

## OPTIMAL SNAPSHOT LOCATION FOR COMPUTING POD BASIS FUNCTIONS\*

KARL KUNISCH<sup>1</sup> AND STEFAN VOLKWEIN<sup>2</sup>

**Abstract.** The construction of reduced order models for dynamical systems using proper orthogonal decomposition (POD) is based on the information contained in so-called snapshots. These provide the spatial distribution of the dynamical system at discrete time instances. This work is devoted to optimizing the choice of these time instances in such a manner that the error between the POD-solution and the trajectory of the dynamical system is minimized. First and second order optimality systems are given. Numerical examples illustrate that the proposed criterion is sensitive with respect to the choice of the time instances and further they demonstrate the feasibility of the method in determining optimal snapshot locations for concrete diffusion equations.

**Mathematics Subject Classification.** 49J20, 49K20, 49M15, 90C53.

Received September 10, 2008.

Published online February 4, 2010.

### 1. INTRODUCTION

Proper orthogonal decomposition (POD) is one of the most popular techniques for model reduction. It was first used for signal analysis and pattern recognition, consequently in the context of dynamical systems, and more recently, also for optimal control and inverse problems [6,12,13,15,18]. The snapshot version of POD assumes the availability of the states  $y(t_j)$  of the dynamical system

$$\begin{cases} \dot{y}(t) = f(t, y(t)) & \text{for } t \in (0, T], \\ y(0) = y_0 \end{cases} \quad (1.1)$$

at times  $\{t_j\}_{j=0}^m$ . Later in the paper we shall assume that (1.1) is linear. This assumption, however, is not essential for the present work, since only the regularity properties of the trajectory  $t \mapsto y(t)$  will enter the analysis. Given the snapshot, and the dimension  $\ell$  of the reduced-order space, the POD-basis is determined

---

*Keywords and phrases.* Proper orthogonal decomposition, optimal snapshot locations, first and second order optimality conditions.

\* Research in part supported through SFB Mathematical Optimization and Applications in Biomedical Sciences, Fonds zur Förderung der wissenschaftlichen Forschung, Austria.

<sup>1</sup> University of Graz, Institute for Mathematics and Scientific Computing, Heinrichstrasse 36, 8010 Graz, Austria.

[karl.kunisch@uni-graz.at](mailto:karl.kunisch@uni-graz.at)

<sup>2</sup> University of Constance, Department for Mathematics and Statistics, Universitätsstraße 10, 78464 Konstanz, Germany.

[stefan.volkwein@uni-konstanz.de](mailto:stefan.volkwein@uni-konstanz.de)

as the solution to

$$\begin{cases} \min_{\{\psi_i\}_{i=1}^{\ell}} \sum_{j=0}^m \left\| y(t_j) - \sum_{i=1}^{\ell} \langle y(t_j), \psi_i \rangle \psi_i \right\|^2 \\ \text{subject to } \langle \psi_i, \psi_j \rangle = \delta_{ij} \text{ for } 1 \leq i, j \leq \ell. \end{cases} \quad (1.2)$$

Thus, snapshot POD consists in choosing an orthonormal basis such that the mean square error between the elements  $\{y(t_j)\}_{j=0}^m$  and the corresponding  $\ell$ -th partial sum is minimized on average.

The norm and inner product in (1.2) depend on the specific setting of (1.1). It is well-known, see, *e.g.*, [1], that the solution to (1.2) is given by the first  $\ell$  eigenvectors of the selfadjoint operator

$$\mathcal{R}\psi = \sum_{j=0}^m \langle y(t_j), \psi \rangle y(t_j),$$

where the eigenvalues are ordered according to  $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$ . Once the basis is determined by (1.2) the POD-approximation  $y^\ell$  to (1.1) is obtained by means of a Galerkin procedure with respect to the basis  $\{\psi_i\}_{i=1}^{\ell}$ .

