Reduced Order Optimal Control Using Proper Orthogonal Decomposition Sensitivities

Tuğba Akman and Bülent Karasözen

Abstract In general, reduced-order model (ROM) solutions obtained using proper orthogonal decomposition (POD) at a single parameter cannot approximate the solutions at other parameter values accurately. In this paper, parameter sensitivity analysis is performed for POD reduced order optimal control problems (OCPs) governed by linear diffusion-convection-reaction equations. The OCP is discretized in space and time by discontinuous Galerkin (dG) finite elements. We apply two techniques, extrapolating and expanding the POD basis, to assess the accuracy of the reduced solutions for a range of parameters. Numerical results are presented to demonstrate the performance of these techniques to analyze the sensitivity of the OCP with respect to the ratio of the convection to the diffusion terms.

1 Introduction

Optimal control problems for nonlinear and time-dependent partial differential equations (PDEs) depending on a set of parameters are very time consuming. To overcome this, in the last years, POD-ROMs are applied to optimal control of PDEs (see for example [4]). The POD is based on projecting the dynamical system onto subspaces of basis elements using the snapshots computed by finite elements. The finite element solutions are not correlated to the physical properties of the system they approximate, whereas the POD bases express the characteristics of the solutions better. Besides POD, reduced basis methods are also used to obtain efficient ROM solutions for parameterized PDEs (see for example [6]). When ROMs should approximate solutions for a wide range of parameters, the cost of basis selection increases because full data are required. In recent years, sensitivity analysis has been used in the POD basis selection process for fluid dynamics [3]. They rely on the continuous or discrete sensitivities, baseline or reference POD modes and their derivatives with respect to parameters. In this

T. Akman (✉) • B. Karasözen
Department of Mathematics, Institute of Applied Mathematics, Middle East Technical University, 06800 Ankara, Turkey
e-mail: takman@metu.edu.tr; bulent@metu.edu.tr

A. Abdulle et al. (eds.), *Numerical Mathematics and Advanced Applications - ENUMATH 2013*, Lecture Notes in Computational Science and Engineering 103, DOI 10.1007/978-3-319-10705-9_40

work, we extend the parameter sensitivity analysis in [3] to time dependent OCPs constrained by linear diffusion-convection-reaction equations. We compute two new POD bases by extrapolating and expanding the baseline POD basis to assess the accuracy of the reduced solutions for a range of parameters. The optimality system is discretized using space-time dG method. DG time discretization schemes combined with the symmetric interior penalty (SIPG) method in space have the pleasant property that discretization and optimization commute. In addition, dG time-stepping methods require less regularity compared to the finite difference schemes in time [7, Chap. 7].

The paper is organized as follows: In Sect. 2, we give the optimality system for the OCP governed by the unsteady diffusion-convection equation. The fully-discrete optimality system using the space-time dG is given in Sect. 3. The POD-ROM for the OCP and the derivation of POD sensitivities are presented in Sect. 4. Numerical results for an OCP with interior and boundary layers are discussed in Sect. 5.

2 The Optimal Control Problem

We consider the following distributed OCP by the unsteady diffusion-convection-reaction equation without control constraints

$$\begin{aligned} \underset{u\in L^2(0,T;L^2(\Omega))}{\text{minimize}} \quad & J(y,u) := \frac{1}{2}\int_0^T \left(\|y-y_d\|^2_{L^2(\Omega)} + \alpha\, \|u\|^2_{L^2(\Omega)} \right) dt, \\ \text{subject to } & \partial_t y - \epsilon\Delta y + \boldsymbol{\beta}\cdot\nabla y + ry = f + u \quad (x,t)\in\Omega\times(0,T], \\ & y(x,t) = 0 \qquad (x,t)\in\partial\Omega\times[0,T], \\ & y(x,0) = y_0(x) \qquad x\in\Omega, \end{aligned} \tag{1}$$

where Ω is a bounded open, convex domain in $\mathbb{R}^2$ with a Lipschitz boundary $\partial\Omega$ and $I = (0,T]$ is the time interval, $f, y_d \in L^2(0,T;L^2(\Omega))$, $y_0(x)\in H_0^1(\Omega)$, $r\in L^\infty(\Omega)$, $\boldsymbol{\beta}\in (W^{1,\infty}(\Omega))^2$ are given functions and $\epsilon,\alpha > 0$ are given scalars. The velocity field $\boldsymbol{\beta}$ does not depend on time and satisfies the incompressibility condition, i.e. $\nabla\cdot\beta = 0$.

