

Reduced order optimal control of the convective FitzHugh–Nagumo equations

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

^a Institute of Applied Mathematics & Department of Mathematics, Middle East Technical University, 06800 Ankara, Turkey

^b Department of Mathematics, Sinop University, 57000 Sinop, Turkey

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

ARTICLE INFO

Article history:

Received 3 March 2019

Received in revised form 6 July 2019

Accepted 14 August 2019

Available online 21 August 2019

Keywords:

FitzHugh–Nagumo equation

Optimal control

Discontinuous Galerkin method

Proper orthogonal decomposition

Discrete empirical interpolation

Dynamic mode decomposition

ABSTRACT

In this paper, we compare three model order reduction methods: the proper orthogonal decomposition (POD), discrete empirical interpolation method (DEIM) and dynamic mode decomposition (DMD) for the optimal control of the convective FitzHugh–Nagumo (FHN) equations. The convective FHN equations consist of the semi-linear activator and the linear inhibitor equations, modeling blood coagulation in moving excitable media. The semilinear activator equation leads to a non-convex optimal control problem (OCP). The most commonly used method in reduced optimal control is POD. We use DEIM and DMD to approximate efficiently the nonlinear terms in reduced order models. We compare the accuracy and computational times of three reduced-order optimal control solutions with the full order discontinuous Galerkin finite element solution of the convection dominated FHN equations with terminal controls. Numerical results show that POD is the most accurate whereas POD–DMD is the fastest.

© 2019 Elsevier Ltd. All rights reserved.

1. Introduction

Optimal control problems (OCPs) governed by semilinear partial differential equations (PDEs) with wave-type solutions have been recently investigated for the Schlögl equation with traveling waves fronts and for the Nagumo equation with spiral waves [1], for the FitzHugh–Nagumo (FHN) equations with spiral and traveling waves [2–4], lambda–omega systems with spiral waves [5], Allee equation [6], optimal harvesting [7] and the convective FHN equations [8] with traveling wave solutions. PDE-constrained optimization is also applied to parameter identification in pattern formation. Optimal parameters leading to Turing patterns in semi-linear reaction–diffusion equations like the Schnakenberg and Gierer–Meinhardt equations are identified in [9–11]. Controlling traveling wave solutions or patterns of semi-linear PDEs is computationally challenging. For stable solutions of these OCPs, at each iteration step of the nonlinear optimization algorithms, large scale linear systems should be solved. Efficient and stable solutions of this kind of PDE constrained optimization problems are the focus of the above mentioned studies.

For space discretization of PDEs, commonly used methods are finite differences, finite elements, finite volume, and spectral methods. One of the most stable and accurate methods is the discontinuous Galerkin (dG) method for convection dominated PDEs like the FHN equations. In this paper, we use the symmetric interior penalty Galerkin (SIPG) method for space discretization [12,13], which is the most relevant and popular dG method. The dG methods possess better properties for convective problems compared with continuous finite element methods, but they require a much larger number

* Corresponding author.

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

of degrees of freedom. Recently OCPs governed by linear stationary and time-dependent convection–diffusion–reaction equations [14–16] and semilinear steady state equations [17] are solved efficiently by the several dG methods. OCPs with PDE constraints are discretized usually following two approaches: the *discretize-then-optimize* approach, where the optimality conditions are imposed in the discrete setting, or the *optimize-then-discretize* approach in which the optimality conditions are derived in variational form and then discretization holds. There is no preferred approach [18,19], here we follow the *optimize-then-discretize* approach: the optimality conditions are discretized in space by the SIPG method and the backward Euler method in time. We employ the projected nonlinear conjugate gradient (CG) method [20] for solving the resulting nonlinear discrete optimization problem.

The optimal control of instationary PDEs requires numerous evaluations of the optimality system, which lead to large scale optimization problems. In the last decade, different types of reduced order modeling (ROM) methods were developed to approximate these by low order or surrogate models. The resulting small optimization problems can be solved efficiently with less computational effort. For an overview of ROMs for OCPs for linear and semilinear PDEs, we refer the reader to [18,19]. Despite its heuristic nature, currently, proper orthogonal decomposition (POD) is the preferred model reduction technique for linear and semilinear OCPs. The POD basis functions are constructed from the snapshots, solutions of space–time discretized PDE at pre-specified time-instances, then the reduced order OCP is solved applying the Galerkin projection. ROMs using the POD–Galerkin projection for OCPs for linear PDE constraints are investigated in [21–23] and for semi-linear PDE constraints in [24–26]. With increasing number of POD basis functions, more accurate solutions are obtained. But the success of POD depends on the type of the problem; for semi-linear PDEs with traveling or spiral wave type solutions like the Schlögl and FHN equations [1,4,27] require more POD modes than for the OCPs with semi-linear parabolic equations [24]. Because the reduced nonlinear term still depends on the dimension of the full order model (FOM), i.e., high dimensional finite element discretized model, hyper reduction techniques are developed. The empirical interpolation method (EIM) [28] and the discrete empirical interpolation method (DEIM) [29] are commonly used as hyperreduction techniques to reduce the computational cost for evaluation of the nonlinear terms. More recently, as an alternative to the DEIM, the dynamic mode decomposition (DMD) [30] is used to approximate the nonlinear terms of the PDEs. DMD was first introduced by Schmid [31], Rowley [32] and it is based on the linear approximation of the infinite dimensional Koopman operator [33,34].

We compare the accuracy and computational efficiency of POD, POD with DEIM and POD with DMD for the optimal control of the convective FHN equations modeling blood coagulation and bio-reactors [35,36]. The FHN equations are the simplest and most widely used model for describing the complex spatio-temporal behavior of traveling waves in excited media. In contrast to the classical FHN equations [37,38] with a semi-linear PDE and with a linear ordinary differential equation (ODE), the convective FHN equations consists of a semi-linear PDE with the monotone cubic nonlinear term for the activator and a linear PDE for the inhibitor, modeling excitable systems in moving media. For semi-linear OCPs like the Burger's equation, quadratic nonlinear terms can be reduced to the linear terms [39]. But for higher order polynomial or general nonlinearities, hyperreduction techniques should be used. As far as we know, reduced order OCPs using POD–DEIM and POD–DMD are not yet investigated for semilinear PDE constraint OCPs. Among these three methods, POD–DMD is the fastest, because after collecting snapshots, the non-linearity disappears and the reduced model becomes a coupled system of linear ODEs. Also, the reduced non-convex optimization problem for POD and POD–DEIM becomes convex for POD–DMD. The POD is the most accurate but, the slowest. The POD–DMD is less accurate than the POD but fastest.

In the sequel, we describe the OCP constrained by the convective FHN equations in Section 2. The fully discrete optimal control system is given in Section 3, based on the first order optimality conditions. In Section 4, the reduced order OCPs (POD, POD–DEIM, and POD–DMD) are constructed. In Section 5, we compare the three ROMs with respect to accuracy and computational efficiency for a test problem with terminal control. The paper ends with concluding remarks in Section 6.

2. Optimal control problem

We consider the following OCP:

$$\begin{aligned} \min_{f \in \mathcal{F}_{ad}} J(u, v, f) &= \frac{1}{2} \int_{\Omega} (u(x, T) - u_T(x))^2 dx + \frac{1}{2} \int_{\Omega} (v(x, T) - v_T(x))^2 dx \\ &+ \frac{\nu}{2} \int_0^T \int_{\Omega} f(x, t)^2 dx dt, \end{aligned} \quad (1)$$

governed by the convective FHN equations

$$\begin{aligned} u_t(x, t) - d_u \Delta u(x, t) + \mathbf{b}(x) \cdot \nabla u(x, t) + g(u(x, t)) + v(x, t) &= f(x, t) && \text{in } Q, \\ v_t(x, t) - d_v \Delta v(x, t) + \mathbf{b}(x) \cdot \nabla v(x, t) + \varepsilon(v(x, t) - c_3 u(x, t)) &= 0 && \text{in } Q, \\ \partial_n u(x, t) = 0, \quad \partial_n v(x, t) = 0 &&& \text{on } \Sigma^N, \\ u(x, t) = u_D(x, t), \quad v(x, t) = v_D(x, t) &&& \text{on } \Sigma^D. \end{aligned} \quad (2)$$

where d_u, d_v are the diffusion coefficients, c_3, ε are real constants, $x = (x_1, x_2)^T$ is the spatial element, $u_T(x)$ and $v_T(x)$ are given desired terminal states, and $u_D(x, t)$ and $v_D(x, t)$ are given functions. The control function $f(x, t)$ satisfies the point-wise box constraints

$$f \in \mathcal{F}_{ad} := \{f \in L^\infty(Q) : f_l \leq f(x, t) \leq f_r \text{ for a.e } (x, t) \in Q\},$$

for some real numbers $-\infty < f_l < f_r < +\infty$. The cubic polynomial $g(u)$ is given by

$$g(u) = c_1 u(u - c_2)(u - 1),$$

where $c_i, i = 1, 2$, are non-negative real numbers. The non-linearity $g(u)$ is monostable for $0 < c_1 < 20$ and $c_2 = 0.02$ [35], whereas for the Schlögl equation [1], the classical FHN equations [2], and the diffusive FHN equations [40] the bistable cubic non-linearity is considered. The velocity field $\mathbf{b}(x) = (b_1(x), b_2(x))^T$ is divergence-free, and for the maximum wave speed $b_{\max}$, it is given by the parabolic profile along the x_1 -direction

$$b_1(x) = ax_2(H - x_2), \quad b_{\max} = \frac{1}{4}aH^2, \quad a > 0, \quad b_2(x) = 0.$$

In (1) and (2), $T > 0$ and $Q := \Omega \times (0, T]$ denote the final time and the time-space domain, respectively. The space domain $\Omega = (0, L) \times (0, H)$ is bounded, Lipschitz in $\mathbb{R}^2$, and $\Sigma = \partial\Omega \times (0, T]$ is the lateral surface. The Dirichlet and Neumann boundaries are $\Sigma^D := \Gamma_D \times (0, T]$ and $\Sigma^N := \Gamma_N \times (0, T]$, where $\Gamma_D = \{x = \{0, L\}, 0 < y < H\}$ and $\Gamma_N = \{0 < x < L, y = \{0, H\}\}$. The outward unit normal vector is denoted by $\mathbf{n}(x)$ and the outward normal derivative on $\partial\Omega$ is defined as $\partial_n u = \nabla u \cdot \mathbf{n}$.