In this work we assume that additional snapshots may be added and focus on the question, where to allocate them. As criterion for optimal placement of additional snapshots at time instances  $\bar{t} = (\bar{t}_1, \dots, \bar{t}_{\bar{k}})$  with  $0 \leq \bar{t}_k \leq T$ ,  $k = 1, \dots, \bar{k}$ , we propose to solve

$$\min_{0 \leq \bar{t}_1, \dots, \bar{t}_{\bar{k}} \leq T} \int_0^T \|y(t) - y^\ell(t)\|^2 dt, \quad (1.3)$$

where we denote by  $y^\ell$  the POD-Galerkin solution in the  $\ell$ -dimensional POD-space span  $\{\psi_1(\bar{t}), \dots, \psi_\ell(\bar{t})\}$  with  $\psi_i(\bar{t})$  generated from the old combined with the new snapshots, *i.e.*, we replace  $\mathcal{R}$  by

$$\mathcal{R}(\bar{t})\psi = \sum_{j=0}^m \langle y(t_j), \psi \rangle y(t_j) + \sum_{k=1}^{\bar{k}} \langle y(t_k), \psi \rangle y(t_k),$$

and (1.3) becomes an optimization problem subject to the eigenvalue constraint

$$\mathcal{R}(\bar{t})\psi = \lambda\psi.$$

Alternative criteria could be considered. For example

$$\max_{0 \leq \bar{t}_1, \dots, \bar{t}_{\bar{k}} \leq T} \sum_{i=1}^{\ell} \lambda_i(\bar{t}), \quad (1.4)$$

*i.e.*, the “energy” captured in the first  $\ell$ -modes of the POD subspace is maximized by properly allocating the snapshots. This criterion is motivated by the fact that in the context of POD for fluid mechanical problems the sum over all eigenvalues of  $\mathcal{R}$  is referred to as the energy of the dynamical system.

Note that no precautions are made in either (1.3) or (1.4) to avoid multiple appearance of a snapshot. In fact, this would simply imply that a specific snapshot location should be given a higher weight than others.

Concerning weights on snapshots, the functional in (1.2) can be considered to be a finite difference approximation to

$$\int_0^T \left\| y(t) - \sum_{i=1}^{\ell} \langle y(t), \psi_i \rangle \psi_i \right\|^2 dt. \quad (1.5)$$

We may then pose the question of optimal discretization of the integral in (1.5) which, in turn, is related to optimal weights on snapshot locations. Further we can allow weights in (1.5), replacing  $dt$  by  $d\eta(t)$  and pose the question of the optimal choice of  $\eta$  for the purpose of model reduction.

While our approach is presented here in the context of choosing optimal snapshots in evolution equations, a similar strategy is applicable in the context of parameter dependent systems.

Let us, very briefly, mention some related issues of interest. In [3,5] the situation of missing snapshot data was investigated and gappy POD was introduced for their reconstruction. In [17], the authors presented a sensitivity for POD approximations of finite-dimensional dynamical system. For the case of linear dynamics, POD-model reduction is closely related to the well-known balanced truncation method [14,19]. An important alternative to POD model reduction is given by reduced basis approximations, we refer to [7,10,18,20] and the references given there. In [7,20] a reduced model is constructed for a parameter dependent family of large scale problems by an iterative procedure that adds new basis variables on the basis of a greedy algorithm. Utilizing the structure in which the parameters enter into the system this can be achieved in a computationally efficient manner. In [4,22] a reduced order model is obtained by minimizing the difference of the outputs of the large scale and the reduced order systems, possibly corresponding to a family of different operating stages, with respect to family of orthogonal bases. In practice this family is parameterized by linear combination of snapshots and the optimization is carried out with respect to the expansion coefficients as optimization variables, which enter into the reduced system, which constitutes a constraint in the optimization problem, in a bilinear fashion. Our approach is different in that the optimization is carried out with respect to the snapshots themselves or, more specifically, with respect to the snapshot times. Thus in our approach, “new” snapshots are added during the optimization procedure, whereas in [4,22] the weights of the preexisting snapshots are optimized.

We also point out the interesting thesis [2], where a model reduction procedure is sought of a class for a family of models corresponding to different operating stages. Here the basis is updated – for example by the means of POD – on the basis of the results of a greedy algorithm which, at each iteration level, determines that parameter, characterizing a specific member of the family of original unreduced systems, at which the output of the reduced model and the desired output of the full model is maximal.

The contents of the paper is the following. Section 2 contains the precise problem formulation leading to a mathematical programming problem. Existence of a solution, first-order optimality conditions and a quasi-Newton algorithm are established. Section 3 is devoted to second-order information for (1.2). Numerical investigations are presented in Section 4. They highlight the feasibility of the proposed approach and the sensitivity of the relevant quantities with respect to  $\bar{t}$ .