In order to write the variational formulation of the problem, we define the bilinear forms $a(y,v) = \int_\Omega (\epsilon\nabla y\cdot\nabla v + \boldsymbol{\beta}\cdot\nabla y v + ryv)\, dx$, $(u,v) = \int_\Omega uv\, dx$, the state and the test space as $Y = V = H_0^1(\Omega)$, $\forall t\in(0,T]$. It is well known that the pair $(y,u)\in H^1(0,T;L^2(\Omega))\cap L^2(0,T;H_0^1(\Omega))\times L^2(0,T;L^2(\Omega))$ is the unique solution of the optimal control problem if and only if there is an adjoint $p\in H^1(0,T;L^2(\Omega))\cap L^2(0,T;H_0^1(\Omega))$ such that (y,u,p) satisfy the following optimality system [8]

$$\begin{aligned} (\partial_t y, v) + a(y,v) &= (f+u, v) \quad \forall v\in V, \quad y(x,0) = y_0, \\ -(\partial_t p, \psi) + a(\psi, p) &= -(y-y_d, \psi) \quad \forall\psi\in V, \quad p(x,T) = 0, \\ \alpha u &= p. \end{aligned} \tag{2}$$

3 Space-Time Discretization of the Optimal Control Problem

Let $\{\mathscr{T}_h\}_h$ be a family of shape regular meshes such that $\overline{\Omega} = \cup_{K \in \mathscr{T}_h} \overline{K}$, $K_i \cap K_j = \emptyset$ for $K_i, K_j \in \mathscr{T}_h$, $i \neq j$. We use discontinuous piecewise finite element space $V_h = \left\{ y \in L^2(\Omega) : y\,|_K \in \mathbb{P}^1(K) \quad \forall K \in \mathscr{T}_h \right\}$ for the control, state and adjoint. Here, $\mathbb{P}^1(K)$ denotes the set of all polynomials on $K \in \mathscr{T}_h$ of degree 1. The diffusion term is discretized by the SIPG method and the convection term is discretized by upwinding [2]. Then, the semi-discrete state equation is given as in the study [1]

$$(\partial_t y_h, v_h) + a_h(y_h, v_h) + b_h(u_h, v_h) = (f_h, v_h) \quad \forall v_h \in V_h, \quad t \in (0, T].$$

For time discretization, we also use dG method. Let $0 = t_0 < t_1 < \cdots < t_N = T$ be a subdivision of $I = (0, T)$ with time intervals $I_n = (t_{n-1}, t_n]$ and time steps $k_n = t_n - t_{n-1}$ for $n = 1, \ldots, N$ and $k = \max_{1 \le n \le N} k_n$. We define the space-time finite element space of piecewise discontinuous functions for test function, state, control and adjoint as

$$V_h^k = \left\{ v \in L^2(0, T; L^2(\Omega)) : v|_{I_m} = \sum_{s=0}^{q} t^s \phi_s, t \in I_m, \phi_s \in V_h, m = 1, \ldots, N \right\}.$$

