \quad \delta > 0. \quad (2.18)$$

Then, under the assumption $I''(u)v^2 \geq 0 \forall v \in C_u \setminus \{0\}$, there exist $\delta > 0$ and $r_0 > 0$ such that

$$J(u) + \frac{\delta}{2} \|v - u\|_{L^2(Q_T)}^2 \leq J(v) \quad \forall u \in U_{ad} \cap B_{r_0}(u), \quad (2.19)$$

where $B_{r_0}(u)$ is the $L^2(Q_T)$ ball centered at u with radius r_0 . This shows the existence of a local minima, see [8, Theorem 3.4] for details.

The verification of the SSCs is difficult because the solution of the infinite dimensional problem is required. Even if it is known, it would still be tedious to check that the SSC holds, since it requires the exact solution of linearized PDEs. However, there exist some numerical studies on the SSCs, see e.g., [26,36,24]. Here we determine the constant δ by computing the smallest eigenvalue of the reduced Hessian as introduced in [36] of the finite dimensional dG discretized OCP.

Now, we can state the following theorem to measure the distance between the local minima u and the local discrete solution u_h obtained by applying the discontinuous Galerkin discretization for spatial discretization and the backward Euler for temporal discretization in Section 3.

Theorem 2.5. Let u be a local minima of (2.5). Assume that u satisfies the second-order sufficient condition (2.18) and (2.19). If u_h is the discrete solution such that $u_h \in B_{r_0}(u)$, then it holds

$$\|u - u_h\|_{L^2(Q_T)} \leq \frac{1}{\delta'} \|\zeta\|_{L^2(Q_T)}, \quad (2.20)$$

where a perturbation function ζ is defined as follows

$$\zeta(x) := \begin{cases} -\min\{0, \omega_u u_h + p_h + \mu \lambda_h\}, & \text{if } u_h = u_a, \\ -(\omega_u u_h + p_h + \mu \lambda_h), & \text{if } u_a < u_h < u_b, \\ -\max\{0, \omega_u u_h + p_h + \mu \lambda_h\}, & \text{if } u_h = u_b \end{cases}$$

and $0 < \delta' < \delta$.

Proof. Let u_h be a discrete solution that need not be optimal for the continuous problem (2.5), and let p_h and λ_h be the associated adjoint and sparse. If u_h were optimal, then $p_h + \omega_u u_h + \mu \lambda_h = 0$ should be satisfied in almost all points $x \in \Omega$, where $u_a \leq u_h \leq u_b$ holds. If not, then $p_h + \omega_u u_h + \mu \lambda_h + \zeta = 0$, where ζ is a perturbation function, adopted from [37]. Although u_h is possibly not be optimal for (2.5), it is optimal for the perturbed optimization problem

$$\min_{u \in \mathcal{U}_{ad}} J(u) + (\zeta, u)_{L^2(Q_T)}.$$

Inserting u in the discrete variational form and u_h in the continuous discrete variational form, we obtain

$$(J'(u_h) + \zeta, u - u_h) \geq 0, \quad (2.21a)$$

$$(J'(u) + \zeta, u_h - u) \geq 0. \quad (2.21b)$$

Adding these equations in (2.21) gives us

$$(J'(u_h) - J'(u), u - u_h) + (\zeta, u - u_h) \geq 0. \quad (2.22)$$

By the mean value theorem, we obtain

$$-J''(\hat{u})(u - u_h)^2 + (\zeta, u - u_h) \geq 0 \quad (2.23)$$

for some $\hat{u} \in \{v \in U_{ad} : v = u + t(u_h - u), t \in (0, 1)\}$. Then, when we apply the SSC (2.18) with Cauchy–Schwarz inequality, we find

$$\delta' \|u_h - u\|_{L^2(Q_T)}^2 \leq \|\zeta\|_{L^2(Q_T)} \|u_h - u\|_{L^2(Q_T)}, \quad (2.24)$$

which is the desired result. $\square$

Here we follow [26, Remark 3.3] and select $\delta' := \delta/2$. If u_h belongs to the neighborhood of u , we can estimate in the following

$$\|u - u_h\|_{L^2(Q_T)} \leq \frac{2}{\delta'} \|\zeta\|_{L^2(Q_T)}. \quad (2.25)$$

Remark 2.6. We also remark that the presence of the so-called Tikhonov parameter $\frac{\omega_u}{2}$ in the cost functional (2.5) is extremely important. The standard second-order optimality conditions do not hold for vanishing Tikhonov parameter $\omega_u = 0$. By introducing new critical cones, the second-order optimality conditions are established in [8] for sparse optimal control governed by FitzHugh–Nagumo system, and in [38] for general nonlinear functions. Hence, the new second order conditions are used for proving the stability of locally optimal solutions with respect to $\omega_u \rightarrow 0$. This theoretical result was confirmed for one dimensional wave solutions of the classical FHN equation with sparse controls in [8,9]. In this paper, we apply the theory introduced in [8,38] for two dimensional wave solutions of the convective FHN equations.

3. Discretization in space and time

We discretize the state system (2.1) and the adjoint system (2.14) by the symmetric interior penalty Galerkin (SIPG) in space and the backward Euler discretization in time.

3.1. SIPG discretization of the state and adjoint equations in space

The interior penalty dG methods are well established in the literature and the details can be found in the classical texts like [39,40].

We assume that the domain Ω is polygonal domain. We denote $\{\mathcal{T}_h\}_h$ as a family of shape-regular simplicial triangulations of Ω , see, e.g., [41]. Each mesh $\mathcal{T}_h$ consists of closed triangles such that $\bar{\Omega} = \bigcup_{K \in \mathcal{T}_h} \bar{K}$ holds. We assume that the mesh is regular in the following sense: for different triangles $K_i, K_j \in \mathcal{T}_h, i \neq j$, the intersection $K_i \cap K_j$ is either empty or a vertex or an edge, i.e., hanging nodes are not allowed. The diameter of an element K and the length of an edge E are denoted by h_K and h_E , respectively.

We split the set of all edges $\mathcal{E}_h$ into the set $\mathcal{E}_h^0$ of interior edges, the set $\mathcal{E}_h^D$ of Dirichlet boundary edges and the set $\mathcal{E}_h^N$ of Neumann boundary edges so that $\mathcal{E}_h = \mathcal{E}_h^B \cup \mathcal{E}_h^0$ with $\mathcal{E}_h^B = \mathcal{E}_h^D \cup \mathcal{E}_h^N$. Let $\mathbf{n}$ denote the unit outward normal to $\partial\Omega$. For the activator y and the inhibitor z , we define the inflow and outflow parts of $\partial\Omega$ by

$$\Gamma^- = \{x \in \partial\Omega : \mathbf{V}(x) \cdot \mathbf{n}(x) < 0\}, \quad \Gamma^+ = \{x \in \partial\Omega : \mathbf{V}(x) \cdot \mathbf{n}(x) \geq 0\}.$$

Similarly, the inflow and outflow boundaries of an element K are defined by

$$\partial K^- = \{x \in \partial K : \mathbf{V}(x) \cdot \mathbf{n}_K(x) < 0\}, \quad \partial K^+ = \{x \in \partial K : \mathbf{V}(x) \cdot \mathbf{n}_K(x) \geq 0\},$$

where $\mathbf{n}_K$ is the unit normal vector on the boundary ∂K of an element K .

Let the edge E be a common edge for two elements K and K^e . For a piecewise continuous scalar function y , there are two traces of y along E , denoted by $y|_E$ from inside K and $y^e|_E$ from inside K^e . Then, the jump and average of y across the edge E are defined by:

$$\llbracket y \rrbracket = y|_E \mathbf{n}_K + y^e|_E \mathbf{n}_{K^e}, \quad \{y\} = \frac{1}{2}(y|_E + y^e|_E). \quad (3.1)$$

Similarly, for a piecewise continuous vector field ∇y , the jump and average across an edge E are given by

$$\llbracket \nabla y \rrbracket = \nabla y|_E \cdot \mathbf{n}_K + \nabla y^e|_E \cdot \mathbf{n}_{K^e}, \quad \{\{\nabla y\}\} = \frac{1}{2}(\nabla y|_E + \nabla y^e|_E). \quad (3.2)$$

For a boundary edge $E \in K \cap \partial\Omega$, we set $\{\{y\}\} = y$ and $\llbracket y \rrbracket = y\mathbf{n}$.