## 2. PROBLEM FORMULATION AND OPTIMALITY CONDITIONS

We consider the linear dynamical system

$$\dot{y}(t) = Ay(t) + f(t) \quad \text{for } t \in (0, T] \quad (2.1a)$$

$$y(0) = y_o, \quad (2.1b)$$

where  $T > 0$  holds,  $V$  and  $H$  are separable real Hilbert spaces, with  $V$  dense and compact in  $H$ , and  $V \subset H \equiv H^* \subset V^*$  is a Gelfand triple, and  $A : V \rightarrow V^*$  denotes a bounded linear operator. We suppose that  $f \in C([0, T], H)$  and  $y_o \in \text{dom}(A)$ , where  $\text{dom}(A) = \{\varphi \in H : A\varphi \in H\}$  denotes the domain of the operator  $A$ .

Throughout this paper we assume that

$$\text{Problem (2.1) has a unique solution in } C^1([0, T], H) \cap C([0, T], \text{dom}(A)). \quad (\mathbf{H1})$$

### Example 2.1.

- (a) For  $V = H = \mathbb{R}^n$  and  $A \in \mathbb{R}^{n \times n}$  equation (2.1) admits a unique solution in  $C^1([0, T], \mathbb{R}^n)$  and thus  $(\mathbf{H1})$  holds.
- (b) If  $A \in \mathcal{L}(V, V^*)$  is coercive, *i.e.*, there exists  $\alpha > 0$  such that

$$\langle Av, v \rangle_{V^*, V} \geq \alpha \|v\|_V^2 \quad \text{for all } v \in V,$$

and  $y_o \in \text{dom}(A)$ , then (2.1) admits a unique solution  $y$  in the space  $C([0, T], \text{dom}(A)) \cap C^1([0, T], H) \cap W^{2,2}(0, T; V^*)$ , see [21], p. 72.

The POD basis is determined from knowledge of the solution  $y$  at discrete time instances. We distinguish here between a predefined set chosen on a uniform grid and additional ones which will be determined in an optimal way. Let  $t_j = j\Delta t$ ,  $0 \leq j \leq m$ , be a fixed time grid with uniform step size  $\Delta t = T/m$ . By (H1) we have  $y_j = y(t_j) \in V$  for  $0 \leq j \leq m$ . The snapshots at the new time instances  $\bar{t}_k \in [0, T]$ ,  $1 \leq k \leq \bar{k}$ , are denoted by  $\bar{y}_k = y(\bar{t}_k)$ ,  $1 \leq k \leq \bar{k}$ , and we set

$$\mathcal{V} = \text{span} \{y_0, \dots, y_m, \bar{y}_1, \dots, \bar{y}_{\bar{k}}\} \subset V$$

endowed with the same topology as in  $V$ . Let  $\ell \in \{1, \dots, \dim \mathcal{V}\}$  be the number of POD basis functions.

For simplicity we henceforth denote by  $\bar{t}$  the vector  $(\bar{t}_1, \dots, \bar{t}_{\bar{k}}) \in \mathbb{R}^{\bar{k}}$  and define the bounded linear selfadjoint operator  $\mathcal{R}(\bar{t}) : H \rightarrow H$  by

$$\mathcal{R}(\bar{t})\psi = \sum_{j=0}^m \langle y_j, \psi \rangle_H y_j + \sum_{k=1}^{\bar{k}} \langle \bar{y}_k, \psi \rangle_H \bar{y}_k. \tag{2.2}$$

Recall that the POD basis functions  $\{\psi_i\}_{i=1}^{\ell}$  are eigenvectors of  $\mathcal{R}(\bar{t})$ , *i.e.*,

$$\mathcal{R}(\bar{t})\psi_i = \lambda_i \psi_i, \tag{2.3}$$

see, *e.g.*, [1], where the eigenvalues are ordered and for the simplicity of presentation are supposed to be simple:

$$\lambda_1 > \lambda_2 > \dots > \lambda_\ell > 0. \tag{2.4}$$

Thus it follows that  $\langle \psi_i, \psi_j \rangle_H = 0$  for  $1 \leq i \leq \ell$  with  $i \neq j$  and we assume that

$$\|\psi_i\|_H = 1 \quad \text{for } 1 \leq i \leq \ell. \tag{2.5}$$

The reduced-order models are based on a Galerkin ansatz with respect to the POD basis  $\{\psi_i\}_{i=1}^{\ell}$  of rank  $\ell$ . We approximate the trajectory  $y(t)$  by the Galerkin ansatz

$$y^\ell(t) = \sum_{j=1}^{\ell} \eta_j(t) \psi_j \in V \quad \text{for } t \in [0, T]. \tag{2.6}$$