We use dG(0) method, i.e. $q = 0$, where the approximating polynomials are piecewise constant in time. We define $y_n = y_{hk}|_{I_n}$, $p_n = p_{hk}|_{I_n}$, $u_n = u_{hk}|_{I_n}$ for $n = 1, \cdots, N$, $y_{hk,0}^- = y_0$, $p_{hk,N}^+ = 0$. Then, the fully-discrete state and the adjoint equation are written as

$$\begin{aligned}
(M + kA^s) y_n &= M y_{n-1} + \frac{k}{2}(f_n + f_{n-1}) + \frac{k}{2} M(u_n + u_{n-1}), \\
(M + kA^a) p_{n-1} &= M p_n - \frac{k}{2} M(y_n + y_{n-1}) + \frac{k}{2}(y_n^d + y_{n-1}^d),
\end{aligned}$$

where M is the mass matrix and A^s, A^a are the stiffness matrices for the state $a_h(y_h, v_h)$ and adjoint equations $a_h(v_h, p_h)$, respectively. We note that the resulting scheme is a variant of the backward Euler method where the temporal terms on the right-hand side of (2) are computed by trapezoidal rule [7, Chap. 7].

4 Reduced-Order Modelling Using POD

In this section, we briefly explain the POD method. Let the matrix W be a real-valued $M \times N$ matrix of rank $d \le \min(M, N)$ representing the snapshot data. We introduce the correlation matrix $K = \tilde{W}^T \tilde{W}$ with $\tilde{W} = M^{1/2} W$. Then, we compute the coefficients of a POD basis of rank l using the eigenvalue decomposition (EVD) of K as follows

$$\Psi_{:,j} = W\tilde{V}_{:,j}/\sqrt{\lambda_j}, \quad j = 1, \cdots, l,$$

where $\tilde{V}_{:,j}$ is the j-th eigenvector of K and λ_j is the associated eigenvalue. On the other hand, the singular value decomposition (SVD) of the matrix $\tilde{W} = U\Sigma V^T$ can also be used. The POD basis coefficients are computed by solving the linear system $(M^{1/2})^T\Psi_{:,l} = U^l$, with the first l columns of U, for $\Psi_{:,l}$. Then, the l POD basis functions are written as a linear combination of the finite element basis functions, $\psi_j(x) = \sum_{i=1}^{M} \Psi_{ij}\varphi_i(x),\ j = 1, \ldots, l.$

In general, the POD basis generated via the snapshots depending on a parameter μ_0 cannot capture the dynamics of the perturbed problem associated to $\mu = \mu_0 + \Delta\mu$. Motivated by the study of fluid flow equations using POD-ROMs [3], POD sensitivities can be used to enrich the low-dimensional space for a wider range of parameters. In order to derive POD sensitivities, the sensitivity of the snapshot set is required. The sensitivity of a term is defined as the derivative of that term with respect to a quantity of interest. In this study, we are interested in the sensitivities with respect to μ corresponding to the ratio $\mu = |\boldsymbol{\beta}|/\epsilon$ in the OCP (1). For the computation of the sensitivities, we compare two different approaches: the continuous sensitivity equation (*CSE*) and finite-difference (*FD*) approximation. In *CSE* approach, state, adjoint and control are assumed to be differentiable with respect to μ. The subscript μ denotes the derivative with respect to μ. Then, with the sensitivities $s = y_\mu, q = p_\mu, \nu = u_\mu$, we derive another optimality system depending on s, q and ν,

$$\begin{aligned} (\partial_t s, v) + a(s, v) &= (f_\mu + \nu, v) - (\nabla y, \nabla v), \quad s(x,0) = (y_0)_\mu, \\ -(\partial_t q, \psi) + a(\psi, q) &= -(s - y_\mu^d, \psi) - (\nabla p, \nabla \psi), \quad q(x,T) = 0, \\ \alpha\nu &= q. \end{aligned} \tag{3}$$

We note that the sensitivity equations are always linear, so *CSE* method would be especially promising for nonlinear problems. The sensitivities s, q and μ can be computed either by inserting the solution of the state and adjoint to the right-hand side of (3) or solving the systems arising from (2) and (3) simultaneously. We use the second approach. In *FD* method, particularly for the centred difference scheme, the solution of the perturbed optimal control problem is required, i.e. depending on $\mu = \mu_0 \pm \Delta\mu$. Then, the sensitivity of the state can be computed via the centred difference as follows

$$y_\mu(\mu_0) \approx \frac{y(\mu_0 + \Delta\mu) - y(\mu_0 - \Delta\mu)}{2\Delta\mu}. \tag{4}$$

We treat each POD mode as a function of both space and the parameter, i.e. $\psi = \psi(x,\mu)$. In order to find POD sensitivities, we differentiate $(M^{1/2})^T\Psi = U^l$ with respect to μ and then solve the resulting equation for Ψ_μ. Then, the sensitivities of the l POD basis functions, namely ψ_μ, are written as a linear combination of the finite element basis functions, $(\psi_j)_\mu = \sum_{i=1}^{M} (\Psi_{ij})_\mu \varphi_i(x),\ j = 1,\dots,l$.