The convective FHN equations (2) consists of an activator equation for $u(x, t)$ and an inhibitor equation for $v(x, t)$ in an excitable medium. In blood coagulation process, the activator variable $u(x, t)$ describes the concentrations of thrombin in the excitation, and $v(x, t)$ the inhibition of this excitation and recovery of the medium variable, which is activated factor XI [36]. The complex process of coagulation consists of cascadic enzymatic reactions and feedback loops [41]. These reactions enable generation of auto-catalytic thrombin outside of the damaged region. The most important property of the blood coagulation process is the formation of auto-waves with the velocity independent initial conditions [35,36,41]. The flow propagates within the impermeable channel walls (Neumann boundary conditions). Different types of standing and triggering waves occur depending on the constants of the convective FHN equations [35]. The waves become more curved when the flow velocity $b_{\max}$ increases. For a detailed discussion of the complex spatio-temporal wave phenomena occurring in the uncontrolled convective FHN equations, we refer the reader to [35,42].

The waves are initiated at $t = 0$ inside an initial narrow excitation region $\{x_a \leq x_1 \leq x_b, 0 \leq x_2 \leq H\}$

$$u_0(x) = \begin{cases} \bar{u}, & \text{if } x_a \leq x_1 \leq x_b, \\ 0, & \text{otherwise,} \end{cases} \quad v_0(x) = 0. \quad (3)$$

The interval from x_a to x_b is the damaged region of the vessel wall. In blood coagulation, for fixed thrombin concentration the initial condition (3) mimics the activation process [35,36,41]. In this study, we take a rectangle as $x_a = 0, x_b = 0.1$, and we set $\bar{u} = 0.1$.

Optimal control aims that the activator $u(x, t)$ and the inhibitor $v(x, t)$ follow desired terminal states $u_T(x)$ and $v_T(x)$ as close as possible in the L_2 norm by minimizing the cost functional $J(u, v, f)$. The inhibitor $v(x, t)$ has only some auxiliary character as for the FHN equations in [2–4], therefore only the activator $u(x, t)$ is to be controlled by $f(x, t)$, which seems to be experimentally feasible [43]. The parameter $\nu > 0$ in the optimal cost functional $J(u, v, f)$ denotes the penalization or the Tikhonov regularization parameter.

Remark 2.1. We have considered the OCP (1) with terminal observation functionals to show the applicability of reduced order modeling. Reduced solutions of OCPs with observation functionals, distributed in space–time for the FHN equations [8] can also be handled using the same approach.

3. Optimality system and discretization

The local optimal solutions are constructed using the first-order necessary optimality conditions. In [8], we consider a more general OCP constrained by the FHN equations with distributed, terminal and sparse controls. Here we consider the OCP (1) only with the terminal controls and without sparse controls. As in [8], we follow the *optimize-then-discretize* approach, first by deriving the first order optimality conditions in the variational form and then discretize the OCP. The second order necessary and sufficient optimality conditions for semi-linear parabolic equations are investigated in [2,44]. The first-order necessary optimality conditions [44, Chapter 1], lead to the following adjoint system

$$\begin{aligned} -p_t(x, t) - d_u \Delta p(x, t) - \mathbf{b}(x) \cdot \nabla p(x, t) + g'(u)p(x, t) - \varepsilon c_3 q(x, t) &= 0, \\ -q_t(x, t) - d_v \Delta q(x, t) - \mathbf{b}(x) \cdot \nabla q(x, t) + \varepsilon q(x, t) + p(x, t) &= 0, \end{aligned} \quad (4)$$

with the mixed boundary and terminal time conditions

$$\begin{aligned} d_u \partial_n p(x, t) + (\mathbf{b} \cdot \mathbf{n})p(x, t) &= 0, \quad d_v \partial_n q(x, t) + (\mathbf{b} \cdot \mathbf{n})q(x, t) = 0, & \text{on } \Sigma^N, \\ p(x, t) &= 0, \quad q(x, t) = 0, & \text{on } \Sigma^D, \\ p(x, T) &= u(x, T) - u_T(x), \quad q(x, T) = v(x, T) - v_T(x), & \text{in } \Omega. \end{aligned}$$

The flow in the adjoint system (4) is in the opposite direction of the flow in the state system (2). In view of the box constraints, the optimal solution $f(x, t)$ has to satisfy the following variational inequality

$$\int_0^T \int_{\Omega} (p(x, t) + \nu f(x, t))(v(x, t) - f(x, t)) \, dx \, dt \geq 0, \quad \forall v \in \mathcal{F}_{ad},$$

which leads to the point-wise projection formula

$$f(x, t) = \mathbb{P}_{[f_l, f_r]} \left(-\frac{1}{v} p(x, t) \right),$$

with the projection operator $\mathbb{P}_{[a, b]}(\phi) = \max\{a, \min\{\phi, b\}\}$.

3.1. Space discretization of the optimality system

The optimality system consisting of the state equation (2) and the adjoint equation (4) is discretized in space using the SIPG method [13]. We denote the family of meshes consisting of non-overlapping triangular elements K by $\{\mathcal{T}_h\}_h$. The diameter of an element K and the length of an edge E are denoted by h_K and h_E , respectively. We assume that the mesh is regular in the following sense: for different triangles $K_i, K_j \in \{\mathcal{T}_h\}_h, i \neq j$, the intersection $K_i \cap K_j$ is either empty or a vertex or an edge, i.e., hanging nodes are not allowed.

The discrete test, state, and control spaces are defined by the space of piecewise discontinuous finite element functions

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

Here, the set of polynomials on $K \in \mathcal{T}_h$ of degree at most l is denoted by $\mathbb{P}^l(K)$. In numerical examples, we use linear discontinuous finite elements, i.e., $l = 1$.

The set $\mathcal{E}_h$ of all edges are split into the sets of interior, Dirichlet and Neumann boundary edges, denoted respectively by $\mathcal{E}_h^0, \mathcal{E}_h^D, \mathcal{E}_h^N$, so that $\mathcal{E}_h = \mathcal{E}_h^0 \cup \mathcal{E}_h^D \cup \mathcal{E}_h^N$. The inflow and outflow boundaries are defined by

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

Similarly, the inflow and outflow boundaries of an element $K \in \mathcal{T}_h$ are defined by

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

where $\mathbf{n}_K(x)$ is the exterior unit normal vector on the boundary ∂K of an element K . Let the common edge for two elements K and K^e be E . Then there are two traces of a function $w \in W_h$ along the edge E , denoted by $w|_E$ from the interior of K and $w^e|_E$ from the interior of K^e . Accordingly, the jump and average of w across the edge E are defined by

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

Similarly, the jump and average of a vector-valued function are given by

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

We set $\llbracket w \rrbracket = w|_E \mathbf{n}$ and $\{\nabla w\} = \nabla w|_E$ on a boundary edge $E \in K \cap \partial\Omega$. Then, the state equations (2) and the adjoint equations (4) are discretized using the SIPG method leading to the system: $\forall w \in W_h$ and a.e. $t \in (0, T]$

$$\begin{aligned} \langle \partial_t u_h, w \rangle + a_{h,u}(u_h, w) + g_h(u_h, w) + \langle v_h, w \rangle &= \ell_{h,u}(t, w) + \langle f_h, w \rangle, \\ \langle u_h(x, 0), w(x) \rangle &= \langle u_0(x), w(x) \rangle, \\ \langle \partial_t v_h, w \rangle + a_{h,v}(v_h, w) + \varepsilon \langle v_h, w \rangle - \varepsilon c_3 \langle u_h, w \rangle &= \ell_{h,v}(t, w), \\ \langle v_h(x, 0), w(x) \rangle &= \langle v_0(x), w(x) \rangle, \end{aligned} \quad (5)$$

$$\begin{aligned} -\langle \partial_t p_h, w \rangle + a_{h,p}(p_h, w) + \langle g'(u_h) p_h, w \rangle - \varepsilon c_3 \langle q_h, w \rangle &= 0, \\ \langle p_h(x, T), w(x) \rangle &= \langle u_h(x, T) - u_T(x), w(x) \rangle, \\ -\langle \partial_t q_h, w \rangle + a_{h,q}(q_h, w) + \varepsilon \langle q_h, w \rangle + \langle p_h, w \rangle &= 0, \\ \langle q_h(x, T), w(x) \rangle &= \langle v_h(x, T) - v_T(x), w(x) \rangle, \end{aligned} \quad (6)$$

where $\langle \cdot, \cdot \rangle$ denotes the usual L^2 -inner product. The bilinear forms $a_{h,u}$ and $a_{h,v}$ in the discrete state system (5) are given for $i = u, v$ by

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

for any $\tilde{w}$, $w \in W_h$, and γ is called the penalty parameter [13]. For $i = u, v$, the linear right-hand side vectors, and the nonlinear form are defined by

$$\begin{aligned}\ell_{h,i}(t, w) &= \sum_{E \in \mathcal{E}_h^D} \int_E i_D(x, t) \left(\frac{d_i \gamma}{h_E} \mathbf{n} \cdot \llbracket w \rrbracket - \{ \{ d_i \nabla w \} \} \right) ds - \sum_{K \in \mathcal{T}_h} \int_{\partial K^- \cap \Gamma^-} \mathbf{b} \cdot \mathbf{n} i_D(x, t) w ds, \\ g_h(u, w) &= \sum_{K \in \mathcal{T}_h} \int_K g(u) w dx.\end{aligned}$$

The bilinear forms $a_{h,p}$ and $a_{h,q}$ in the adjoint system (6) are defined similar to the bilinear forms $a_{h,u}$ and $a_{h,v}$ for states, respectively, but with the negative velocity field. They contain, respectively, the additional terms