Recall that for dG methods, we do not impose continuity constraints on the trial and test functions across the element interfaces. As a consequence, the weak formulation must include jump terms across interfaces, and typically penalty terms are added to control the jump terms. Then, we define the discontinuous discrete space as follows:

$$W_h = \{w \in L^2(\Omega) : w|_K \in \mathbb{P}^1(K), \forall K \in \mathcal{T}_h\}, \quad (3.3)$$

where $\mathbb{P}^1(K)$ is the set of piecewise linear polynomials defined on K . We note that the space of discrete states and the space of test functions are identical due to the weak treatment of boundary conditions for dG methods.

Then, the semi-discrete formulation of the state system (2.1) for $\forall w \in W_h$ and $t \in (0, T]$ becomes

$$\left(\frac{dy_h}{dt}, w\right) + a_{h,y}(y_h, w) + b_{h,y}(y_h, w) + c_{h,z}(z_h, w) = \ell_{h,y}(w) + (u_h, w), \quad (3.4a)$$

$$(y_h(\cdot, 0), w) = (y_0, w), \quad (3.4b)$$

$$\left(\frac{dz_h}{dt}, w\right) + a_{h,z}(z_h, w) + b_{h,z}(z_h, w) + c_{h,y}(y_h, w) = \ell_{h,z}(w), \quad (3.4c)$$

$$(z_h(\cdot, 0), w) = (z_0, w), \quad (3.4d)$$

where the (bi)-linear terms are defined for $i = y, z$ and $\forall w \in W_h$

$$\begin{aligned} a_{h,i}(v, w) &= \sum_{K \in \mathcal{T}_h} \int_K d_i \nabla v \cdot \nabla w \, dx - \sum_{E \in \mathcal{E}_h^0 \cup \mathcal{E}_h^D} \int_E \left(\{\{d_i \nabla v\}\} \cdot \llbracket w \rrbracket + \{\{d_i \nabla w\}\} \cdot \llbracket v \rrbracket \right) ds \\ &\quad + \sum_{E \in \mathcal{E}_h^0 \cup \mathcal{E}_h^D} \frac{\sigma d_i}{h_E} \int_E \llbracket v \rrbracket \cdot \llbracket w \rrbracket \, ds + \sum_{K \in \mathcal{T}_h} \int_K \mathbf{v} \cdot \nabla v w \, dx \\ &\quad + \sum_{K \in \mathcal{T}_h} \int_{\partial K^- \setminus \partial\Omega} \mathbf{v} \cdot \mathbf{n} (v^e - v) w \, ds - \sum_{K \in \mathcal{T}_h} \int_{\partial K^- \cap \Gamma^-} \mathbf{v} \cdot \mathbf{n} v w \, ds, \end{aligned} \quad (3.5a)$$

$$\ell_{h,i}(w) = \sum_{E \in \mathcal{E}_h^D} \int_E i_D \left(\frac{\sigma d_i}{h_E} \mathbf{n} \cdot \llbracket w \rrbracket - \{\{d_i \nabla w\}\} \right) ds - \sum_{K \in \mathcal{T}_h} \int_{\partial K^- \cap \Gamma^-} \mathbf{v} \cdot \mathbf{n} i_D w \, ds, \quad (3.5b)$$

$$b_{h,y}(y, w) = \sum_{K \in \mathcal{T}_h} \int_K g(y) w \, dx, \quad c_{h,y}(y, w) = \sum_{K \in \mathcal{T}_h} \int_K -\varepsilon c_3 y w \, dx, \quad (3.5c)$$

$$b_{h,z}(z, w) = \sum_{K \in \mathcal{T}_h} \int_K \varepsilon z w \, dx, \quad c_{h,z}(z, w) = \sum_{K \in \mathcal{T}_h} \int_K z w \, dx, \quad (3.5d)$$

where the parameter $\sigma \in \mathbb{R}_0^+$ is called the penalty parameter which should be sufficiently large to ensure the stability of the dG discretization; independent of the mesh size h and of the diffusion coefficients d_i as described in [40, Sec. 2.7.1] with a lower bound depending only on the polynomial degree. Large penalty parameters decrease the jumps across element interfaces, which can affect the numerical approximation. However, the dG approximation can converge to the continuous Galerkin approximation as the penalty parameter goes to infinity (see, e.g., [42] for details).

For each time step, we can expand the discrete solutions of the states y, z , and the control u as

$$y_h(t) = \sum_{i=1}^N \sum_{j=1}^{n_k} \tilde{y}_j^i \phi_j^i, \quad z_h(t) = \sum_{i=1}^N \sum_{j=1}^{n_k} \tilde{z}_j^i \phi_j^i, \quad \text{and} \quad u_h(t) = \sum_{i=1}^N \sum_{j=1}^{n_k} \tilde{u}_j^i \phi_j^i, \quad (3.6)$$

where $\tilde{y}_j^i, \tilde{z}_j^i, \tilde{u}_j^i$, and ϕ_j^i are the unknown coefficients and the basis functions, respectively, for $j = 1, 2, \dots, n_k$ and $i = 1, 2, \dots, N$. The number N denotes the number of dG elements and $n_k = (k+1)(k+2)/2$ is the local dimension of each dG element depending on the order k of the polynomial basis.

Inserting (3.6) into (3.4), we obtain

$$\mathbf{M} \frac{d\tilde{\mathbf{y}}}{dt} + \mathbf{A}_y \tilde{\mathbf{y}} + \mathbf{b}_y(\tilde{\mathbf{y}}) + \mathbf{C}_z \tilde{\mathbf{z}} = \mathbf{I}_y + \mathbf{M} \tilde{\mathbf{u}}, \quad (3.7a)$$

$$\mathbf{M} \frac{d\tilde{\mathbf{z}}}{dt} + \mathbf{A}_z \tilde{\mathbf{z}} + \mathbf{B}_z \tilde{\mathbf{z}} + \mathbf{C}_y \tilde{\mathbf{y}} = \mathbf{I}_z, \quad (3.7b)$$

where $\vec{y}$, $\vec{z}$ and $\vec{u}$ are the unknown coefficient vectors, i.e.,

$$\vec{y} = (\vec{y}_1^1, \dots, \vec{y}_{n_k}^1, \dots, \vec{y}_1^N, \dots, \vec{y}_{n_k}^N),$$

$$\vec{z} = (\vec{z}_1^1, \dots, \vec{z}_{n_k}^1, \dots, \vec{z}_1^N, \dots, \vec{z}_{n_k}^N),$$

$$\vec{u} = (\vec{u}_1^1, \dots, \vec{u}_{n_k}^1, \dots, \vec{u}_1^N, \dots, \vec{u}_{n_k}^N),$$

$\mathbf{M}$ is the mass matrix, $\mathbf{A}_y$ and $\mathbf{A}_z$ are the stiffness matrices corresponding to $a_h(y_h, w)$ and $a_h(z_h, w)$, respectively, and $\mathbf{l}_y$ and $\mathbf{l}_z$ are vectors corresponding to $\ell_{h,y}(w)$ and $\ell_{h,z}(w)$, respectively. Moreover, $\mathbf{C}_y$, $\mathbf{B}_z$, and $\mathbf{C}_z$ are the matrices corresponding to $c_{h,y}$, $b_{h,z}$, and $c_{h,z}$, respectively. Further, $\mathbf{b}_y(\vec{y})$ is the vector corresponding to the nonlinear part $b_{h,y}(y, w)$.

In the literature, there exist two approaches for the solution the OCP (2.5) numerically, i.e., *discretize-then-optimize* (DO) and *optimize-then-discretize* (OD). In the DO approach, the objective function (2.5) and the state equation (2.1) are discretized first, and then the discrete optimality system is formed. In the OD approach, the optimality conditions consisting of the state system (2.1), the adjoint system (2.14) and the variational form (2.13) are derived. Then, the infinite dimensional optimality system is discretized. We here apply the *optimize-then-discretize* approach. The discretization and optimization commute for the SIPG discretized linear steady state convection–diffusion–reaction OCPs [43,17]. For linear time dependent OPCs the OD and DO approaches commute for the dG discretization in time and the SIPG discretization in space [18]. The commutativity property is no more valid for semi-linear elliptic and parabolic OCPs.