Then, the Galerkin projection of (2.1) is given by

$$\begin{aligned} \langle \dot{y}^\ell(t), \psi_i \rangle_H &= \langle Ay^\ell(t), \psi_i \rangle_{V^*, V} + \langle f(t), \psi_i \rangle_H, & t \in (0, T], \quad 1 \leq i \leq \ell, \\ \langle y^\ell(0), \psi_i \rangle_H &= \langle y_o, \psi_i \rangle_H, & 1 \leq i \leq \ell. \end{aligned}$$

Introducing the vectors

$$\eta(t) = \begin{pmatrix} \eta_1(t) \\ \vdots \\ \eta_\ell(t) \end{pmatrix}, \quad f(t) = \begin{pmatrix} \langle f(t), \psi_1 \rangle_H \\ \vdots \\ \langle f(t), \psi_\ell \rangle_H \end{pmatrix}, \quad \eta_o = \begin{pmatrix} \langle y_o, \psi_1 \rangle_H \\ \vdots \\ \langle y_o, \psi_\ell \rangle_H \end{pmatrix}$$

in  $\mathbb{R}^\ell$  and the matrix

$$A^\ell = ((A_{ij}^\ell)) \in \mathbb{R}^{\ell \times \ell} \quad \text{with } A_{ij}^\ell = \langle A\psi_j, \psi_i \rangle_{V^*, V}$$

the Galerkin approximation of (2.1) can be expressed as

$$\dot{\eta}(t) = A^\ell \eta(t) + f(t) \quad \text{for } t \in (0, T], \quad (2.7a)$$

$$\eta(0) = \eta_0. \quad (2.7b)$$

Next we formulate the optimization problem. For that purpose we define the spaces

$$\mathcal{X} = H^1(0, T; \mathbb{R}^\ell) \times \mathbb{R}^{\bar{k}} \times H^\ell \times \mathbb{R}^\ell \quad \text{and} \quad \mathcal{Y} = L^2(0, T; \mathbb{R}^\ell) \times \mathbb{R}^\ell \times H^\ell \times \mathbb{R}^\ell,$$

supplied with the common product topologies, where we set  $H^\ell = \bigotimes_{i=1}^\ell H$ . The cost functional  $J : \mathcal{X} \rightarrow [0, \infty)$  quantifies the difference between the trajectory  $y$  of (2.1) and its POD-Galerkin approximation (2.7):

$$J(\eta, \bar{t}, \psi, \lambda) = \frac{1}{2} \int_0^T \left\| y(t) - \sum_{i=1}^\ell \eta_i(t) \psi_i \right\|_H^2 dt.$$

The equality constraints are given by (2.3), (2.5), and (2.7). Therefore, we define the nonlinear mapping  $e = (e_1, e_2, e_3, e_4) : \mathcal{X} \rightarrow \mathcal{Y}$  by

$$\begin{aligned} e_1(\eta, \bar{t}, \psi, \lambda) &= \dot{\eta} - A^\ell \eta - f && \text{in } L^2(0, T; \mathbb{R}^\ell), \\ e_2(\eta, \bar{t}, \psi, \lambda) &= \eta(0) - \eta_0 && \text{in } \mathbb{R}^\ell, \\ e_3(\eta, \bar{t}, \psi, \lambda) &= \begin{pmatrix} (\mathcal{R}(\bar{t}) - \lambda_1) \psi_1 \\ \vdots \\ (\mathcal{R}(\bar{t}) - \lambda_\ell) \psi_\ell \end{pmatrix} && \text{in } H^\ell, \\ e_4(\eta, \bar{t}, \psi, \lambda) &= \begin{pmatrix} 1 - \|\psi_1\|_H^2 \\ \vdots \\ 1 - \|\psi_\ell\|_H^2 \end{pmatrix} && \text{in } \mathbb{R}^\ell \end{aligned}$$

for  $x = (\eta, \bar{t}, \psi, \lambda) \in \mathcal{X}$ . The inequality constraints  $0 \leq t_k \leq T$  for  $1 \leq k \leq \bar{k}$  are expressed by  $g(x) \leq 0$ , where  $g = (g_1, g_2) : \mathcal{X} \rightarrow \mathbb{R}^{\bar{k}} \times \mathbb{R}^{\bar{k}}$  is defined as

$$g_1(\eta, \bar{t}, \psi, \lambda) = -(\bar{t}_1, \dots, \bar{t}_{\bar{k}}) \quad \text{and} \quad g_2(\eta, \bar{t}, \psi, \lambda) = (\bar{t}_1 - T, \dots, \bar{t}_{\bar{k}} - T)$$

for  $x = (\eta, \bar{t}, \psi, \lambda) \in \mathcal{X}$ . Now the minimization problem is

$$\min J(x) \text{ subject to } x \in \mathcal{F}(\mathbf{P}), \quad (\mathbf{P})$$

where the feasible set for (P) is given by

$$\mathcal{F}(\mathbf{P}) = \{x = (\eta, \bar{t}, \psi, \lambda) \in \mathcal{X} \mid e(x) = 0 \text{ and } g(x) \leq 0\}.$$