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

which come from the mixed boundary conditions. The SIPG discrete states and the control in (5) are given in the form

$$u_h(x, t) = \sum_{i=1}^{N_K} \sum_{j=1}^{N_l} u_j^i(t) \phi_j^i(x), \quad v_h(x, t) = \sum_{i=1}^{N_K} \sum_{j=1}^{N_l} v_j^i(t) \phi_j^i(x), \quad f_h(x, t) = \sum_{i=1}^{N_K} \sum_{j=1}^{N_l} f_j^i(t) \phi_j^i(x).$$

In a similar way, the discrete desired terminal states (just L^2 -projections) are given by

$$u_{h,T}(x) = \sum_{i=1}^{N_K} \sum_{j=1}^{N_l} u_{T,j}^i \phi_j^i(x), \quad v_{h,T}(x) = \sum_{i=1}^{N_K} \sum_{j=1}^{N_l} v_{T,j}^i \phi_j^i(x).$$

The number of (triangular) elements are denoted by N_K , and the local dimension of each element is given by N_l . ϕ_j^i denotes the j th finite element basis function defined on the i th triangle. Setting dG degrees of freedom $N := N_K \times N_l$, the corresponding time dependent unknown coefficient vectors can be written as

$$\begin{aligned}\mathbf{u}(t) &= (u_1^1(t), \dots, u_{N_l}^1(t), \dots, u_1^{N_K}(t), \dots, u_{N_l}^{N_K}(t)) \in \mathbb{R}^N, \\ \mathbf{v}(t) &= (v_1^1(t), \dots, v_{N_l}^1(t), \dots, v_1^{N_K}(t), \dots, v_{N_l}^{N_K}(t)) \in \mathbb{R}^N, \\ \mathbf{f}(t) &= (f_1^1(t), \dots, f_{N_l}^1(t), \dots, f_1^{N_K}(t), \dots, f_{N_l}^{N_K}(t)) \in \mathbb{R}^N.\end{aligned}$$

On the other hand, the known coefficient vectors of the discrete desired terminal states are given by

$$\begin{aligned}\mathbf{u}_T &= (u_{T,1}^1, \dots, u_{N_l}^{T,1}, \dots, u_1^{T,N_K}, \dots, u_{N_l}^{T,N_K}) \in \mathbb{R}^N, \\ \mathbf{v}_T &= (v_{T,1}^1, \dots, v_{N_l}^{T,1}, \dots, v_1^{T,N_K}, \dots, v_{N_l}^{T,N_K}) \in \mathbb{R}^N,\end{aligned}$$

which are computed from the projection identities

$$\langle u_{h,T}(x), \phi_i(x) \rangle = \langle u_T(x), \phi_i(x) \rangle, \quad \langle v_{h,T}(x), \phi_i(x) \rangle = \langle v_T(x), \phi_i(x) \rangle, \quad i = 1, \dots, N.$$

Then, the SIPG semi-discretized system of the state equations (5) leads to the following system of ODEs

$$\begin{aligned}\mathbf{M}\mathbf{u}_t + \mathbf{S}_u \mathbf{u} + \mathbf{g}(\mathbf{u}) + \mathbf{M}\mathbf{v} &= \boldsymbol{\ell}_u + \mathbf{M}\mathbf{f}, \\ \mathbf{M}\mathbf{v}_t + \mathbf{S}_v \mathbf{v} + \varepsilon \mathbf{M}\mathbf{v} - \varepsilon c_3 \mathbf{M}\mathbf{u} &= \boldsymbol{\ell}_v,\end{aligned} \tag{7}$$

and the semi-discrete objective functional takes form

$$\begin{aligned}J_h(\mathbf{u}, \mathbf{v}, \mathbf{f}) &= \frac{1}{2}(\mathbf{u} - \mathbf{u}_T)^T \mathbf{M}(\mathbf{u} - \mathbf{u}_T) + \frac{1}{2}(\mathbf{v} - \mathbf{v}_T)^T \mathbf{M}(\mathbf{v} - \mathbf{v}_T) \\ &\quad + \frac{\nu}{2} \int_0^T \mathbf{f}^T \mathbf{M} \mathbf{f} dt.\end{aligned}$$

Here $\mathbf{M}$ denotes the mass matrix, and $\mathbf{S}_u$ and $\mathbf{S}_v$ are the stiffness matrices corresponding to the bilinear forms $a_{h,u}$ and $a_{h,v}$, respectively. The nonlinear vector $\mathbf{g}(\mathbf{u})$ corresponds to the nonlinear form g_h , and the time-dependent vectors $\boldsymbol{\ell}_u := \boldsymbol{\ell}_u(t)$ and $\boldsymbol{\ell}_v := \boldsymbol{\ell}_v(t)$ are the vectors corresponding to the linear forms $\ell_{h,u}$ and $\ell_{h,v}$, respectively. By a similar setting, the SIPG semi-discretized system of the adjoint equations (6) are given as the following system of ODEs

$$\begin{aligned}-\mathbf{M}\mathbf{p}_t + \mathbf{S}_p \mathbf{p} + \mathbf{R}_g(\mathbf{u}) \mathbf{p} - \varepsilon c_3 \mathbf{M}\mathbf{q} &= 0, \\ -\mathbf{M}\mathbf{q}_t + \mathbf{S}_q \mathbf{q} + \varepsilon \mathbf{M}\mathbf{q} + \mathbf{M}\mathbf{p} &= 0,\end{aligned} \tag{8}$$

where the state $u(x, t)$ dependent matrix $\mathbf{R}_g(\mathbf{u})$ is related to the term $\langle g'(u_h) p_h, w \rangle$ in the adjoint system (6), and $\mathbf{S}_p$ and $\mathbf{S}_q$ are the stiffness matrices corresponding to the bilinear forms $a_{h,p}$ and $a_{h,q}$, respectively.

3.2. Fully discrete optimality system

The time interval $[0, T]$ is partitioned uniformly as $0 = t_0 < t_1 < \dots < t_{N_T} = T$, with the step-size $\Delta t = T/N_T$. For $n = 0, 1, \dots, N_T$, we denote by $\mathbf{u}_n, \mathbf{v}_n, \mathbf{f}_n, \mathbf{p}_n$ and $\mathbf{q}_n$ the approximate solution vectors to the semi-discrete solutions $\mathbf{u}(t), \mathbf{v}(t), \mathbf{f}(t), \mathbf{p}(t)$ and $\mathbf{q}(t)$ at the time instance $t = t_n$, respectively. Also, we set $\ell_u^n := \ell_u(t_n)$ and $\ell_v^n := \ell_v(t_n)$. Then, applying the backward Euler method to the semi-discrete state equation (7) and to the semi-discrete adjoint equation (8), we obtain the following fully discrete state system

$$\begin{aligned} \frac{1}{\Delta t} \mathbf{M}(\mathbf{u}_n - \mathbf{u}_{n-1}) + \mathbf{S}_u \mathbf{u}_n + \mathbf{g}(\mathbf{u}_n) + \mathbf{M} \mathbf{v}_n &= \ell_u^n + \mathbf{M} \mathbf{f}_n, \\ \frac{1}{\Delta t} \mathbf{M}(\mathbf{v}_n - \mathbf{v}_{n-1}) + \mathbf{S}_v \mathbf{v}_n + \varepsilon \mathbf{M} \mathbf{v}_n - \varepsilon \mathbf{C}_3 \mathbf{M} \mathbf{u}_n &= \ell_v^n, \\ (\mathbf{M} \mathbf{u}_0)_i &= \langle u_0(x), \phi_i \rangle, \quad (\mathbf{M} \mathbf{v}_0)_i = \langle v_0(x), \phi_i \rangle, \end{aligned} \quad (9)$$

for $n = 1, 2, \dots, N_T$, and the following fully discrete adjoint system

$$\begin{aligned} \frac{1}{\Delta t} \mathbf{M}(\mathbf{p}_{n-1} - \mathbf{p}_n) + \mathbf{S}_p \mathbf{p}_{n-1} + \mathbf{R}_g(\mathbf{u}_n) \mathbf{p}_{n-1} - \varepsilon \mathbf{C}_3 \mathbf{M} \mathbf{q}_{n-1} &= 0, \\ \frac{1}{\Delta t} \mathbf{M}(\mathbf{q}_{n-1} - \mathbf{q}_n) + \mathbf{S}_q \mathbf{q}_{n-1} + \varepsilon \mathbf{M} \mathbf{q}_{n-1} + \mathbf{M} \mathbf{p}_{n-1} &= 0, \\ \mathbf{p}_{N_T} &= \mathbf{u}_{N_T} - \mathbf{u}_T, \quad \mathbf{q}_{N_T} = \mathbf{v}_{N_T} - \mathbf{v}_T, \end{aligned} \quad (10)$$

for $n = N_T, \dots, 2, 1$, where $\mathbf{u}_T$ and $\mathbf{v}_T$ are the known coefficient vectors of the desired terminal states. We note that the fully discrete state system (9) is solved forward in time, whereas the fully discrete adjoint system (10) is solved backward in time. Eqs. (9) and (10) are semi-explicit in time. Because the mass matrix $\mathbf{M}$ is symmetric positive definite and therefore invertible, the discrete solutions of (9) and (10) exist and are unique. The resulting semilinear discrete OCP can be solved by several optimization algorithms. We use the projected nonlinear CG method [20], which is applied for solving the OCP governed by Schlögl and FHN equations in [1–4] and by the convective FHN equations in [8]. The details of the implementation of the projected nonlinear CG algorithm can be found in these papers.