In the *optimize-then-discretize* approach, we derive directly semi-discrete form of the adjoint equations (2.14) using the SIPG discretization as for the state equation (3.4):

$$\left(-\frac{dp_h}{dt}, w\right) + a_{h,p}(p_h, w) + b_{h,p}(p_h, w) + c_{h,q}(q_h, w) = \omega_Q^y(y_h, w) + \ell_{h,p}(w), \quad (3.8a)$$

$$(p_h(\cdot, T), w) = \omega_T^y((y_h(\cdot, T) - y_T(x)), w), \quad (3.8b)$$

$$\left(-\frac{dq_h}{dt}, w\right) + a_{h,q}(q_h, w) + b_{h,q}(q_h, w) + c_{h,p}(p_h, w) = \omega_Q^z(z_h, w) + \ell_{h,q}(w), \quad (3.8c)$$

$$(q_h(\cdot, T), w) = \omega_T^z((z_h(\cdot, T) - z_T(x)), w), \quad (3.8d)$$

where the bilinear forms $a_{h,p}$ and $a_{h,q}$ are similar to the state ones (3.5) with the negative convection terms, i.e., $-\mathbf{V}(x, y)$. As an extra term, they just contain the contribution of the mixed boundary conditions, i.e.,

$$\sum_{E \in \mathcal{E}_h^N} \int_E (\mathbf{v} \cdot \mathbf{n}) p_h w \, ds \quad \text{and} \quad \sum_{E \in \mathcal{E}_h^N} \int_E (\mathbf{v} \cdot \mathbf{n}) q_h w \, ds.$$

The other terms are given by

$$b_{h,p}(p, w) = \sum_{K \in \mathcal{T}_h} \int_K g_y(y) p w \, dx, \quad c_{h,q}(q, w) = \sum_{K \in \mathcal{T}_h} \int_K -\varepsilon c_3 q w \, dx, \quad (3.9a)$$

$$b_{h,q}(q, w) = \sum_{K \in \mathcal{T}_h} \int_K \varepsilon q w \, dx, \quad c_{h,p}(p, w) = \sum_{K \in \mathcal{T}_h} \int_K p w \, dx, \quad (3.9b)$$

$$\ell_{h,p}(w) = -\omega_Q^y \sum_{K \in \mathcal{T}_h} \int_K y_Q w \, dx, \quad \ell_{h,q}(w) = -\omega_Q^z \sum_{K \in \mathcal{T}_h} \int_K z_Q w \, dx. \quad (3.9c)$$

3.2. Time discretization

We solve the nonlinear system of ordinary differential equation (3.7) in time by the backward Euler method. We first divide the time interval $[0, T]$ $0 = t_0 < t_1 < \dots < t_{N_T} = T$, with step size $\tau_n = t_n - t_{n-1}$, $n = 1, 2, \dots, N_T$ and then obtain the following fully discretized state system:

$$\frac{1}{\tau_n} \mathbf{M}(\vec{y}_n - \vec{y}_{n-1}) + \mathbf{A}_y \vec{y}_n + \mathbf{b}_y(\vec{y}_n) + \mathbf{C}_z \vec{z}_n = \mathbf{l}_y^n + \mathbf{M} \vec{u}_n, \quad (3.10a)$$

$$\frac{1}{\tau_n} \mathbf{M}(\vec{z}_n - \vec{z}_{n-1}) + \mathbf{A}_z \vec{z}_n + \mathbf{B}_z \vec{z}_n + \mathbf{C}_z \vec{y}_n = \mathbf{l}_z^n \quad (3.10b)$$

for $n = 1, 2, \dots, N_T$. We solve the nonlinear system by Newton's method.

We discretize the adjoint variables p and q analogously to the state variables y , and z as

$$p_h(t) = \sum_{i=1}^N \sum_{j=1}^{n_k} \tilde{p}_j^i \phi_j^i \quad \text{and} \quad q_h(t) = \sum_{i=1}^N \sum_{j=1}^{n_k} \tilde{q}_j^i \phi_j^i. \quad (3.11)$$

By inserting (3.11) into (3.8) and applying the backward Euler discretization in time, we obtain the following fully discretized adjoint system:

$$\frac{1}{\tau_n} \mathbf{M}(\bar{p}_{n-1} - \bar{p}_n) + \mathbf{A}_p \bar{p}_{n-1} + \mathbf{B}_p(\bar{y}_n) \bar{p}_{n-1} + \mathbf{C}_q \bar{q}_{n-1} = \omega_Q^y (\mathbf{M} \bar{y}_n + \mathbf{I}_p^n), \quad (3.12a)$$

$$\frac{1}{\tau_n} \mathbf{M}(\bar{q}_{n-1} - \bar{q}_n) + \mathbf{A}_q \bar{q}_{n-1} + \mathbf{B}_q \bar{q}_{n-1} + \mathbf{C}_p \bar{p}_{n-1} = \omega_Q^z (\mathbf{M} \bar{z}_n + \mathbf{I}_q^n) \quad (3.12b)$$

for $n = N_T, \dots, 2, 1$. The matrices $\mathbf{A}_p, \mathbf{A}_q$ correspond to the discretized bilinear forms $a_{h,p}$ and $a_{h,q}$, respectively, whereas the matrices $\mathbf{B}_p(y), \mathbf{C}_p, \mathbf{B}_q, \mathbf{C}_q$ and the vectors $\mathbf{I}_p, \mathbf{I}_q$ are the discretized forms in (3.9), respectively.

4. Numerical results

In this section, we provide numerical results for the OCP governed by the convective FHN equation (2.5). There exists several optimization algorithms for OCPs governed by semi-linear equations. We have used in this paper the projected nonlinear conjugate gradient (CG) method [44,22], which was applied to the Schlögl and FHN equations [6–8]. The projected nonlinear CG algorithm is outlined below:

Initialization Choose an initial guess u_0 , an initial step size s_0 and stopping tolerances ToI_1 and ToI_2 . Then, compute

- initial states $(y_0, z_0) = (y_{u_0}, z_{u_0})$ by solving (3.10),
- initial adjoints $(p_0, q_0) = (p_{y_0, z_0}, q_{y_0, z_0})$ by solving (3.12),
- subgradient of j by (2.16), i.e., $\lambda_0 = \lambda_{u_0, p_0}$,
- subgradient of J , i.e., $g_0 = \omega_c u_0 + p_0 + \mu \lambda_0$,
- anti-subgradient of J , i.e., $d_0 = -g_0$.

Set $k := 0$.

Step 1. (New subgradients) Update

- control, i.e., $u_{k+1} = u_k + s_k d_k$,
- states $(y_{k+1}, z_{k+1}) = (y_{u_{k+1}}, z_{u_{k+1}})$ by solving (3.10),
- adjoints $(p_{k+1}, q_{k+1}) = (p_{y_{k+1}, z_{k+1}}, q_{y_{k+1}, z_{k+1}})$ by solving (3.12),
- subgradient of j by (2.16), i.e., $\lambda_{k+1} = \lambda_{u_{k+1}, p_{k+1}}$,
- subgradient of J , i.e., $g_{k+1} = \omega_c u_{k+1} + p_{k+1} + \mu \lambda_{k+1}$.

Step 2. (Stopping Criteria) Stop if $\|g_{k+1}\| < \text{ToI}_1$ or $\|J_{k+1} - J_k\| \leq \text{ToI}_2$.

Step 3. (Direction of descent) Compute the conjugate direction β_{k+1} according to one of the update formulas such as Hestenes–Stiefel, Polak–Ribiere, Fletcher–Reeves, and Hager–Zhang, see e.g., [6,22] for details.

$$d_{k+1} = -g_{k+1} + \beta_{k+1} d_k.$$

Step 4. (Step size) Select step size s_{k+1} according to some standard options such as bisection, strong Wolfed–Powell, see e.g., [6,44] for details. Set $k := k + 1$ and go to Step 1.

For the computation of the reduced Hessian, we use the BFGS algorithm [45,46]:

- Set $H_0 = I$.
- Update for $k = 1, 2, \dots$

$$H_{k+1} = H_k + \frac{q_k q_k^T}{q_k^T r_k} - \frac{(H_k r_k)(H_k r_k)^T}{r_k^T H_k r_k},$$

where $r_k = u_{k+1} - u_k$ and $q_k = g_{k+1} - g_k$.

After setting the reduced Hessian matrix at each time level, we compute the smallest eigenvalue of the Hessian by using MATLAB® eig.

In all numerical experiments, we started with the initial control $u = 0$. The optimization algorithm is terminated, when the gradient is less than $\text{ToI}_1 = 10^{-3}$ or the difference between the successive cost functionals is less than $\text{ToI}_2 = 10^{-5}$. We use uniform step sizes in time $\Delta t = 0.05$. The penalty parameter σ in (3.5) is chosen as $\sigma = 6$ on the interior edges and $\sigma = 12$ on the boundary edges.