We have taken the same range of parameters as in [3]. For larger parameter variations, the applicability of this approach might not be useful, because the sensitivities are based on the asymptotic expansion of μ in (4).

The connection between the state and the POD sensitivities is realised through the relation

$$U^l_\mu = (\tilde{W} V^l \Sigma^\dagger)_\mu = \tilde{W}_\mu V^l \Sigma^\dagger + \tilde{W} V^l_\mu \Sigma^\dagger + \tilde{W} V^l \Sigma^\dagger_\mu.$$

For the computation of V^l_μ and $\Sigma^\dagger_\mu$, we consider the equation $B = A^T A$ which leads to the following eigenvalue problem $BV^k = V^k\lambda^k$ with the kth column of V.

After differentiation, one obtains

$$(V^k)^T(B_\mu - \lambda^k_\mu I)V^k = 0. \tag{5}$$

Equation (5) is solved in the least-squares sense and we denote one particular solution by s^k. $\Sigma^\dagger_\mu$ is computed using the relation $\sigma^2 = \lambda$. For details, we refer to [3, Sec. 3.2].

We use the sensitivity information in two ways, i.e. extrapolating POD (ExtPOD) and expanding POD (ExpPOD) basis. In ExtPOD, the POD basis depending on μ is written using the first-order Taylors expansion as follows

$$\psi(x,\mu) = \psi(x,\mu_0) + \Delta\mu \frac{\partial\psi}{\partial\mu}(x,\mu_0) + \mathcal{O}(\Delta\mu^2).$$

In ExpPOD, the POD basis sensitivities are also added to the original POD basis as $[\psi_1,\dots,\psi_l,(\psi_1)_\mu,\dots,(\psi_l)_\mu]$ and the reduced order solution is written as

$$y^r_h(x,t) = \sum_{j=1}^{l} y^r_j(t)\psi_j(x) + \sum_{j=l+1}^{2l} y^r_j(t)(\psi_j(x))_\mu,$$

where the dimension of the reduced basis is doubled.

5 Numerical Results

We consider the optimal control problem with

$$Q = (0,1] \times \Omega, \ \Omega = (0,1)^2, \ \epsilon = 10^{-2}, \ \boldsymbol{\beta} = \frac{1}{\sqrt{2}}(1,1)^T, \ r = 1, \ \alpha = 1.$$

The source function f, the desired state y_d and the initial condition y_0 are computed from the optimality system (2) using the following exact solutions of the state and control, respectively,

$$y(x,t) = (1-e^{-t})xye^{-\frac{1-x}{\epsilon}-1}e^{-\frac{1-y}{\epsilon}-1},$$
$$u(x,t) = (1-t)xy(1-x)(1-y)\arctan\left(\frac{x-y}{\epsilon}\right).$$

We observe that the state contains boundary layers along $x = 1$ and $y = 1$, while the control exhibit an interior layer along $x = y$ of the width ϵ. The full problem is solved for $\Delta x = 1/40$, $\Delta t = 1/60$. The conjugate gradient method is used in the optimization step. The error between the full and reduced solution of the control is measured with respect to $L^2(0,T;L^2(\Omega))$.