4. Reduced order optimal control

We construct ROMs for the state equation (2) and the adjoint equation (4). For the construction of the POD modes to the state variables u and v , the POD method uses the snapshot matrices generated by the coefficient vectors $\mathbf{u}$ and $\mathbf{v}$ of the optimal state solution vectors from the full order optimal control problem, and similarly, the snapshot matrix generated by the coefficient vector $\mathbf{f}$ of the optimal control solution is used to construct the POD modes to the control variable f . For the construction of the POD modes to the adjoint variables p and q , we use the same POD modes generated by the snapshots of the states u and v , motivated by the error analysis in [19]. This approach might be not the best option, but the construction of different POD modes using the snapshot matrices of the adjoint states does not improve the suboptimal solutions much [24] and requires more computational work. It is not known a-priori how far the optimal solution of the reduced POD problem is from the exact one unless the snapshots are generating a sufficiently large state space. Different approaches are discussed in [18] for the proper choice of the POD basis in the reduced OCP, which affects the accuracy of the reduced solutions. Here we construct the POD basis for the reduced OCP from the snapshots of controls of the FOM.

4.1. POD Galerkin discretization

The reduced order system for the state equations (2) of lower dimensions k_u and k_v is formed by the Galerkin projection of the equations onto the subspaces

$$W_{u,h}^r = \text{span}\{\psi_{u,1}, \dots, \psi_{u,k_u}\} \subset W_h, \quad W_{v,h}^r = \text{span}\{\psi_{v,1}, \dots, \psi_{v,k_v}\} \subset W_h.$$

For any time t , the reduced order states $u_h^r(x, t) \in W_{u,h}^r$ and $v_h^r(x, t) \in W_{v,h}^r$ stand for the approximations of the full order states $u_h(x, t)$ and $v_h(x, t)$, respectively, and they are given by

$$u_h(x, t) \approx u_h^r(x, t) = \sum_{i=1}^{k_u} u_i^r(t) \psi_{u,i}(x), \quad v_h(x, t) \approx v_h^r(x, t) = \sum_{i=1}^{k_v} v_i^r(t) \psi_{v,i}(x),$$

where $\mathbf{u}^r(t) := (u_1^r(t), \dots, u_{k_u}^r(t))^T$ and $\mathbf{v}^r(t) := (v_1^r(t), \dots, v_{k_v}^r(t))^T$ are the reduced coefficient vectors, and $\{\psi_{u,i}\}_{i=1}^{k_u}$ and $\{\psi_{v,i}\}_{i=1}^{k_v}$ are the L^2 -orthogonal reduced basis functions. Similarly, the full order control $f_h(x, t)$ is approximated by the reduced order control $f_h^r(x, t)$ from the k_f dimensional subspace $W_{f,h}^r$

$$f_h(x, t) \approx f_h^r(x, t) = \sum_{i=1}^{k_f} f_i^r(t) \psi_{f,i}(x), \quad W_{f,h}^r = \text{span}\{\psi_{f,1}, \dots, \psi_{f,k_f}\} \subset W_h,$$

where $\mathbf{f}^r(t) := (f_1^r(t), \dots, f_{k_f}^r(t))^T$ is the reduced coefficient vector for the control, and $\{\psi_{f,i}\}_{i=1}^{k_f}$ are the L^2 -orthogonal reduced basis functions. All the reduced basis functions are linear combinations of dG basis functions

$$\psi_{u,i} = \sum_{j=1}^N \Psi_{u,j,i} \phi_j(x), \quad \psi_{v,i} = \sum_{j=1}^N \Psi_{v,j,i} \phi_j(x), \quad \psi_{f,i} = \sum_{j=1}^N \Psi_{f,j,i} \phi_j(x),$$

where $\Psi_{\cdot,\cdot,i}$ are the coefficient vectors of the i th reduced basis functions. For $w \in \{u, v, f\}$, the following POD matrices are constructed

$$\Psi_w := [\Psi_{w,1} \cdots \Psi_{w,k_w}] \in \mathbb{R}^{N \times k_w}, \quad \Psi_{w,i} = (\Psi_{w,1,i}, \dots, \Psi_{w,N,i})^T \in \mathbb{R}^N,$$

where the columns $\Psi_{w,i}$ are the POD modes. The POD modes are computed through the snapshot matrices $\mathcal{U} = [\mathbf{u}_1, \dots, \mathbf{u}_{N_T}]$, $\mathcal{V} = [\mathbf{v}_1, \dots, \mathbf{v}_{N_T}]$ and $\mathcal{F} = [\mathbf{f}_1, \dots, \mathbf{f}_{N_T}]$ in $\mathbb{R}^{N \times N_T}$ [40,45], respectively, where the columns of the snapshot matrices are simply the coefficient vectors of the full order states and control from the FOM (9) at the time instances t_i , $i = 1, \dots, N_T$. Then, for $w \in \{u, v, f\}$, the reduced basis functions $\{\psi_{w,i}\}$, $i = 1, 2, \dots, k_w$, are given by the solution of the following minimization problem

$$\min_{\psi_{w,1}, \dots, \psi_{w,k_w}} \frac{1}{N_T} \sum_{j=1}^{N_T} \left\| \mathbf{w}_j - \sum_{i=1}^{k_w} \langle \mathbf{w}_j, \psi_{w,i} \rangle \psi_{w,i} \right\|_{L^2(\Omega)}^2$$

subject to $\langle \psi_{w,i}, \psi_{w,j} \rangle = \Psi_{w,i}^T \mathbf{M} \Psi_{w,j} = \delta_{ij}$, $1 \leq i, j \leq k_w$,

where δ_{ij} is the Kronecker delta. Again for $w \in \{u, v, f\}$, and for $\mathcal{W} \in \{\mathcal{U}, \mathcal{V}, \mathcal{F}\}$, the above minimization problem is equivalent to the eigenvalue problems

$$\mathcal{W} \mathcal{W}^T \mathbf{M} \Psi_{w,\cdot,i} = \sigma_{w,i}^2 \Psi_{w,\cdot,i}, \quad i = 1, 2, \dots, k_w, \quad (11)$$

for the coefficient vectors $\Psi_{w,\cdot,i}$ of the reduced basis functions $\psi_{w,i}$. Setting $\widehat{\mathcal{W}} = R \mathcal{W}$ (R^T is the Cholesky factor of the mass matrix $\mathbf{M}$), we obtain the equivalent formulation of (11) as

$$\widehat{\mathcal{W}} \widehat{\mathcal{W}}^T \widehat{\Psi}_{w,\cdot,i} = \sigma_{w,i}^2 \widehat{\Psi}_{w,\cdot,i}, \quad i = 1, 2, \dots, k_w,$$

where $\widehat{\Psi}_{w,\cdot,i} = R \Psi_{w,\cdot,i}$. Because the singular value decomposition (SVD) is more stable and efficient than the eigenvalue decomposition, we apply the SVD to compute the first k_w left singular vectors, $\widehat{\Psi}_{w,\cdot,i} := \zeta_{w,i}$, of the matrix $\widehat{\mathcal{W}}$

$$\widehat{\mathcal{W}} = \zeta_w \Sigma_w \beta_w^T,$$

where the diagonal matrix Σ_w contains the singular values $\sigma_{w,i}$, and $\zeta_{w,i}$ are the columns of the orthogonal matrix ζ_w . Using the fact that $\widehat{\Psi}_{w,\cdot,i} = R \Psi_{w,\cdot,i}$, the coefficient vectors $\Psi_{w,\cdot,i}$ of the reduced basis functions are computed as

$$\Psi_{w,\cdot,i} = R^{-1} \widehat{\Psi}_{w,\cdot,i}, \quad i = 1, 2, \dots, k_w.$$

By the above settings, the following relations can be found between the coefficient vectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{f}$ of the FOM solutions, and the reduced coefficient vectors $\mathbf{u}^r$, $\mathbf{v}^r$ and $\mathbf{f}^r$

$$\begin{aligned} \mathbf{u} &\approx \Psi_u \mathbf{u}^r, \quad \mathbf{v} \approx \Psi_v \mathbf{v}^r, \quad \mathbf{f} \approx \Psi_f \mathbf{f}^r, \\ \mathbf{u}^r &\approx \Psi_u^T \mathbf{M} \mathbf{u}, \quad \mathbf{v}^r \approx \Psi_v^T \mathbf{M} \mathbf{v}, \quad \mathbf{f}^r \approx \Psi_f^T \mathbf{M} \mathbf{f}. \end{aligned}$$

Then the reduced system for the state equations takes the form

$$\begin{aligned} \frac{d}{dt} \mathbf{u}^r + \mathbf{S}_u^r \mathbf{u}^r + \Psi_u^T \mathbf{g}(\Psi_u \mathbf{u}^r) + \mathbf{M}_{u,v}^r \mathbf{v}^r &= \Psi_u^T \ell_u + \mathbf{M}_{u,f}^r \mathbf{f}^r, \\ \frac{d}{dt} \mathbf{v}^r + \mathbf{S}_v^r \mathbf{v}^r + \varepsilon \mathbf{v}^r - \varepsilon c_3 \mathbf{M}_{v,u}^r \mathbf{u}^r &= \Psi_v^T \ell_v, \end{aligned} \quad (12)$$

and using the L^2 -orthogonality of the reduced basis functions, the semi-discrete reduced order objective functional takes form

$$\begin{aligned} J_h^r(\mathbf{u}^r, \mathbf{v}^r, \mathbf{f}^r) &= \frac{1}{2} (\mathbf{u}^r - \mathbf{u}_T^r)^T (\mathbf{u}^r - \mathbf{u}_T^r) + \frac{1}{2} (\mathbf{v}^r - \mathbf{v}_T^r)^T (\mathbf{v}^r - \mathbf{v}_T^r) \\ &\quad + \frac{\nu}{2} \int_0^T (\mathbf{f}^r)^T \mathbf{f}^r dt, \end{aligned}$$

which is absolutely cheaper to calculate, comparing to the full order objective functional J_h . Using the same POD modes generated by the snapshot matrix of the states, we can obtain the reduced adjoint system in a similar way

$$\begin{aligned} -\frac{d}{dt} \mathbf{p}^r + \mathbf{S}_p^r \mathbf{p}^r + \mathbf{R}_g^r(\mathbf{u}^r) \mathbf{p}^r - \varepsilon c_3 \mathbf{M}_{u,v}^r \mathbf{q}^r &= 0, \\ -\frac{d}{dt} \mathbf{q}^r + \mathbf{S}_q^r \mathbf{q}^r + \varepsilon \mathbf{q}^r + \mathbf{M}_{v,u}^r \mathbf{p}^r &= 0. \end{aligned} \quad (13)$$

In the ROMs (12) and (13), the reduced matrices are defined by

$$\begin{aligned} \mathbf{S}_u^r &= \Psi_u^T \mathbf{S}_u \Psi_u, & \mathbf{S}_v^r &= \Psi_v^T \mathbf{S}_v \Psi_v, & \mathbf{S}_p^r &= \Psi_u^T \mathbf{S}_p \Psi_u, & \mathbf{S}_q^r &= \Psi_v^T \mathbf{S}_q \Psi_v \\ \mathbf{M}_{u,v}^r &= \Psi_u^T \mathbf{M} \Psi_v, & \mathbf{M}_{u,f}^r &= \Psi_u^T \mathbf{M} \Psi_f, & \mathbf{M}_{v,u}^r &= \Psi_v^T \mathbf{M} \Psi_u, \\ \mathbf{R}_g^r(\mathbf{u}^r) &= \Psi_u^T \mathbf{R}_g(\Psi_u \mathbf{u}^r) \Psi_u. \end{aligned}$$

The sub-optimality systems (12) and (13) are solved by backward Euler method in time.