If it is not specified, we use the following parameters in all numerical examples

$$c_1 = 9, \quad c_2 = 0.02, \quad c_3 = 5, \quad \varepsilon = 0.1, \quad d_1 = d_2 = 1$$

on a rectangular box $\Omega = [0, L] \times [0, H]$ with $L = 100$ and $H = 5$ as in [4]. Further, the final time is taken as $T = 1$.

All simulations are generated on Intel(R) Core(TM) i7-4720HQ CPU@2.60 GHz, 16 GB RAM, Windows 8, with MATLAB® R2014a (64-bit).

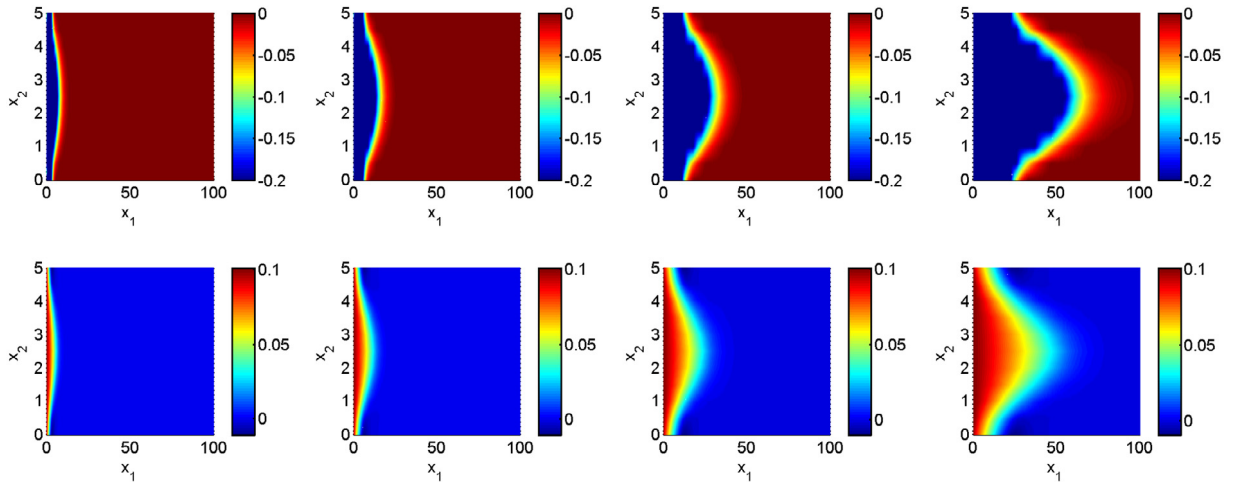


Fig. 1. Example 4.1: optimal controls u (top) and associated states y (bottom) at $t = 0.75$ without sparse control for $V_{\max} = 16, 32, 64, 128$ (from left to right).

Table 1

Example 4.1: cost functional J , number of nonlinear CG iterations, line searches, Newton steps without sparse control.

$V_{\max}$	J	#ite.	#search	#Newton
16	2.91e−2	64	245	787
32	5.80e−2	77	297	956
64	1.16e−1	96	373	1203
128	2.30e−1	123	481	1554

4.1. Optimal control in the space–time domain

We first consider the OCP with the desired state functions defined in the whole space–time domain Q_T with $\omega_Q^y = \omega_Q^z = 1$, $\omega_T^y = \omega_T^z = 0$, and $\omega_u = 10^{-5}$. The step sizes in space are taken as $\Delta x_1 = \Delta x_2 = 0.5$. The desired states are chosen as the solution of uncontrolled FHN equation (2.1)

$$y_Q(x, t) = \begin{cases} y_{\text{nat}}(x, t), & \text{if } t \leq T/2, \\ 0, & \text{otherwise,} \end{cases} \quad z_Q(x, t) = \begin{cases} z_{\text{nat}}(x, t), & \text{if } t \leq T/2, \\ 0, & \text{otherwise,} \end{cases}$$

with the initial conditions given by

$$y_0(x, t) = \begin{cases} 0.1, & \text{if } 0 \leq x_1 \leq 0.1, \\ 0, & \text{otherwise,} \end{cases} \quad z_0(x, 0) = 0.$$

Further, we define the admissible set of controls as

$$\mathcal{U}_{\text{ad}} := \{u \in L^\infty(Q) : -0.2 \leq u(x, t) \leq 0 \text{ for a.e. } (x, t) \in Q\}.$$

We first investigate the numerical solutions of the optimization problem (2.5) without the sparse parameter, i.e., $\mu = 0$. Fig. 1 demonstrates the computed solutions of the control u and their associated states y at $t = 0.75$ for various values of the $V_{\max}$, respectively. The control u , bounded by the box constraints, exhibits the same behavior of the state y . When the value of $V_{\max}$ increases, the controlled solutions become more curved, as for the uncontrolled solutions [4]. Optimal values of cost functional J , and the number of iterations, line searches, and Newton steps in Table 1 are increasing for the higher values of $V_{\max}$ because the linear system of equations to be solved become more stiff due to convection dominated character of the OCP problem.

One of the important features of the optimal control is the robustness of the control with respect to the parameters. However, in the PDE-constrained optimization context, we do not always have an explicit control function as in [47,48]. In that case, we check the robustness of the control by solving the optimal control problem for different values of the parameters. For this example, we study the effect of the maximum velocity $V_{\max}$ on the controlled wave solutions and of the parameter c_1 on the stability of the waves. As $V_{\max}$ increases, the wave profiles are elongated along x_1 directions and control is adjusted accordingly by shifting to the right in Fig. 2. Fig. 3 shows that the solutions display oscillations for higher values of c_1 as the uncontrolled solutions in [4]. Since the control variable stays inside the bounds of the control, we can conclude that the control variable is robust with respect to variations of the system parameters.

Now, we look into the effect of the sparse parameter μ on the optimization problem (2.5). Higher values of μ cause the sparsity of the optimal control, which can be clearly seen from Figs. 4 and 5. Fig. 5 shows profiles of the controls u

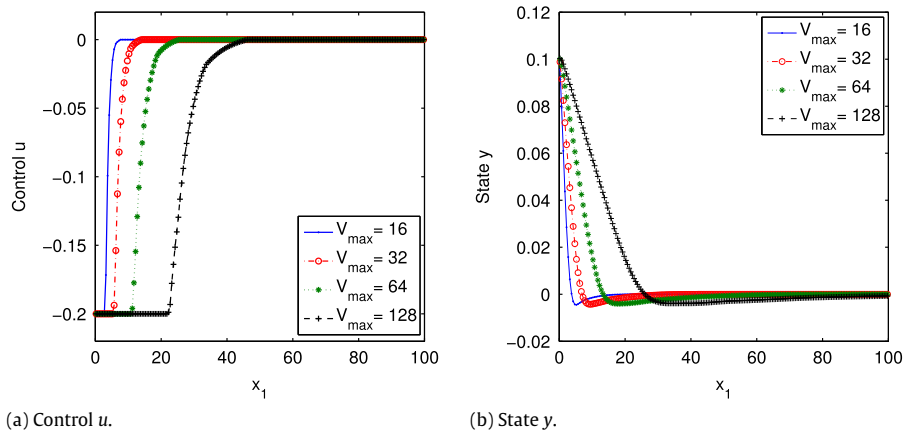


Fig. 2. Example 4.1: profiles of optimal controls u and associated states y along x_1 axis at $t = 0.75$ without sparse control for various values of $V_{\max}$.

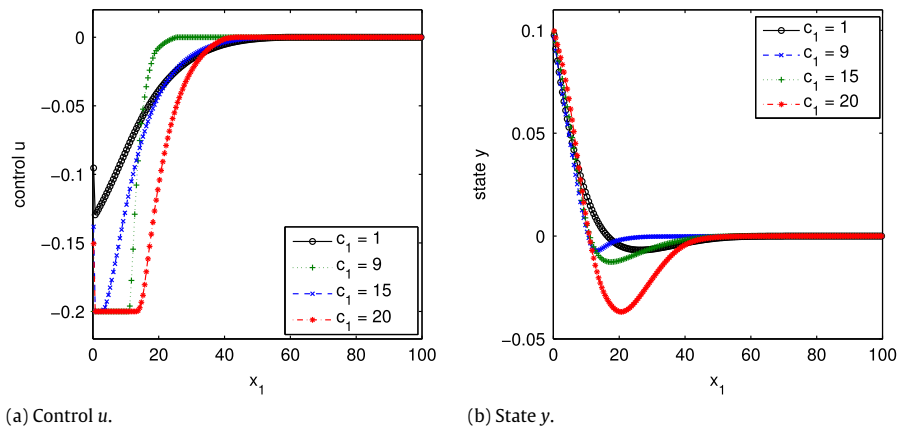


Fig. 3. Example 4.1: profiles of optimal controls u and associated states y along x_1 axis for $V_{\max} = 64$ without sparse control for various values of c_1 .