We choose the parameter range for the ratio $\mu =\mid \boldsymbol{\beta} \mid /\epsilon$ as $1/\epsilon = 80:5:120$. We compute l POD basis functions associated to the nominal diffusion parameter, i.e. $\epsilon = 1/100$, and compare the resulting error with the ExtPOD and ExpPOD basis. Three different snapshot sets for W are used to generate the POD basis functions, namely state Y, adjoint P and the combination of them $Y \cup P$, as in [5]. The state, adjoint and the control are written in terms of the same POD basis functions associated to W and then the optimality system is projected onto the low-dimensional subspace.

We choose the number of POD basis functions, namely l, according to the relative information content, that is, the ratio of the modelled energy to the total energy contained in the system $\mathscr{E}(l) = \sum_{i=1}^{l} \lambda_i / \sum_{i=1}^{d} \lambda_i$. It is fixed up to $100(1-\gamma)\%$ by keeping the most energetic POD modes. In this study, we choose 10 POD basis functions setting $\gamma = 10^{-2}$.

Because the velocity field is constant in our example, we proceed with the diffusion term to calculate the sensitivities. In Fig. 1, we present the decrease of the first 15 eigenvalues of the snapshot ensemble $Y, P, Y \cup P$ on the left and their sensitivities $Y_\mu, P_\mu, Y_\mu \cup P_\mu$ on the right. The sensitivities are computed using the centered *FD* quotient and *CSE* method. We observe that *FD* and *CSE* methods yield almost the same eigenvalues. We note that the eigenvalues and their sensitivities are decreasing.

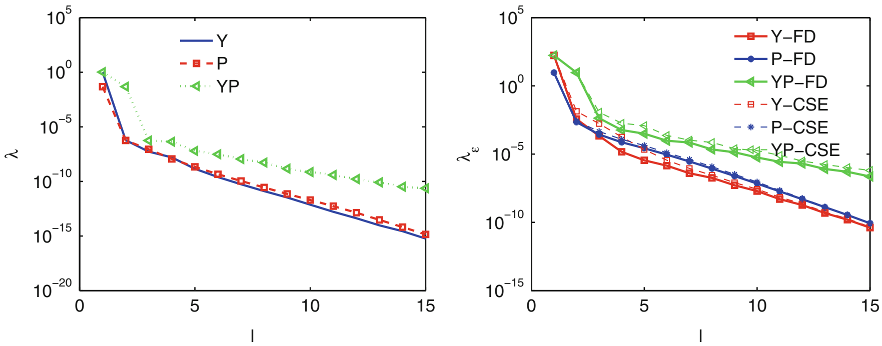

Fig. 1 Eigenvalues(*left*) and their sensitivities(*right*)

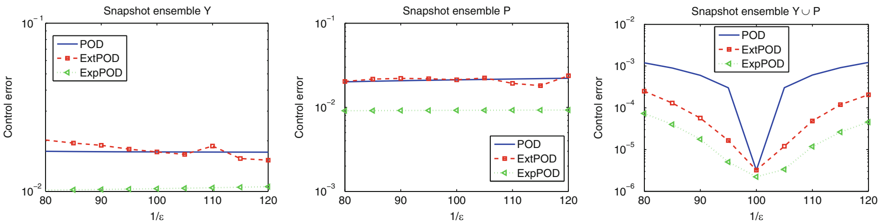

Fig. 2 Error versus parameter for 10 POD basis functions generated with the snapshot set Y (*left*), P (*middle*) and $Y \cup P$ (*right*)

In Fig. 2, we present the error for the control with respect to ϵ with 10 POD bases functions. The control approximated with the POD bases generated from the state solution is poor because the characteristics of the control are totally different from the state solution. The inclusion of the adjoint information in W improves the performance of the method, because the relation between the adjoint and the control is determined through the optimality condition (2). In addition, a good approximation to the control influences the state solution directly due to acting on the right-hand side of the state equation. The figures on the left and in the middle indicate that the snapshot sets Y and P cannot reveal the sensitivity of the control with respect to ϵ. Although ExpPOD gives the smallest error, it is too large for the reduced solution to be accepted. The solution plotted in the right of Fig. 2 is obtained using the snapshot set $Y \cup P$ and it reveals the sensitivity of the problem with respect to ϵ. As we move away the parameter, the error in the reduced solution increases. For the reduced solution of the perturbed problem, ExpPOD basis generated with the snapshot ensemble $Y \cup P$ is the most promising basis among POD and ExtPOD.