4.2. Discrete empirical interpolation method

In this part, we rewrite the reduced state equation (12) resulting from the application of the DEIM [29] to the cubic nonlinear term $\mathbf{g}(\Psi_u \mathbf{u}^r)$. The DEIM aims to approximate the nonlinear vector $\mathbf{g}(\Psi_u \mathbf{u}^r)$ by projecting it onto a subspace of the space generated by the non-linear functions and spanned by a basis of dimension $m \ll N$

$$\mathbf{g}(\Psi_u \mathbf{u}^r) \approx \mathbf{W} \mathbf{s}(t), \quad (14)$$

where $\mathbf{s}(t) \in \mathbb{R}^m$ is the corresponding coefficient vector, and the DEIM basis matrix $\mathbf{W} \in \mathbb{R}^{N \times m}$ is calculated by the application of the POD to the nonlinear snapshot matrix $\mathcal{G} = [\mathbf{g}(\mathbf{u}_1) \cdots \mathbf{g}(\mathbf{u}_{N_T})] \in \mathbb{R}^{N \times N_T}$. Since the system (14) is overdetermined, we find a projection matrix $\mathbf{P}$ such that $\mathbf{P} = [\mathbf{e}_{p_1} \cdots \mathbf{e}_{p_m}] \in \mathbb{R}^{N \times m}$ where $\mathbf{e}_{p_i}$ is the i th column of the identity matrix $\mathbf{I} \in \mathbb{R}^{N \times N}$. After elimination of the coefficient vector $\mathbf{s}(t)$ in (14), we arrive at the following ROM for the state equation with the DEIM approximation as follows

$$\begin{aligned} \frac{d}{dt} \mathbf{u}^r + \mathbf{S}_u^r \mathbf{u}^r + \mathbf{Q} \mathbf{g}^r(\mathbf{u}^r) + \mathbf{M}_{u,v}^r \mathbf{v}^r &= \Psi_u^T \ell_u + \mathbf{M}_{u,f}^r \mathbf{f}^r, \\ \frac{d}{dt} \mathbf{v}^r + \mathbf{S}_v^r \mathbf{v}^r + \varepsilon \mathbf{v}^r - \varepsilon c_3 \mathbf{M}_{v,u}^r \mathbf{u}^r &= \Psi_v^T \ell_v, \end{aligned} \quad (15)$$

where the matrix $\mathbf{Q} := \Psi_u^T \mathbf{W} (\mathbf{P}^T \mathbf{W})^{-1}$ is computed once in the off-line stage. The computation of the reduced nonlinear vector $\mathbf{g}^r(\mathbf{u}^r) := \mathbf{P}^T \mathbf{g}(\Psi_u \mathbf{u}^r) \in \mathbb{R}^m$ and its Jacobian matrix require only $m \ll N$ and $m \times N_l$ integral evaluations, respectively, while they need N and $N \times N_l$ integral evaluations when DEIM approximation is not used.

4.3. Dynamic mode decomposition

The DMD extracts dynamically relevant spatio-temporal information content from numerical or experimental data sets [34]. Without explicit knowledge of the dynamical system, the DMD algorithm determines eigenvalues, eigenmodes, and spatial structures for each mode. The snapshot data is decomposed in spatio-temporal modes both by POD and DMD that correlates the spatial features of the data. Additionally, DMD associates the snapshot data to the temporal Fourier modes. DMD is a special case of the Koopman operator [33] approximating nonlinear systems via an associated infinite dimensional system. The connection between the DMD and Koopman operator was established in [31,32]. The Koopman operator $\mathcal{K}$ acts on a set of scalar observable functions $g : \mathcal{M} \rightarrow \mathbb{C}$

$$\mathcal{K} g(\mathbf{y}) = g(\mathbf{N}(\mathbf{y})),$$

for the nonlinear dynamical system

$$\frac{d\mathbf{y}}{dt} = \mathbf{N}(\mathbf{y}),$$

where $\mathbf{y} \in \mathcal{M}$, and $\mathcal{M}$ is an n -dimensional manifold. The DMD determines the Koopman eigenvalues and modes directly from the data, when linear observables are considered as state space, $g(\mathbf{y}) = \mathbf{y}$, directly from the data.

We consider a snapshot matrix and a time shifted version of the snapshot matrix

$$\mathbf{Y} = [\mathbf{y}(t_0) \cdots \mathbf{y}(t_{m-1})], \quad \mathbf{Y}' = [\mathbf{y}(t_1) \cdots \mathbf{y}(t_m)]. \quad (16)$$

The DMD involves the decomposition of the unknown best-fit linear operator $\mathbf{A} \in \mathbb{R}^{n \times n}$ relating the matrices above

$$\mathbf{Y}' = \mathbf{A} \mathbf{Y}.$$

The DMD modes are computed by the exact DMD algorithm [46].

In [30], the DMD is proposed for producing low-rank approximations of the nonlinearities in PDEs as an alternative to DEIM. Here, after building the POD basis functions of rank k_u , we collect snapshots $\mathcal{G}, \mathcal{G}'$ for the non-linear term $\mathbf{g}(\mathbf{u})$ from the FOM (7) at N_{T+1} time instances and divide them into sets as in (16)

$$\mathcal{G} = [\mathbf{g}(\mathbf{u}_0) \cdots \mathbf{g}(\mathbf{u}_{N_T-1})] \in \mathbb{R}^{N \times N_T}, \quad \mathcal{G}' = [\mathbf{g}(\mathbf{u}_1) \cdots \mathbf{g}(\mathbf{u}_{N_T})] \in \mathbb{R}^{N \times N_T}.$$

The snapshots matrices satisfy

$$\mathcal{G}' = \mathbf{A}_{\mathcal{G}} \mathcal{G},$$

Algorithm 1 Exact DMD Algorithm

$\mathcal{G}$ and $\mathcal{G}'$ are the snapshot matrices

Computation of SVD of $\mathcal{G}$, $\mathcal{G} = U \Sigma V^*$.

Define $\tilde{\mathbf{A}}_{\mathcal{G}} = U^* \mathcal{G}' V \Sigma^{-1}$.

Computation of the eigenvalues and eigenvectors (λ_j, w_j) from $\tilde{\mathbf{A}}_{\mathcal{G}} w_j = \lambda_j w_j, j = 1, \dots, \tilde{m}$.

Computation of DMD modes $\Psi_j^{\text{DMD}} := \mathcal{G}' V \Sigma^{-1} w_j, j = 1, \dots, \tilde{m}$.

where the unknown matrix $\mathbf{A}_{\mathcal{G}}$ is given as the solution of the minimization problem

$$\min \|\mathcal{G}' - \mathbf{A}_{\mathcal{G}} \mathcal{G}\|_F^2,$$

where $\|\cdot\|_F$ denotes the Frobenius norm, $\mathbf{A}_{\mathcal{G}} = \mathcal{G}' \mathcal{G}^\dagger$, and $\dagger$ is the Moore–Penrose pseudoinverse. The Algorithm 1 gives the exact DMD procedure [46] to find the DMD basis matrix $\Psi^{\text{DMD}} = [\Psi_1^{\text{DMD}} \dots \Psi_{\tilde{m}}^{\text{DMD}}] \in \mathbb{R}^{N \times \tilde{m}}$. After calculating the DMD basis function Ψ^{DMD} by the Algorithm 1, the nonlinear vector $\mathbf{g}(\mathbf{u})$ in the full order state equations (7) is approximated as the following time-dependent vector

$$\mathbf{g}^{\text{DMD}}(t) = \sum_{j=1}^{\tilde{m}} \alpha_j \Psi_j^{\text{DMD}} \exp(\omega_j t) = \Psi^{\text{DMD}} \text{diag}(e^{\omega t}) \alpha, \quad (17)$$

where $\alpha = [\alpha_1, \dots, \alpha_{\tilde{m}}]^T$ is the initial amplitudes given by $\alpha = (\Psi^{\text{DMD}})^\dagger \mathbf{g}(\mathbf{u}_1)$, and $\omega = [\omega_1, \dots, \omega_{\tilde{m}}]^T$ includes the eigenvalues λ_j as $\omega_j = \log(\lambda_j)/\Delta t, j = 1, \dots, \tilde{m}$.

Inserting the identity (17) into the reduced state equations (12), we obtain the following linear ROM for the activator equation

$$\frac{d}{dt} \mathbf{u}^r + \mathbf{S}_u^r \mathbf{u}^r + \Psi_u^T \mathbf{g}^{\text{DMD}}(t) + \mathbf{M}_{u,v}^r \mathbf{v}^r = \Psi_u^T \ell_u + \mathbf{M}_{u,f}^r \mathbf{f}^r. \quad (18)$$

Although the dimension of the reduced system (18) with $\tilde{m}$ DMD modes is small as the reduced system (15) with m DEIM modes, the POD–DMD reduced state equation (18) is linear, and the OCP problem becomes convex and Newton's iterations are no more needed. Therefore, the POD–DMD is much faster than the POD and the POD–DEIM.