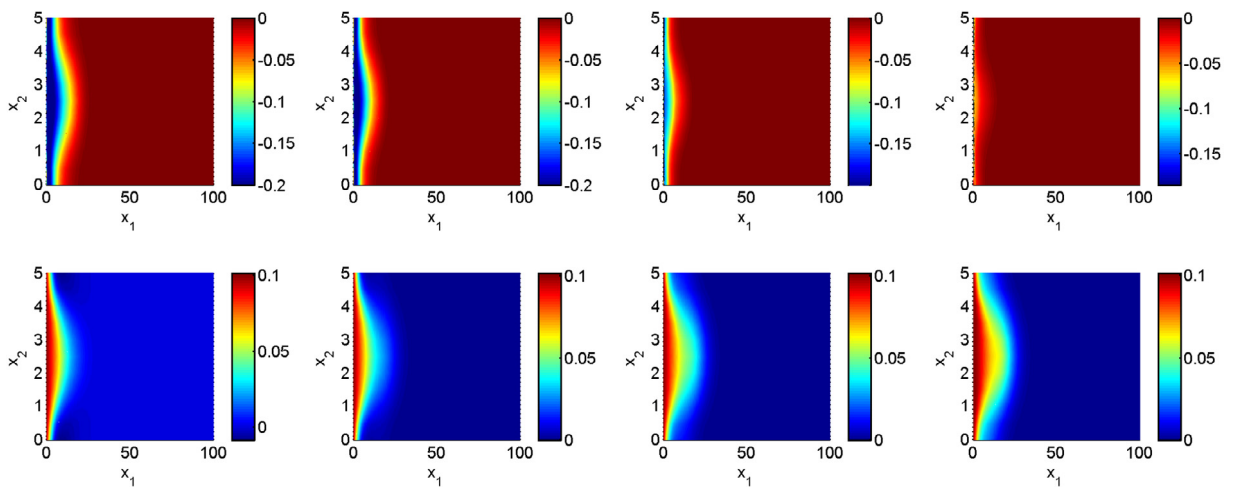


Fig. 4. Example 4.1: optimal controls u (top) and associated states y (bottom) for the sparse parameters $\mu = 1/500, 1/100, 1/50, 1/35$ (from left to right) and $V_{\max} = 32$ at $t = 0.75$.

and associated states y along x_1 direction for various values of μ . The values of the optimal cost functional and number of iterations with line searches and Newton steps are given in Table 2. When the sparse parameter μ increases, the smooth and non-smooth cost functionals increase as well as the number nonlinear CG iterations, line searches and Newton iterations. However, we will show in the next example with terminal control, the number of nonlinear CG iterations, line searches and

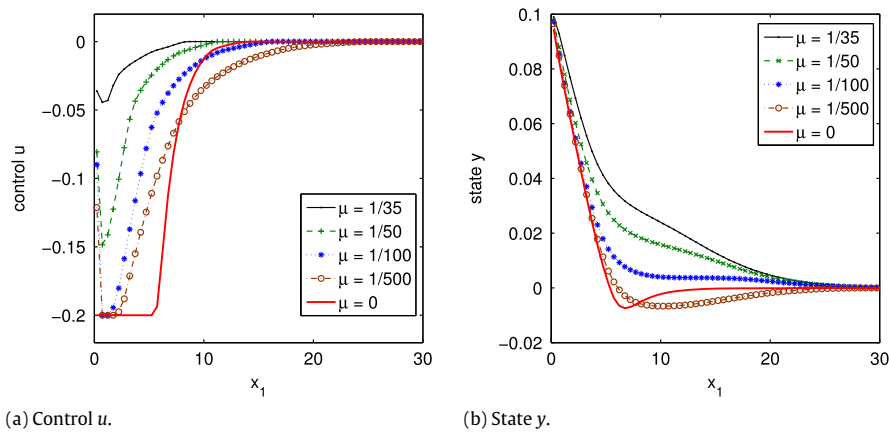


Fig. 5. Example 4.1: profiles of optimal controls u and associated states y at $t = 0.75$ with $V_{\max} = 32$ for various values of μ .

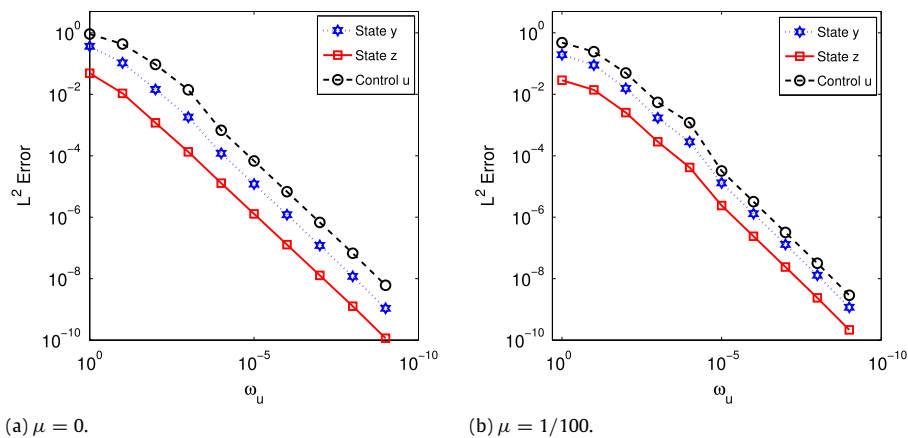


Fig. 6. Example 4.1: $L^2(Q)$ errors for $\|\bar{y}_{\omega_u} - \bar{y}_{\text{ref}}\|$, $\|\bar{z}_{\omega_u} - \bar{z}_{\text{ref}}\|$, and $\|\bar{u}_{\omega_u} - \bar{u}_{\text{ref}}\|$.

Table 2

Example 4.1: optimal values of J , μj , and number of iterations, line searches, and Newton steps for $V_{\max} = 32$.

μ	J	μj	#ite.	#search	#Newton
1/35	1.10e-0	1.30e-1	170	666	2193
1/50	7.31e-1	8.21e-2	156	610	1971
1/100	3.60e-1	3.47e-2	121	470	1516
1/500	1.21e-1	5.93e-3	88	341	1099
0	5.80e-2	0	77	297	956

Newton iterations can decrease when the sparse parameter increases. Similar behavior was observed in [7] for different problems with sparse controls.

Finally, in this example we study the convergence of the optimal control and its associated state for vanishing Tikhonov parameter ω_u for $\mu = 0$ and $\mu \neq 0$. The convergence of optimal states and control for vanishing Tikhonov parameter ω_u plays an important role for checking the SSCs according to the theory in [8,38]. It was demonstrated for the one dimensional wave solutions of the classical FHN equation with sparse controls in [8,9]. Here we obtain similar results for two dimensional wave solutions of the convective FHN equations as shown in Fig. 6.

We fix the maximum velocity $V_{\max} = 64$. We take $\bar{y}_{\text{ref}} := y_{1e-10}$, $\bar{z}_{\text{ref}} := z_{1e-10}$, and $\bar{u}_{\text{ref}} := u_{1e-10}$ as reference solutions to determine the order of convergence as $\omega_u \downarrow 0$. We look for the errors, i.e., $\|\bar{y}_{\omega_u} - \bar{y}_{\text{ref}}\|$, $\|\bar{z}_{\omega_u} - \bar{z}_{\text{ref}}\|$, and $\|\bar{u}_{\omega_u} - \bar{u}_{\text{ref}}\|$ as $\omega_u \downarrow 0$. We observe that the errors in L^2 -norm decay linearly as $\omega_u \downarrow 0$ in Fig. 6 for both cases $\mu = 0$ and $\mu = 1/100$. For the norm of $L^\infty(Q_T)$ this looks similar. The trajectories of the target functions, i.e., y_Q and z_Q , are quite achieved as shown in Table 3.