The state variables for (P) are  $x = (\eta, \bar{t}, \psi, \lambda) \in \mathcal{X}$ , where  $\eta$  denotes the vector of modal coefficients in the POD Galerkin ansatz,  $\bar{t}$  are the new snapshot locations,  $\psi_1, \dots, \psi_\ell$  stands for the POD basis functions, and  $\lambda_1 > \lambda_2 > \dots > \lambda_\ell$  are the corresponding positive, distinct eigenvalues of  $\mathcal{R}$ . The adjoint variables  $z = (\mathfrak{p}, \mathfrak{p}_0, \mu, \eta) \in \mathcal{Y}$  are the Lagrange multipliers corresponding the constraints given by the Galerkin approximation (2.7), the spectral conditions (2.3), and normalization conditions (2.5).

**Proposition 2.2.** *Problem (P) admits a solution  $x^* = (\eta^*, \bar{t}^*, \psi^*, \lambda^*)$ .*

*Proof.* Let  $x^n = (\eta^n, \bar{t}^n, \psi^n, \lambda^n) \in \mathcal{X}$ ,  $n \in \mathbb{N}$ , be a minimizing sequence for **(P)**. Since  $\{\bar{t}^n\}_{n \in \mathbb{N}}$  is uniformly bounded, there exists  $\bar{t}^* \in \mathbb{R}^{\bar{k}}$  such that, on a subsequence,  $\lim_{n \rightarrow \infty} \bar{t}^n = \bar{t}^*$  and  $0 \leq \bar{t}_i^* \leq T$ . Consequently  $\mathcal{R}(\bar{t}^n) \rightarrow \mathcal{R}(\bar{t}^*)$  in operator norm in  $\mathcal{L}(H)$  and also in the generalized sense, [11], p. 206, *i.e.*,

$$\max[\delta(G(\mathcal{R}(\bar{t}^n)), G(\mathcal{R}(\bar{t}^*))), \delta(G(\mathcal{R}(\bar{t}^*)), G(\mathcal{R}(\bar{t}^n)))] \rightarrow 0 \quad \text{for } n \rightarrow \infty,$$

where  $G(\mathcal{R}(\bar{t}^*))$  stands for the graph of the operator  $\mathcal{R}(\bar{t}^*)$  and

$$\delta(M, N) = \sup_{\{u \in M: \|u\|_H=1\}} \text{dist}(u, N),$$

where  $M, N$  are closed linear manifolds in  $H$ . Perturbation results on spectra of operators imply the convergence of the associated eigenvalues and eigenvectors

$$\lim_{n \rightarrow \infty} \lambda_i^n = \lambda_i^* \quad \text{and} \quad \lim_{n \rightarrow \infty} \psi_i^n = \psi_i^* \quad \text{for } 1 \leq i \leq \ell,$$

where  $(\lambda_i^n, \psi_i^n)$  eigenvalue-eigenvector pairs of  $\mathcal{R}(\bar{t}^n)$  and  $(\lambda_i^*, \psi_i^*)$  eigenvalue-eigenvector pairs of  $\mathcal{R}(\bar{t}^*)$ ; see [11], p. 212–214, in particular, we have

$$\mathcal{R}(\bar{t}^*)\psi_i^* = \lambda_i^*\psi_i^* \quad \text{for } 1 \leq i \leq \ell.$$

It is simple to argue that  $\lim_{n \rightarrow \infty} A^\ell(\psi^n) = A^\ell(\psi^*)$  in  $\mathbb{R}^{\ell \times \ell}$ ,  $\lim_{n \rightarrow \infty} f(\psi^n) = f(\psi^*)$  in  $L^2(0, T; \mathbb{R}^\ell)$  and  $\lim_{n \rightarrow \infty} \eta_\circ(\psi^n) = \eta_\circ(\psi^*)$  in  $\mathbb{R}^l$ . Consequently,  $\eta^n \rightarrow \eta^* = \eta(\psi^*)$  in  $W^{1,2}(0, T; \mathbb{R}^\ell)$  as  $n \rightarrow \infty$ . We have now established that  $e_i(x^*) = 0$  for  $i = 1, \dots, 4$  and  $g_i(x^*) \leq 0$  for  $i = 1, 2$ , where  $x^* = (\eta^*, \bar{t}^*, \psi^*, \lambda^*)$ , and thus  $x^* \in \mathcal{F}(\mathbf{P})$ . Continuity of  $x \mapsto J(x)$  implies that  $x^*$  is a solution of **(P)**.  $\square$