In Fig. 3, we present the error for the control with respect to increasing number of POD basis functions by taking $\epsilon = 1/120$. The figure on the left has been obtained with the state snapshots Y and shows that the error decays slowly and is oscillating due to a poor approximation to the control is used. The figure in the middle depicts the error obtained by the POD basis generated with the adjoint snapshot set P.

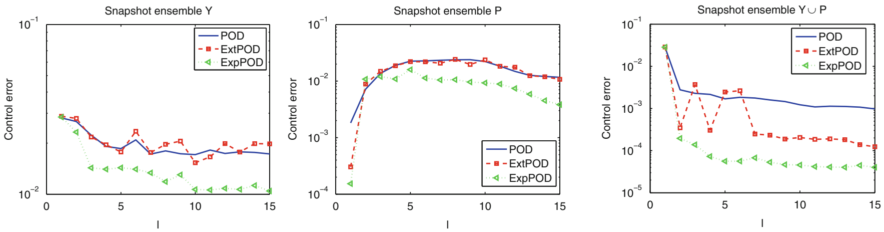

Fig. 3 Error versus the number of POD basis functions for $\epsilon = 1/120$ generated with the snapshot set Y (*left*), P (*middle*) and $Y \cup P$ (*right*)

Although the error for the first POD mode is around 10^{-4}, error increases up to 10^{-2}, which is not usual when the number POD basis functions are increased. The figure on the right shows that snapshot ensemble $Y \cup P$ leads to the smallest error. Moreover, the error for ExtPOD oscillates until the 6th POD mode and then surpasses the error in the nominal POD. However, the benefit of using ExpPOD is revealed at the most, because the error in the nominal POD basis is improved almost 2 digits. In addition, the decay of the errors for $Y \cup P$ is much faster than the one obtained with Y or P using a smaller number of POD basis functions.

We observe that although the eigenvalues of the snapshot sets Y and P decreases as shown in Fig. 1, the quality of the reduced-order control obtained using Y or P is not sufficient. The state and adjoint snapshots might be a good choice if associated POD basis is used to approximate the state and adjoint independently. The POD basis generated via the snapshot ensemble $Y \cup P$, containing information about both state and adjoint, give more accurate reduced order solutions and capture the sensitivity of the problem better. For the perturbed problem, expanding the POD basis increases accuracy without solving the nominal problem for each parameter.

Acknowledgements This work has been supported by METU BAP-07-05-2013-003.

References

1. 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), 703–729 (2014)
2. 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), 1749–1779 (2001/2002)
3. A. Hay, J.T. Borggaard, D. Pelletier, Local improvements to reduced-order models using sensitivity analysis of the proper orthogonal decomposition. J. Fluid Mech. **629**, 41–72 (2009)
4. M. Hinze, S. Volkwein, Proper orthogonal decomposition surrogate models for nonlinear dynamical systems: error estimates and suboptimal control, in *Dimension Reduction of Large-Scale Systems*, ed. by P. Benner, V. Mehrmann, D.C. Sorensen. Lecture Notes in Computational Science and Engineering, vol. 45 (Springer, Berlin, 2005), pp. 261–306

5. M. Hinze, S. Volkwein, Error estimates for abstract linear-quadratic optimal control problems using proper orthogonal decomposition. Comput. Optim. Appl. **39**(3), 319–345 (2008)
6. A. Quarteroni, G. Rozza, A. Manzoni, Certified reduced basis approximation for parametrized partial differential equations and applications. J. Math. Ind. **1**, Art. 3, 44 (2011)
7. V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*. Springer Series in Computational Mathematics, vol. 25, 2nd edn. (Springer, Berlin, 2006)
8. F. Tröltzsch, *Optimal Control of Partial Differential Equations: Theory, Methods and Applications*. Graduate Studies in Mathematics, vol. 112 (American Mathematical Society, Providence, 2010)