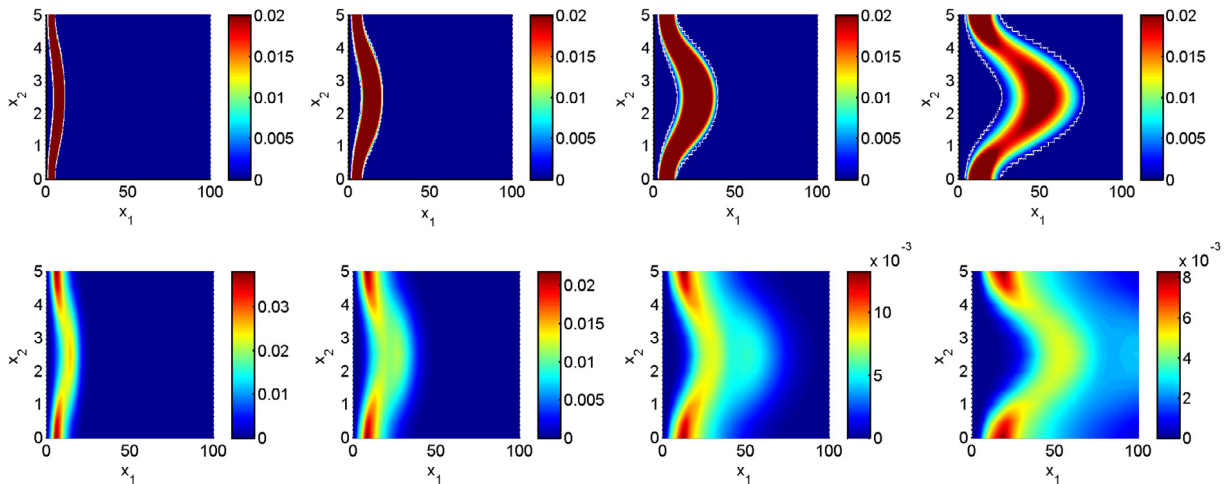


Fig. 7. Example 4.2: optimal controls u (top) and associated states y (bottom) at $t = 1$ without sparse control for $V_{\max} = 16, 32, 64, 128$ (from left to right).

Table 3

Example 4.1: $L^2(Q)$ errors for $\|\bar{y}_{\omega_u} - y_Q\|$ and $\|\bar{z}_{\omega_u} - z_Q\|$.

ω_u	$\mu = 0$		$\mu = 1/100$	
	$\ \bar{y}_{\omega_u} - y_Q\ _{L^2(Q)}$	$\ \bar{z}_{\omega_u} - z_Q\ _{L^2(Q)}$	$\ \bar{y}_{\omega_u} - y_Q\ _{L^2(Q)}$	$\ \bar{z}_{\omega_u} - z_Q\ _{L^2(Q)}$
1	5.77167e-1	1.09495e-1	5.28730e-1	9.90557e-2
1e-1	3.84021e-1	7.32190e-2	4.48344e-1	8.48398e-2
1e-2	3.42134e-1	6.53493e-2	4.02207e-1	7.44464e-2
1e-3	3.39051e-1	6.46970e-2	3.95422e-1	7.25312e-2
1e-4	3.38605e-1	6.46413e-2	3.94826e-1	7.23321e-2
1e-5	3.38590e-1	6.46329e-2	3.94671e-1	7.22975e-2
1e-6	3.38588e-1	6.46321e-2	3.94668e-1	7.22959e-2
1e-7	3.38588e-1	6.46320e-2	3.94667e-1	7.22957e-2
1e-8	3.38588e-1	6.46320e-2	3.94667e-1	7.22957e-2
1e-9	3.38588e-1	6.46320e-2	3.94667e-1	7.22957e-2
1e-10	3.38588e-1	6.46320e-2	3.94667e-1	7.22957e-2

4.2. Terminal control

Our second test example is that the target functions, that is, $y_T(\cdot, T)$ and $z_T(\cdot, T)$, are only given at the final time. Regularization parameters are chosen as $\omega_Q^y = \omega_Q^z = 0$, $\omega_T^y = \omega_T^z = 1$, and $\omega_u = 10^{-3}$. We take the step sizes in space as $\Delta x_1 = \Delta x_2 = 0.125$. We construct the desired functions as done in the previous example. They are given by

$$y_T(x, T) = y_{\text{nat}}(x, T/2) \quad \text{and} \quad z_T(x, T) = z_{\text{nat}}(x, T/2),$$

where y_{nat} and z_{nat} are the solutions of the uncontrolled convective FHN equation at the final time $T = 1$ with the initial conditions

$$y_0(x, t) = \begin{cases} 1, & \text{if } x_a \leq x_1 \leq x_b, \\ 0, & \text{otherwise,} \end{cases} \quad z_0(x, 0) = 0,$$

where $x_a = 2$ and $x_b = 2.2$. The admissible set of controls is defined as

$$\mathcal{U}_{\text{ad}} := \{u \in L^\infty(Q) : 0 \leq u(x, t) \leq 0.2 \text{ for a.e. } (x, t) \in Q\}.$$

We observe similar behavior for the controls and associated states in Fig. 7 and for the optimal value of cost functional J , the number of nonlinear CG iterations, line searches, and Newton steps in Table 4 as obtained for the whole space-time domain example 4.2.

The effect of the sparse parameter for the controlled solutions is displayed in Figs. 8 and 9. As stated before, in this case for higher values of the sparse parameter μ the number of nonlinear CG iterations, line searches and Newton steps decrease as given in Table 5. We observe the same linear decay for the states and the control for vanishing Tikhonov parameter in Fig. 10 as in the whole space-time cylinder, cf. Fig. 6. The values of $\|\bar{y}_{\omega_u} - y_T\|$ and $\|\bar{z}_{\omega_u} - z_T\|$ in L^2 -norm are given in Table 6. We observe that the features of the desired trajectories are quite achieved.

Finally, in this example, we check the distance of discrete control u_h from the local minima u for the verification of SSCs (2.19). Since we do not know the exact optimal solutions of this example, we take the discrete solutions computed with

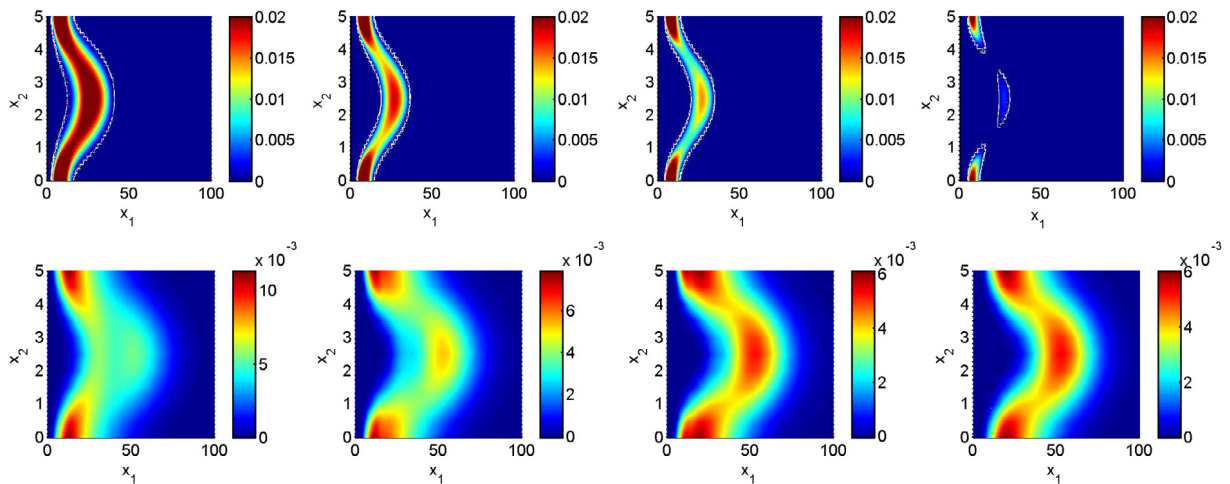


Fig. 8. Example 4.2: optimal controls u (top) and associated states y (bottom) at $t = 1$ for the sparse parameters $\mu = 1/2000, 1/200, 1/150$, and $1/100$ (from left to right) and $V_{\max} = 64$.

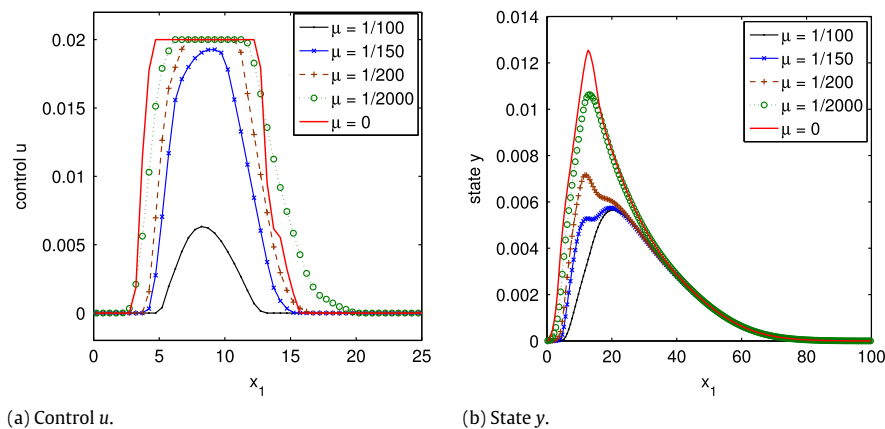


Fig. 9. Example 4.2: profiles of optimal controls u and associated states y (right) at $t = 1$ for various values of the sparse parameter μ .