For further reference it will be convenient to specify the derivatives of  $e$  at  $x = (\eta, \bar{t}, \psi, \lambda) \in \mathcal{X}$  in direction  $\delta x = (\delta\eta, \delta\bar{t}, \delta\psi, \delta\lambda) \in \mathcal{X}$ :

$$\begin{aligned} \nabla e_1(x)\delta x &= \dot{\delta}\eta - A^\ell \delta\eta - \delta A^\ell \eta - \delta f, \\ \nabla e_2(x)\delta x &= \delta\eta(0) - \delta\eta_\circ, \\ \nabla e_3(x)\delta x &= \begin{pmatrix} (\mathcal{R}(\bar{t}) - \lambda_1)\delta\psi_1 - \delta\lambda_1\psi_1 \\ \vdots \\ (\mathcal{R}(\bar{t}) - \lambda_\ell)\delta\psi_\ell - \delta\lambda_\ell\psi_\ell \end{pmatrix} \\ &\quad + \sum_{k=1}^{\bar{k}} (\delta\bar{t})_k \begin{pmatrix} \langle \dot{y}(\bar{t}_k), \psi_1 \rangle_H y(\bar{t}_k) + \langle y(\bar{t}_k), \psi_1 \rangle_H \dot{y}(\bar{t}_k) \\ \vdots \\ \langle \dot{y}(\bar{t}_k), \psi_\ell \rangle_H y(\bar{t}_k) + \langle y(\bar{t}_k), \psi_\ell \rangle_H \dot{y}(\bar{t}_k) \end{pmatrix}, \\ \nabla e_4(x)\delta x &= - \begin{pmatrix} 2 \langle \psi_1, \delta\psi_1 \rangle_H \\ \vdots \\ 2 \langle \psi_\ell, \delta\psi_\ell \rangle_H \end{pmatrix} \end{aligned}$$

where

$$\begin{aligned} \delta A^\ell \in \mathbb{R}^{\ell \times \ell} &\quad \text{with} \quad (\delta A^\ell)_{ij} = \langle A\delta\psi_j, \psi_i \rangle_{V^*, V} + \langle A\psi_j, \delta\psi_i \rangle_{V^*, V}, \\ \delta f(t) \in \mathbb{R}^\ell &\quad \text{with} \quad (\delta f(t))_i = \langle f(t), \delta\psi_i \rangle_H, \\ \delta\eta_\circ \in \mathbb{R}^\ell &\quad \text{with} \quad (\delta\eta_\circ)_i = \langle y_\circ(t), \delta\psi_i \rangle_H \end{aligned} \tag{2.8}$$

for  $1 \leq i, j \leq \ell$ .

**Remark 2.3.** Note that in  $\nabla e_3$  the terms  $\dot{y}(\bar{t}_k)$  are well defined due to (H1) and can be replaced by  $Ay(\bar{t}_k) + f(\bar{t}_k)$ .

To derive first-order optimality conditions for (P) we introduce the Lagrangian function  $\mathcal{L} : \mathcal{X} \times \mathcal{Y} \times \mathbb{R}^{2\bar{k}} \rightarrow \mathbb{R}$  by

$$\begin{aligned} \mathcal{L}(x, z, \nu) &= J(x) + \langle e(x), z \rangle_{\mathcal{Y}} + \langle g(x), \nu \rangle_{\mathbb{R}^{2\bar{k}}} \\ &= J(\eta, \bar{t}, \psi, \lambda) + \int_0^T (\dot{\eta}(t) - A^\ell \eta(t) - f(t))^T \mathbf{p}(t) dt + (\eta(0) - \eta_0)^T \mathbf{p}_0 \\ &\quad + \sum_{i=1}^{\ell} \langle (\mathcal{R}(\bar{t}) - \lambda_i) \psi_i, \mu_i \rangle_H + \sum_{i=1}^{\ell} \eta_i (1 - \|\psi_i\|_H^2) \\ &\quad + \sum_{k=1}^{\bar{k}} (\bar{t}_k - T) \nu_k^b - \sum_{k=1}^{\bar{k}} \bar{t}_k \nu_k^a, \end{aligned}$$

where  $x \in \mathcal{X}$ ,  $z \in \mathcal{Y}$  and  $\nu = (\nu^a, \nu^b) \in \mathbb{R}^{\bar{k}} \times \mathbb{R}^{\bar{k}}$  Lagrange multipliers corresponding to the inequality constraint.