5. Numerical results

We consider the OCP (1)–(2) with a fast wave speed $V_{\max} = 64$, and with the parameters and initial condition given by

$$c_1 = 9, \quad c_2 = 0.02, \quad c_3 = 5, \quad \varepsilon = 0.1, \quad d_y = d_z = 1, \quad v = 10^{-3},$$

$$u_0(x, t) = \begin{cases} 0.1, & \text{if } 0 \leq x_1 \leq 0.1, \quad 0 \leq x_2 \leq H \\ 0, & \text{otherwise,} \end{cases} \quad v_0(x, 0) = 0.$$

The desired terminal states are chosen as

$$u_T(x) = u_{\text{nat}}(x, T/2) \quad \text{and} \quad v_T(x) = v_{\text{nat}}(x, T/2),$$

where u_{nat} and v_{nat} stand for the solutions of the uncontrolled convective FHN equations. The box constraint for the control is given by

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

The space domain is taken as a rectangle with $H = 5$ and $L = 100$. We use uniform step size in space, $\Delta x_1 = \Delta x_2 = 0.5$, and in time, $\Delta t = 0.05$ for the final time $T = 1$. Stopping criteria for the OCP is the relative error $|J_{\text{old}} - J|/|J_{\text{old}}| \leq 10^{-3}$ on the objective functional. For $w \in \{u, v, f\}$, the number of POD modes are determined according to the following energy criteria (relative information content (RIC))

$$\varepsilon(k_w) = \frac{\sum_{i=1}^{k_w} \sigma_{w,i}^2}{\sum_{i=1}^{d_w} \sigma_{w,i}^2},$$

which represents the energy captured by the first k_w POD modes over all d_w POD modes, d_w is the rank of the related snapshot matrix, and $\sigma_{w,i}$ is the corresponding singular value of the i th mode. In the computations, the number of POD modes are taken as the same, $k := \max\{k_u, k_v, k_f\}$, so that the energy criteria $\min_{k_w} \varepsilon(k_w) \geq 99.99\%$ is satisfied for $w \in \{u, v, f\}$. Moreover, for $w \in \{u, v, f\}$, we define the following relative L^2 -error

$$\|w_h - w_h^r\| := \frac{\|w_h(x, t) - w_h^r(x, t)\|_{L^2(\Omega)}}{\|w_h(x, t)\|_{L^2(\Omega)}},$$

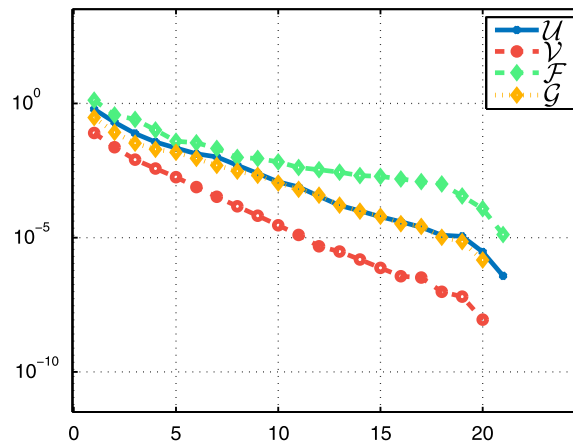


Fig. 1. Singular values of the snapshot matrices.

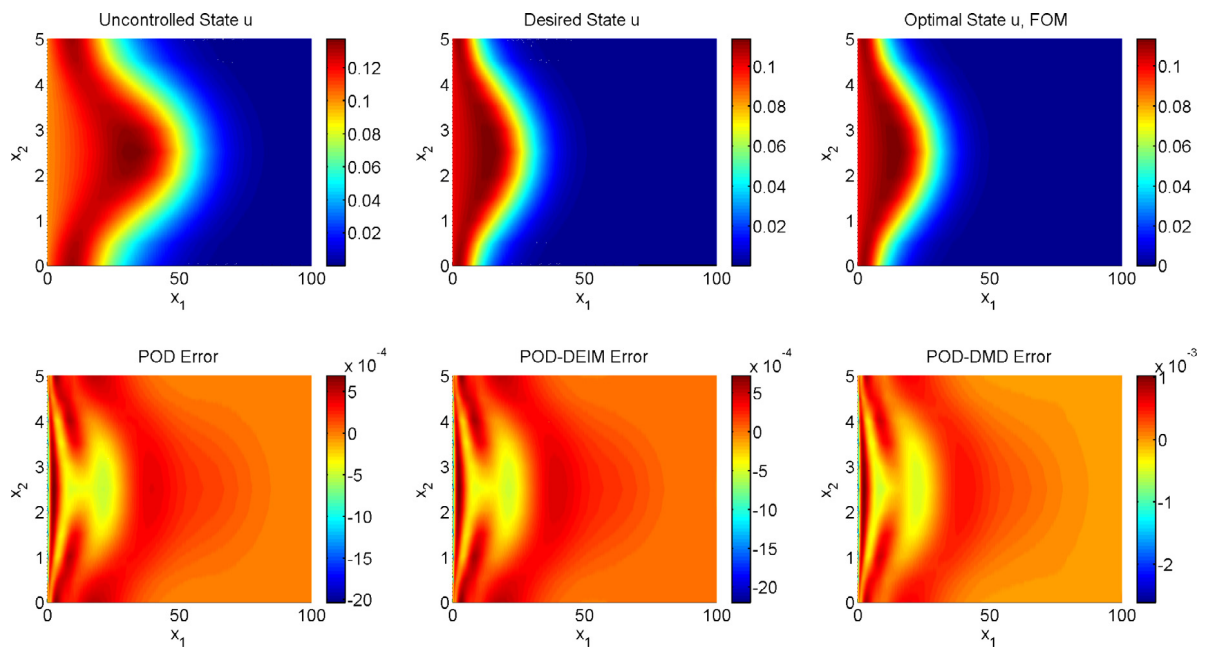


Fig. 2. State u profiles at the final time: (Top) Uncontrolled, desired and optimal FOM states; (Bottom) FOM-ROM errors for $k = 9$ POD modes, $m = 14$ DEIM basis and $\tilde{m} = 18$ DMD basis.

to compute the errors between full and reduced order solutions at a given time t .

The singular values of the snapshot matrices of the states u , v , of the control f , and of the non-linearity $g(u)$ in Fig. 1 decay slowly almost at the same rate, which is typical for convection dominated problems. According to the RIC for %99.99 of energy captured, we have chosen $k = 9$ POD modes for the states u , v and the control f .

The FOM state solutions at the final time in Figs. 2–3, top, show the same wave type characteristics as in [8]. The error plots between full and reduced order state solutions in Figs. 2–3, bottom, agree for POD, POD-DEIM and POD-DMD with $m = 14$ and $\tilde{m} = 18$ DEIM and DMD modes, respectively. In Fig. 4, we give the full and reduced order control profiles at the final time. We see that there is a bit difference between the full and reduced order profiles, but the reduced order profiles are similar in between. The relative L^2 -errors between the full and reduced order solutions of the states and the control at the final time are shown in Table 1. Both states u and v are approximated by the reduced models with the same accuracy, whereas the errors in control are larger for POD-DMD due to linearization of the reduced OCP.

Because the Galerkin projected POD-DMD system is linear, the resulting OCP is convex. Therefore, no Newton iteration is recorded for the POD-DMD in Table 2, and instead of nonlinear CG, we use linear CG method. Comparing the speedup factors (the ratio of the CPU times (in seconds) required for evaluation of FOM to the CPU times required for evaluation of ROM) in Table 2, the computational efficiency of the POD-DMD approach is clearly visible. The POD reduced sub-optimal

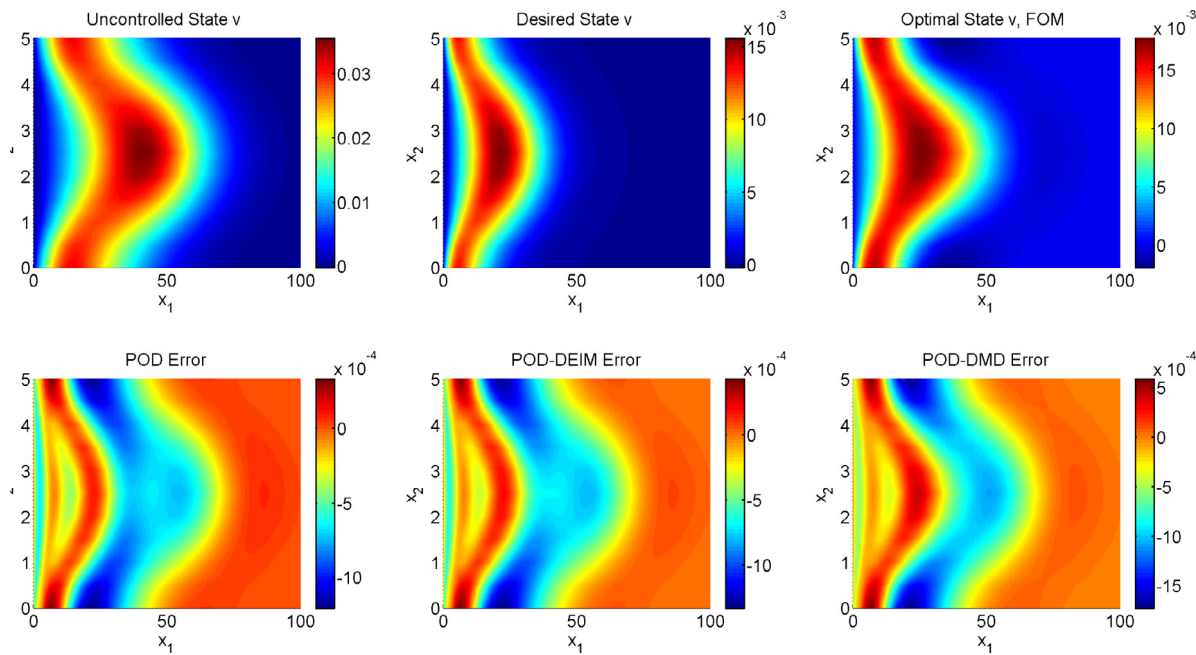


Fig. 3. State v profiles at the final time: (Top) Uncontrolled, desired and optimal FOM states; (Bottom) FOM–ROM errors for $k = 9$ POD modes, $m = 14$ DEIM basis and $\tilde{m} = 18$ DMD-basis.