Table 4

Example 4.2: optimal value of cost functional J , and number of nonlinear CG iterations, line searches, and Newton steps without sparse control.

$V_{\max}$	J	#ite.	#search	#Newton
16	1.12e−2	15	61	198
32	4.30e−2	16	65	211
64	1.42e−3	22	89	289
128	2.89e−4	25	95	304

Table 5

Example 4.2: optimal value of J , μJ , and number of nonlinear CG iterations, line searches, and Newton steps for $V_{\max} = 64$.

μ	J_{opt}	μJ_{opt}	#ite.	#search	#Newton
1/100	1.37e−1	1.15e−2	2	5	16
1/150	9.24e−2	7.70e−3	3	7	18
1/200	7.00e−2	5.88e−3	4	8	19
1/2000	8.79e−3	6.43e−4	19	74	238
0	1.42e−3	0	22	89	289

$\Delta x_1 = \Delta x_2 = 0.3125$ as reference solutions. Table 7 shows the numerical errors of $\|u - u_h\|$ and the error estimates given in (2.25) for $\mu = 1/200$. The numerical results verify Theorem 2.5. However, the error and estimator are not reduced sufficiently for finer discretizations.

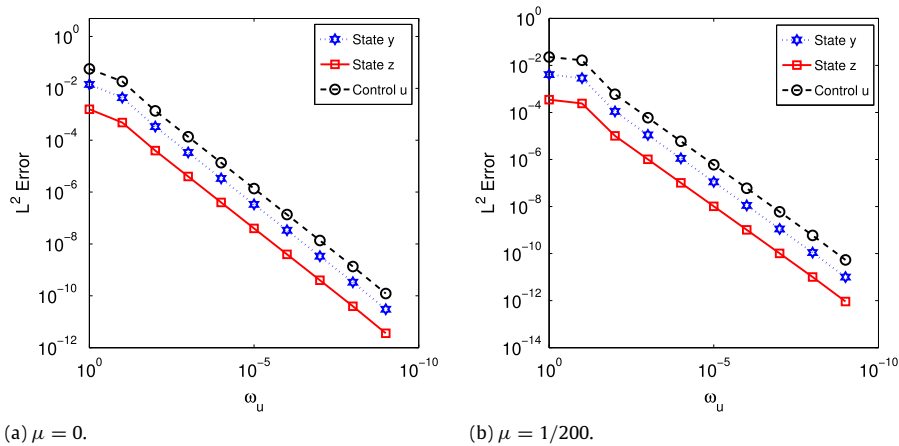


Fig. 10. Example 4.2: $L^2(Q)$ errors for $\|\bar{y}_{\omega_u} - \bar{y}_{\text{ref}}\|$, $\|\bar{z}_{\omega_u} - \bar{z}_{\text{ref}}\|$, and $\|\bar{u}_{\omega_u} - \bar{u}_{\text{ref}}\|$.

Table 6

Example 4.2: $L^2(\Omega)$ errors for $\|\bar{y}_{\omega_u} - y_T\|$ and $\|\bar{z}_{\omega_u} - z_T\|$.

ω_u	$\mu = 0$		$\mu = 1/200$	
	$\ \bar{y}_{\omega_u} - y_T\ _{L^2(\Omega)}$	$\ \bar{z}_{\omega_u} - z_T\ _{L^2(\Omega)}$	$\ \bar{y}_{\omega_u} - y_T\ _{L^2(\Omega)}$	$\ \bar{z}_{\omega_u} - z_T\ _{L^2(\Omega)}$
1	6.53791e-2	2.77918e-2	7.92435e-2	2.81181e-2
1e-1	5.21169e-2	2.74520e-2	7.68586e-2	2.80154e-2
1e-2	4.87670e-2	2.73638e-2	7.13606e-2	2.77851e-2
1e-3	4.86308e-2	2.73684e-2	7.11792e-2	2.77772e-2
1e-4	4.86178e-2	2.73689e-2	7.11610e-2	2.77764e-2
1e-5	4.86165e-2	2.73690e-2	7.11591e-2	2.77763e-2
1e-6	4.86164e-2	2.73690e-2	7.11589e-2	2.77763e-2
1e-7	4.86164e-2	2.73690e-2	7.11589e-2	2.77763e-2
1e-8	4.86164e-2	2.73690e-2	7.11589e-2	2.77763e-2
1e-9	4.86164e-2	2.73690e-2	7.11589e-2	2.77763e-2
1e-10	4.86164e-2	2.73690e-2	7.11589e-2	2.77763e-2

Table 7

Example 4.2: Numerical errors of $\|u - u_h\|$ and error estimates $\frac{2}{\delta} \|\zeta\|$ for sparse controls with $\mu = 1/200$.

$\Delta x_1 = \Delta x_2$	$\ u - u_h\ $	$\frac{2}{\delta} \ \zeta\ $
2.5	2.43e-2	5.92e-2
1.25	2.43e-2	5.60e-2
0.625	2.46e-2	4.78e-2

5. Conclusions

Numerical results of the optimal control governed by the convective FHN model with traveling waves reveal different aspects of the parabolic semi-linear optimal control problems investigated. The second order optimality conditions for local solutions in the form of 2D traveling waves are verified numerically for vanishing Tikhonov regularization parameter ω_u as done for one dimensional waves of the classical FHN equation in [8,9]. By using the second order optimality conditions, we estimate the measure between the discrete solution and the local minima. The control of the convection dominated problems with wave solutions requires a large amount of computing time. In a future study we will investigate the reduced order modeling with proper orthogonal decomposition (POD) in space and model predictive control (MPC) in time [49].

Acknowledgments

The authors would like to thank Fredi Tröltzsch for helpful discussions about this research. The authors also would like to express their sincere thanks to the referees for most valuable suggestions.

References

- [1] J.D. Murray, *Mathematical Biology: I. An Introduction*, Springer-Verlag, New York, 2002.
- [2] B. Karasözen, T. Küçükseyhan, M. Uzunca, Structure preserving integration and model order reduction of skew-gradient reaction–diffusion systems, *Ann. Oper. Res.* (2015) 1–28.