Setting the first derivatives of  $\mathcal{L}$  with respect to  $x, z$  and  $\nu$  equal to zero gives first-order optimality conditions, provided that certain regularity conditions are satisfied by  $e$  and  $g$ . For this purpose let  $x^*$  denote a local solution to (P), and denote by

$$\underline{\mathcal{A}} = \{i \in \{1, \dots, \bar{k}\} \mid \bar{t}_i^* = 0\} \quad \text{and} \quad \bar{\mathcal{A}} = \{i \in \{1, \dots, \bar{k}\} \mid \bar{t}_i^* = T\}$$

the active sets. Note that  $\underline{\mathcal{A}} \cap \bar{\mathcal{A}} = \emptyset$ . Then,  $x^*$  is also a local solution of (P), if  $\mathcal{F}(\mathbf{P})$  is replaced by

$$\tilde{\mathcal{F}}(\mathbf{P}) = \{x \in \mathcal{X} \mid e(x) = 0, (g_1(x))_i = 0 \text{ for } i \in \underline{\mathcal{A}}, (g_2(x))_i = 0 \text{ for } i \in \bar{\mathcal{A}}\}.$$

For the Lagrangian to provide first-order necessary conditions it is then sufficient that the linearisation of  $(e, (g_1)_{\underline{\mathcal{A}}}, (g_2)_{\bar{\mathcal{A}}})$  at  $x^*$  is surjective, see, e.g., [16]. This is addressed in the following result that is proved in the Appendix.

**Proposition 2.4.** *Let  $x^*$  denote a local solution to (P). Then*

$$(\nabla e(x^*), \nabla g_{1, \underline{\mathcal{A}}}(x^*), \nabla g_{2, \bar{\mathcal{A}}}(x^*)) : \mathcal{X} \rightarrow \mathcal{Y} \times \mathbb{R}^{\sharp(\underline{\mathcal{A}})} \times \mathbb{R}^{\sharp(\bar{\mathcal{A}})}$$

is surjective. Here,  $\sharp(\underline{\mathcal{A}})$  denotes the cardinality of  $\underline{\mathcal{A}}$ . Moreover  $\nabla_{(\eta, \psi, \lambda)} e(x) \delta = r$  has a unique solution  $\delta \in H^1(0, T; \mathbb{R}^\ell) \times H^\ell \times \mathbb{R}^\ell$  for any  $r \in \mathcal{Y}$ .

Proposition 2.4 implies the existence of Lagrange multipliers (or dual variables)  $z^* = (\mathbf{p}^*, \mathbf{p}_0^*, \mu^*, \eta^*) \in \mathcal{Y}$ , and  $\nu^* = (\nu^{a,*}, \nu^{b,*}) \in \mathbb{R}^{2\bar{k}}$  such that

$$\frac{\partial \mathcal{L}}{\partial x}(x^*, z^*, \nu^*) = 0 \tag{2.9}$$

and the following complementarity holds for the inequality constraints:

$$\nu_k^{a,*} \bar{t}_k = 0, \quad \nu_k^{b,*} (T - \bar{t}_k) = 0, \quad \nu^{a,*} \geq 0, \quad \nu^{b,*} \geq 0, \quad 0 \leq \bar{t} \leq T. \tag{2.10}$$

We next explore (2.9). Partial differentiation with respect to  $\eta$  in direction of  $\partial \eta$  implies

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \eta}(x^*, z^*, \nu^*) \delta \eta &= \int_0^T \sum_{j=1}^{\ell} \delta \eta_j(t) \left( \left\langle \sum_{i=1}^{\ell} \eta_i(t) \psi_i^* - y(t), \psi_j^* \right\rangle_H + (-\dot{\mathbf{p}}_j^*(t)) - (A^{\ell, T} \mathbf{p}^*(t))_j \right) dt \\ &\quad + \mathbf{p}^*(T) \delta \eta(T) + (\mathbf{p}_0^* - \mathbf{p}^*(0))^T \delta \eta(0) = 0. \end{aligned}$$