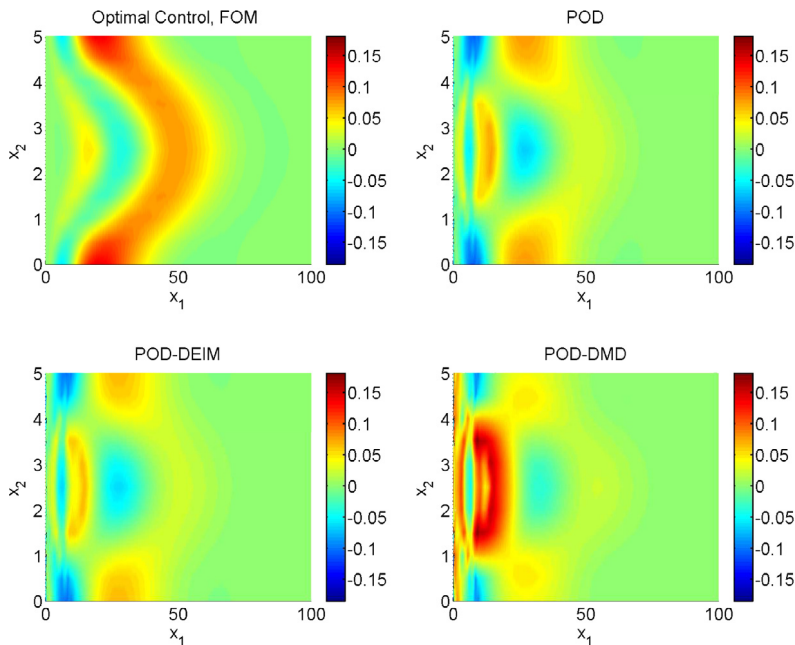


Fig. 4. Control profiles at the final time for $k = 9$ POD modes $k = 9$, $m = 14$ DEIM basis and $\tilde{m} = 18$ DMD basis.

Table 1 Relative L^2 -errors between full and reduced order states and control at the final time.			
	$\ u_h - u_h^r\ $	$\ v_h - v_h^r\ $	$\ f_h - f_h^r\ $
POD	4.644e−03	6.167e−02	6.614e−01
POD–DEIM	4.988e−03	6.511e−02	7.123e−01
POD–DMD	5.787e−03	7.670e−02	1.061e+00

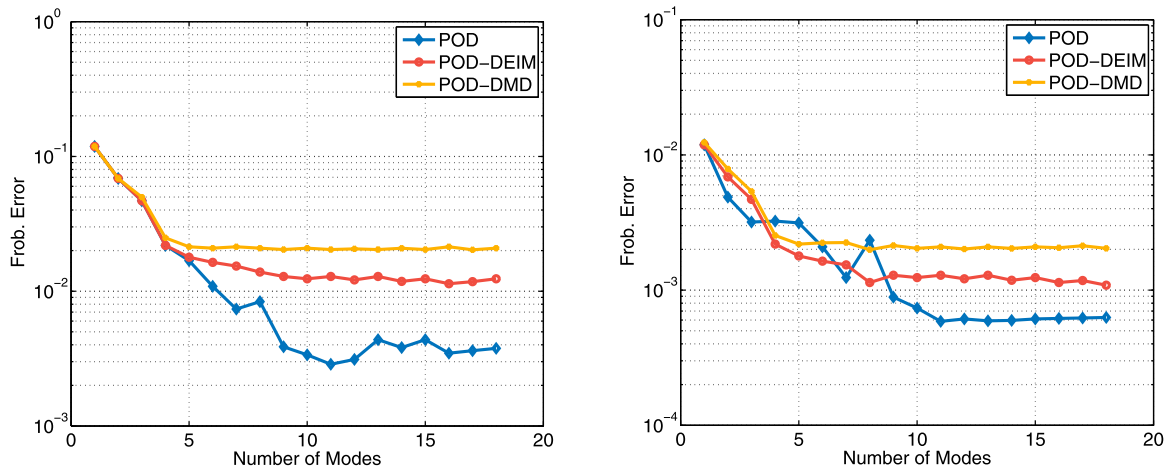


Fig. 5. Relative FOM-ROM Frobenius errors for state u (left) and state v (right).

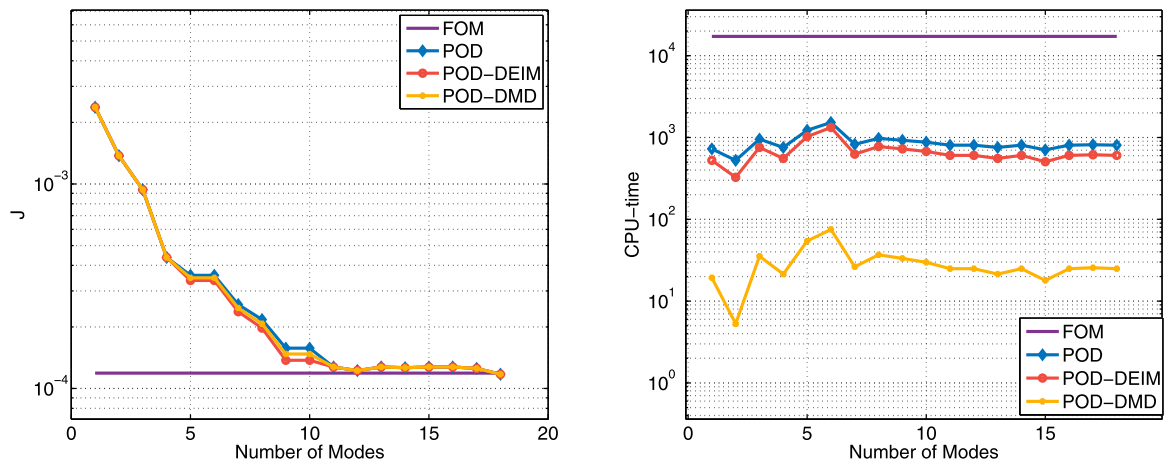


Fig. 6. The values of the objective functionals (left) and CPU times (right) with increasing number of POD modes.

Table 2

The values of the optimal objective functionals J , nonlinear CG iterations and line searches, average number of Newton iterations per time step, CPU times and speedup factors.

	J	# CG iterations	# Line searches	# Newton iterations	CPU times	Speedup factors
FOM	2.376e-03	184	710	3.22	15496.7	–
POD	2.396e-03	85	317	3.19	828.3	18.7
POD-DEIM	2.392e-03	88	332	3.20	386.2	40.1
POD-DMD	2.396e-03	99	370	1	8.0	1937.1

solutions are the most accurate, the POD-DEIM and POD-DMD solutions are less accurate, whereas the POD solutions oscillate at higher modes, as shown for the relative FOM-ROM errors in the Frobenius norm in Fig. 5, and in Table 1 as well. The value of the objective function J obtained by full and reduced order solutions are almost the same with increasing number of POD modes in Fig. 6, left. The ROMs require significantly less CPU time compared to the FOM. In terms of computational cost, the POD-DMD is the fastest, Fig. 6, right. For a small number of POD modes, the CPU times oscillate until $k \approx 6$, and after they do not change much. Therefore it is not necessary to use a larger number of POD modes considering the CPU times, and the accuracy of the reduced order solutions and reduced order objective functional.

6. Conclusions

In this study, the computational efficiency and accuracy of three ROM methods are investigated for solving semilinear convection dominated OCP problem. We derived the ROMs for the state and adjoint equations using the same POD basis

functions generated by the snapshot matrices of the states only. Two hyperreduction techniques, the DEIM and DMD are applied to reduce the computational cost arising from the nonlinear term in the activator equation. Among the three ROM techniques the POD without DEIM/DMD is the most accurate as expected. The POD–DEIM and POD–DMD errors are close, but the POD–DMD is the fastest due to the fact that the reduced state equations are no more nonlinear and the OCP is convex. In a future study, we plan to compare these ROM techniques for model predictive control and feedback control problems using compressive POD and DMD.

Acknowledgments

The authors would like to thank the reviewers for the comments and suggestions that helped to improve the manuscript.