- [3] E.A. Ermakova, M.A. Pantelev, E.E. Shnol, Blood coagulation and propagation of autowaves in flow, *Pathophysiol. Haemost. Thromb.* 34 (2005) 135–142.
- [4] E.A. Ermakova, E.E. Shnol, M.A. Pantelev, A.A. Butylin, V. Volpert, F.I. Ataullakhanov, On propagation of excitation waves in moving media: The FitzHugh–Nagumo model, *PLoS One* 4 (2) (2009) E4454.
- [5] A. Lobanov, T. Starozhilova, The effect of convective flows on blood coagulation processes, *Pathophysiol. Haemost. Thromb.* 34 (2005) 121–134.
- [6] R. Buchholz, H. Engel, E. Kammann, F. Tröltzsch, On the optimal control of the Schlögl-model, *Comput. Optim. Appl.* 56 (1) (2013) 153–185.
- [7] E. Casas, C. Ryll, F. Tröltzsch, Sparse optimal control of the Schlögl and FitzHugh–Nagumo systems, *Comput. Methods Appl. Math.* 13 (4) (2013) 415–442.
- [8] E. Casas, C. Ryll, F. Tröltzsch, Second order and stability analysis for optimal sparse control of the FitzHugh–Nagumo equation, *SIAM J. Control Optim.* 53 (4) (2015) 2168–2202.
- [9] C. Ryll, J. Löber, S. Martens, H. Engel, F. Tröltzsch, Analytical, optimal, and sparse optimal control of traveling wave solutions to reaction–diffusion systems, in: E. Schöll, L.S.H. Klapp, P. Hövel (Eds.), *Control of Self-Organizing Nonlinear Systems*, Springer International Publishing, 2016, pp. 189–210. (Chapter 10).
- [10] M. Stoll, J.W. Pearson, P.K. Maini, Fast solvers for optimal control problems from pattern formation, *J. Comput. Phys.* 304 (2016) 27–45.
- [11] T. Breiten, K. Kunisch, Riccati-based feedback control of the monodomain equations with the FitzHugh–Nagumo model, *SIAM J. Control Optim.* 52 (6) (2014) 4057–4081.
- [12] G. Stadler, Elliptic optimal control problems with L^1 -control cost and applications for the placement of control devices, *Comput. Optim. Appl.* 44 (1) (2009) 159–181.
- [13] E. Casas, R. Herzog, G. Wachsmuth, Optimality conditions and error analysis of semilinear elliptic control problems with L^1 cost functional, *SIAM J. Optim.* 22 (34) (2012) 795–820.
- [14] G. Wachsmuth, D. Wachsmuth, Convergence and regularization results for optimal control problems with sparsity function, *ESAIM Control Optim. Calc. Var.* 17 (3) (2010) 858–886.
- [15] D. Leykekhman, M. Heinkenschloss, Local error analysis of discontinuous Galerkin methods for advection-dominated elliptic linear-quadratic optimal control problems, *SIAM J. Numer. Anal.* 50 (4) (2012) 2012–2038.
- [16] H. Yücel, P. Benner, Distributed optimal control problems governed by coupled convection dominated pdes with control constraints, in: A. Abdulle, S. Deparis, D. Kressner, F. Nobile, M. Picasso (Eds.), *Numerical Mathematics and Advanced Applications, ENUMATH 2013*, in: *Lecture Notes in Computational Science and Engineering*, vol. 103, Springer International Publishing, 2015, pp. 469–478.
- [17] H. Yücel, B. Karasözen, Adaptive symmetric interior penalty Galerkin (SIPG) method for optimal control of convection diffusion equations with control constraints, *Optimization* 63 (2014) 145–166.
- [18] T. Akman, B. Karasözen, Variational time discretization methods for optimal control problems governed by diffusion-convection-reaction equations, *J. Comput. Appl. Math.* 272 (2014) 41–56.
- [19] T. Akman, H. Yücel, B. Karasözen, A priori error analysis of the upwind symmetric interior penalty Galerkin (SIPG) method for the optimal control problems governed by unsteady convection diffusion equations, *Comput. Optim. Appl.* 57 (2014) 703–729.
- [20] H. Yücel, P. Benner, Adaptive discontinuous Galerkin methods for state constrained optimal control problems governed by convection diffusion equations, *Comput. Optim. Appl.* 62 (2015) 291–321.
- [21] H. Yücel, M. Stoll, P. Benner, A discontinuous Galerkin method for optimal control problems governed by a system of convection–diffusion PDEs with nonlinear reaction terms, *Comput. Math. Appl.* 70 (2015) 2414–2431.
- [22] W.W. Hager, H. Zhang, Algorithm 851: CG_DESCENT, a conjugate gradient method with guaranteed descent, *ACM Trans. Math. Software* 32 (1) (2006) 113–137.
- [23] E. Casas, F. Tröltzsch, Second order optimality conditions and their role in PDE control, *Jahresber. Deutsch. Math.-Verein.* 117 (1) (2015) 3–44.
- [24] A. Rösch, D. Wachsmuth, Numerical verification of optimality conditions, *SIAM J. Control Optim.* 47 (5) (2008) 2557–2581.
- [25] A. Rösch, D. Wachsmuth, A-posteriori error estimates for optimal control problems with state and control constraints, *Numer. Math.* 120 (2012) 733–762.
- [26] E. Kammann, F. Tröltzsch, S. Volkwein, A posteriori error estimation for semilinear parabolic optimal control problems with application to model reduction by POD, *ESAIM: M2AN* 47 (2) (2013) 555–581.
- [27] O. Lass, S. Trenz, S. Volkwein, Optimality conditions and POD a-posteriori error estimates for a semilinear parabolic optimal control, *Konstanzer Schr. Math.* 345 (2015).
- [28] C. Pao, *Nonlinear Parabolic and Elliptic Equations*, Plenum Press, New York, 1992.
- [29] F. Tröltzsch, *Optimal Control of Partial Differential Equations: Theory, Methods and Applications*, in: *Graduate Studies in Mathematics*, vol. 112, American Mathematical Society, Providence, RI, 2010.
- [30] W. Barthel, C. John, F. Tröltzsch, Optimal boundary control of a system of reaction diffusion equations, *ZAMM Z. Angew. Math. Mech.* 90 (12) (2010) 966–982.
- [31] R. Griesse, Parametric sensitivity analysis for control-constrained optimal control problems governed by parabolic partial differential equations (Ph.D. thesis), Institut für Mathematik, Universität Bayreuth, Bayreuth, Germany, 2003.
- [32] D.E. Jackson, Existence and regularity for the FitzHugh–Nagumo equations with inhomogeneous boundary conditions, *Nonlinear Anal.-Theor.* 14 (3) (1990) 201–216.
- [33] T. Gudi, A.K. Pani, Discontinuous Galerkin methods for quasi-linear elliptic problems of nonmonotone type, *SIAM J. Numer. Anal.* 45 (1) (2007) 163–192.
- [34] R. Griesse, S. Volkwein, A primal–dual active set strategy for optimal boundary control of a nonlinear reaction–diffusion system, *SIAM J. Control Optim.* 44 (2) (2005) 467–494.
- [35] J.Y. Park, S.H. Park, Optimal control problems for anti-periodic quasi-linear hemivariational inequalities, *Optim. Control Appl. Methods* 28 (4) (2007) 275–287.
- [36] H.D. Mittelmann, Verification of second-order sufficient optimality conditions for semilinear elliptic and parabolic control problems, *Comput. Optim. Appl.* 20 (1) (2001) 93–110.
- [37] N. Arada, E. Casas, F. Tröltzsch, Error estimates for the numerical approximation of a semilinear elliptic control problem, *Comput. Optim. Appl.* 23 (2) (2002) 201–229.
- [38] E. Casas, F. Tröltzsch, Second-order optimality conditions for weak and strong local solutions of parabolic optimal control problems, *Vietnam J. Math.* 44 (2016) 181–202.
- [39] D.N. Arnold, F. Brezzi, B. Cockburn, L.D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, *SIAM J. Numer. Anal.* 39 (5) (2002) 1749–1779.
- [40] B. Rivière, *Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations. Theory and Implementation*, in: *Frontiers in Applied Mathematics*, vol. 35, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008.
- [41] P.G. Ciarlet, The Finite Element Method for Elliptic Problems, in: *Classics Appl. Math.*, vol. 40, SIAM, Philadelphia, PA, 2002.
- [42] A. Cangiani, J. Chapman, E.H. Georgoulis, M. Jensen, On local super-penalization of interior penalty discontinuous Galerkin methods, *Int. J. Numer. Anal. Model.* 11 (3) (2014) 478–495.
- [43] H. Yücel, M. Heinkenschloss, B. Karasözen, Distributed optimal control of diffusion-convection-reaction equations using discontinuous Galerkin methods, in: *Numerical Mathematics and Advanced Applications 2011*, Springer-Verlag, Berlin, Heidelberg, 2013, pp. 389–397.
- [44] W.W. Hager, H. Zhang, A survey of nonlinear conjugate gradient methods, *Pac. J. Optim.* 2 (1) (2006) 35–58.
- [45] J. Nocedal, S.J. Wright, *Numerical Optimization*, second ed., Springer Verlag, Berlin, Heidelberg, New York, 2006.
- [46] R. Herzog, K. Kunisch, Algorithms for pde-constrained optimization, *GAMM-Mitt.* 33 (2) (2010) 163–176.
- [47] N.J. Peruzzi, F.R. Chavarette, J.M. Balthazar, A.M. Tusset, A.L.P.M. Perticarrari, R.M.L.R.F. Brasil, The dynamic behavior of a parametrically excited time-periodic MEMS taking into account parametric errors, *J. Vib. Control* 22 (2016) 4101–4110.
- [48] A.M. Tusset, V. Piccirillo, A. Bueno, J.M. Balthazar, D. Sado, J.L.P. Felix, R.M.L.R.F. Brasil, Chaos control and sensitivity analysis of a double pendulum arm excited by an RLC circuit based nonlinear shaker, *J. Vib. Control* 17 (2016) 3621–3637.
- [49] C. Ryll, F. Tröltzsch, Proper orthogonal decomposition in sparse optimal control of some reaction diffusion equations using model predictive control, *Proc. Appl. Math. Mech.* 14 (1) (2014) 883–884.