References

- [1] R. Buchholz, H. Engel, E. Kammann, F. Tröltzsch, On the optimal control of the Schlögl-model, *Comput. Optim. Appl.* 56 (1) (2013) 153–185, <http://dx.doi.org/10.1007/s10589-013-9550-y>.
- [2] 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, <http://dx.doi.org/10.1515/cmam-2013-0016>.
- [3] 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, <http://dx.doi.org/10.1137/140978855>.
- [4] 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, S.H. L. Klapp, P. Hövel (Eds.), *Control of Self-Organizing Nonlinear Systems*, Springer International Publishing, 2016, pp. 189–210, http://dx.doi.org/10.1007/978-3-319-28028-8_10.
- [5] A. Borzi, R. Griesse, Distributed optimal control of lambda-omega systems, *J. Numer. Math.* 14 (1) (2006) 17–40, <http://dx.doi.org/10.1163/156939506776382120>.
- [6] E. Trélat, J. Zhu, E. Zuazua, Allee optimal control of a system in ecology, *Math. Models Methods Appl. Sci.* 28 (09) (2018) 1665–1697, <http://dx.doi.org/10.1142/S021820251840002X>.
- [7] H. Öz Bakan, F. Yılmaz, G.-W. Weber, A discrete optimality system for an optimal harvesting problem, *Comput. Manag. Sci.* 14 (4) (2017) 519–533, <http://dx.doi.org/10.1007/s10287-017-0286-5>.
- [8] M. Uzunca, T. Küçükseyhan, H. Yücel, B. Karasözen, Optimal control of convective FitzHugh-Nagumo Equation, *Comput. Math. Appl.* 73 (9) (2017) 2151–2169, <http://dx.doi.org/10.1016/j.camwa.2017.02.028>.
- [9] M.R. Garvie, P.K. Maini, C. Trenchea, An efficient and robust numerical algorithm for estimating parameters in Turing systems, *J. Comput. Phys.* 229 (19) (2010) 7058–7071, <http://dx.doi.org/10.1016/j.jcp.2010.05.040>.
- [10] M. Garvie, C. Trenchea, Identification of space–time distributed parameters in the Gierer–Meinhardt reaction–diffusion system, *SIAM J. Appl. Math.* 74 (1) (2014) 147–166, <http://dx.doi.org/10.1137/120885784>.
- [11] M. Stoll, J.W. Pearson, P.K. Maini, Fast solvers for optimal control problems from pattern formation, *J. Comput. Phys.* 304 (2016) 27–45, <http://dx.doi.org/10.1016/j.jcp.2015.10.006>.
- [12] 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, <http://dx.doi.org/10.1137/S0036142901384162>.
- [13] B. Riviere, *Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations: Theory and Implementation*, SIAM, 2008.
- [14] 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, <http://dx.doi.org/10.1016/j.cam.2014.05.002>.
- [15] 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 (3) (2014) 703–729, <http://dx.doi.org/10.1007/s10589-013-9601-4>.
- [16] H. Yücel, B. Karasözen, Adaptive symmetric interior penalty Galerkin (SIPG) method for optimal control of convection diffusion equations with control constraints, *Optimization* 63 (1) (2014) 145–166, <http://dx.doi.org/10.1080/02331934.2013.801474>.
- [17] 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 (10) (2015) 2414–2431, <http://dx.doi.org/10.1016/j.camwa.2015.09.006>.
- [18] P. Benner, E. Sachs, S. Volkwein, Model order reduction for pde constrained optimization, in: G. Leugering, P. Benner, S. Engell, A. Griewank, H. Harbrecht, M. Hinze, R. Rannacher, S. Ulbrich (Eds.), *Trends in PDE Constrained Optimization*, Springer International Publishing, Cham, 2014, pp. 303–326, http://dx.doi.org/10.1007/978-3-319-05083-6_19.
- [19] M. Gubisch, S. Volkwein, Proper orthogonal decomposition for linear-quadratic optimal control, in: P. Benner, A. Cohen, M. Ohlberger, K. Willcox (Eds.), *Model Reduction and Approximation: Theory and Algorithms*, Computational Science & Engineering, SIAM, 2017, pp. 3–63, <http://dx.doi.org/10.1137/1.9781611974829.ch1>.
- [20] 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, <http://dx.doi.org/10.1145/1132973.1132979>.
- [21] M. Hinze, S. Volkwein, Error estimates for abstract linear–quadratic optimal control problems using proper orthogonal decomposition, *Comput. Optim. Appl.* 39 (3) (2008) 319–345, <http://dx.doi.org/10.1007/s10589-007-9058-4>.
- [22] K. Kunisch, S. Volkwein, Proper orthogonal decomposition for optimality systems, *ESAIM Math. Model. Numer. Anal.* 42 (1) (2008) 1–23, <http://dx.doi.org/10.1051/m2an:2007054>.
- [23] F. Tröltzsch, S. Volkwein, POD a-posteriori error estimates for linear-quadratic optimal control problems, *Comput. Optim. Appl.* 44 (1) (2008) 83–115, <http://dx.doi.org/10.1007/s10589-008-9224-3>.
- [24] E. Kammann, F. Tröltzsch, S. Volkwein, A posteriori error estimation for semilinear parabolic optimal control problems with application to model reduction by POD, *ESAIM Math. Model. Numer. Anal.* 47 (2013) 555–581, <http://dx.doi.org/10.1051/m2an/2012037>.
- [25] A. Studinger, S. Volkwein, Numerical analysis of POD a-posteriori error estimation for optimal control, in: K. Bredies, C. Clason, K. Kunisch, G. von Winckel (Eds.), *Control and Optimization with PDE Constraints*, Springer Basel, Basel, 2013, pp. 137–158, http://dx.doi.org/10.1007/978-3-0348-0631-2_8.
- [26] K. Kunisch, M. Müller, Uniform convergence of the POD method and applications to optimal control, *Discrete Contin. Dyn. Syst.* 35 (9) (2015) 4477–4501, <http://dx.doi.org/10.3934/dcds.2015.35.4477>.
- [27] C. Ryll, F. Tröltzsch, Proper orthogonal decomposition in sparse optimal control of some reaction diffusion equations using model predictive control, *PAMM* 14 (1) (2014) 883–884, <http://dx.doi.org/10.1002/pamm.201410422>.

- [28] M. Barrault, Y. Maday, N.C. Nguyen, A.T. Patera, An empirical interpolation method: application to efficient reduced-basis discretization of partial differential equations, *C. R. Math.* 339 (9) (2004) 667–672, <http://dx.doi.org/10.1016/j.crma.2004.08.006>.
- [29] S. Chaturantabut, D.C. Sorensen, Nonlinear model reduction via discrete empirical interpolation, *SIAM J. Sci. Comput.* 32 (5) (2010) 2737–2764, <http://dx.doi.org/10.1137/090766498>.
- [30] A. Alla, J. Kutz, Nonlinear model order reduction via dynamic mode decomposition, *SIAM J. Sci. Comput.* 39 (5) (2017) B778–B796, <http://dx.doi.org/10.1137/16M1059308>.
- [31] P.J. Schmid, Dynamic mode decomposition of numerical and experimental data, *J. Fluid Mech.* 656 (2010) 5–28, <http://dx.doi.org/10.1017/S0022112010001217>.
- [32] C.W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, D.S. Henningson, Spectral analysis of nonlinear flows, *J. Fluid Mech.* 641 (2009) 115–127, <http://dx.doi.org/10.1017/S0022112009992059>.
- [33] B.O. Koopman, Hamiltonian systems and transformation in Hilbert space, *Proc. Natl. Acad. Sci.* 17 (5) (1931) 315–318.
- [34] J.N. Kutz, S.L. Brunton, B.W. Brunton, J.L. Proctor, *Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems*, SIAM, 2016.
- [35] E.A. Ermakova, E.E. Shnol, M.A. Panteleev, A.A. Butylin, V. Volpert, F.I. Ataullakhanov, On propagation of excitation waves in moving media: The FitzHugh-Nagumo model, *PLoSOne* 4 (2) (2009) E4454, <http://dx.doi.org/10.1371/journal.pone.0004454>.
- [36] E. Ermakova, M. Panteleev, E. Shnol, Blood coagulation and propagation of autowaves in flow, *Pathophysiol. Haemost. Thromb.* 34 (2005) 135–142, <http://dx.doi.org/10.1159/000089933>.
- [37] R. FitzHugh, Impulses and physiological states in theoretical models of nerve membrane, *Biophys. J.* 1 (6) (1961) 445–466, [http://dx.doi.org/10.1016/S0006-3495\(61\)86902-6](http://dx.doi.org/10.1016/S0006-3495(61)86902-6).
- [38] J. Nagumo, S. Arimoto, S. Yoshizawa, An active pulse transmission line simulating nerve axon, *Proc. IRE* 50 (10) (1962) 2061–2070, <http://dx.doi.org/10.1109/jrproc.1962.288235>.
- [39] M. Baumann, P. Benner, J. Heiland, Space-time Galerkin POD with application in optimal control of semilinear partial differential equations, *SIAM J. Sci. Comput.* 40 (3) (2018) A1611–A1641, <http://dx.doi.org/10.1137/17M1135281>.
- [40] B. Karasözen, T. Küçükseyhan, M. Uzunca, Structure preserving integration and model order reduction of skew-gradient reaction-diffusion systems, *Ann. Oper. Res.* 258 (2017) 79–106, <http://dx.doi.org/10.1007/s10479-015-2063-6>.
- [41] A. Lobanov, T. Starozhilova, The effect of convective flows on blood coagulation processes, *Pathophysiol. Haemost. Thromb.* 34 (2005) 121–134, <http://dx.doi.org/10.1159/000089932>.
- [42] F.I. Ataullakhanov, V.I. Zarnitsina, A.V. Pokhilko, A.I. Lobanov, O.L. Morozova, Spatio-temporal dynamics of blood coagulation and pattern formation: A theoretical approach, *Int. J. Bifurcation Chaos* 12 (09) (2002) 1985–2002, <http://dx.doi.org/10.1142/S0218127402005649>.
- [43] C. Vilas, M.R. García, J.R. Banga, A.A. Alonso, Robust feed-back control of travelling waves in a class of reaction-diffusion distributed biological systems, *Physica D* 237 (18) (2008) 2353–2364, <http://dx.doi.org/10.1016/j.physd.2008.02.019>.
- [44] F. Tröltzsch, *Optimal Control of Partial Differential Equations*, in: Graduate Studies in Mathematics, vol. 112, American Mathematical Society, Providence, RI, 2010, <http://dx.doi.org/10.1090/gsm/112>.
- [45] K. Kunisch, S. Volkwein, Galerkin proper orthogonal decomposition methods for parabolic problems, *Numer. Math.* 90 (1) (2001) 117–148, <http://dx.doi.org/10.1007/s002110100282>.
- [46] J.H. Tu, C.W. Rowley, D.M. Luchtenburg, S.L. Brunton, J.N. Kutz, On dynamic mode decomposition: theory and applications, *J. Comput. Dyn.* 1 (2) (2014) 391–421, <http://dx.doi.org/10.3934/jcd.2014.1.391>.