Hence, from (2.6) and (2.9) we derive the adjoint system

$$-\dot{\mathbf{p}}^*(t) = (A^{\ell, \text{T}})\mathbf{p}^*(t) - \begin{pmatrix} \langle y^{\ell, *}(t) - y(t), \psi_1^* \rangle_H \\ \vdots \\ \langle y^{\ell, *}(t) - y(t), \psi_\ell^* \rangle_H \end{pmatrix}, \quad t \in (0, T], \quad (2.11a)$$

$$\mathbf{p}^*(T) = 0 \quad (2.11b)$$

as well as the relation

$$\mathbf{p}_\circ^* = \mathbf{p}^*(0). \quad (2.12)$$

By Gronwall's inequality there exists a constant  $\tilde{C} > 0$  depending on  $\|A\|_{\mathcal{L}(\text{dom } A, H)}$  and  $\|\psi_i^*\|_{\text{dom } A}$  such that

$$\|\mathbf{p}^*\|_{L^2(0, T; \mathbb{R}^\ell)} \leq \tilde{C} \|y^{\ell, *} - y\|_{L^2(0, T; H)}. \quad (2.13)$$

Utilizing (2.1), Remark 2.3 and (H1) we find for  $k = 1, \dots, \bar{k}$

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial t_k}(x^*, z^*, \nu^*) &= \sum_{i=1}^{\ell} \langle Ay(\bar{t}_k^*) + f(\bar{t}_k^*), \psi_i^* \rangle_H \langle y(\bar{t}_k^*), \mu_i^* \rangle_H + \nu_k^{a, *} - \nu_k^{b, *} \\ &+ \sum_{i=1}^{\ell} \langle Ay(\bar{t}_k^*) + f(\bar{t}_k^*), \mu_i^* \rangle_H \langle y(\bar{t}_k^*), \psi_i^* \rangle_H = 0. \end{aligned} \quad (2.14)$$

We find for any direction  $\delta\psi \in H^\ell$

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \psi}(x^*, z^*, \nu^*)\delta\psi &= \sum_{i=1}^{\ell} \langle (\mathcal{R}(\bar{t}^*) - \lambda_i^*)\mu_i^*, \delta\psi_i \rangle_H - \int_0^T (\delta A^\ell \eta^*(t) + \delta f(t))^{\text{T}} \mathbf{p}^*(t) dt \\ &+ \sum_{i=1}^{\ell} \left\langle \int_0^T \eta_i^*(t)(y^{\ell, *}(t) - y(t)) dt - (\mathbf{p}_\circ^*)_i y_\circ - 2\eta_i^* \psi_i^*, \delta\psi_i \right\rangle_H = 0 \end{aligned}$$

where  $\delta A^\ell \in \mathbb{R}^{\ell \times \ell}$  and  $\delta f(t) \in \mathbb{R}^\ell$  has been introduced in (2.8). Since  $A\psi_i \in H$  for  $i = 1, \dots, \ell$  as a consequence of (H1) we find

$$(\mathcal{R}(\bar{t}^*) - \lambda_i^*)\mu_i^* = \mathcal{G}_i(\eta^*, \psi^*, \mathbf{p}^*) + 2\eta_i^* \psi_i^*, \quad (2.15)$$

where

$$\begin{aligned} \mathcal{G}_i(\eta^*, \psi^*, \mathbf{p}^*) &= \int_0^T \eta_i^*(t) \left( A^* \left( \sum_{j=1}^{\ell} \mathbf{p}_j^*(t) \psi_j^* \right) + (y(t) - y_*^\ell(t)) \right) dt \\ &+ \int_0^T \mathbf{p}_i^*(t) \left( A \left( \sum_{j=1}^{\ell} \eta_j^*(t) \psi_j^* \right) + f(t) \right) dt + (\mathbf{p}_\circ^*)_i y_\circ \\ &= \int_0^T \eta_i^*(t) (A^* p^{\ell, *}(t) + y(t) - y_*^\ell(t)) dt + \int_0^T \mathbf{p}_i^*(t) (Ay^{\ell, *}(t) + f(t)) dt + (\mathbf{p}_\circ^*)_i y_\circ \end{aligned}$$

and  $p^\ell(t) = \sum_{i=1}^{\ell} \mathbf{p}_i(t) \psi_i$ .

For any  $k = 1, \dots, \ell$  we have

$$\frac{\partial \mathcal{L}}{\partial \lambda_k}(x^*, z^*, \nu^*) = -\langle \psi_k^*, \mu_k^* \rangle_H = 0 \quad (2.16)$$



and hence by (2.15)

$$\eta_i^* = -\frac{1}{2} \langle \mathcal{G}_i(\mathfrak{h}^*, \psi^*, \mathfrak{p}^*), \psi_i^* \rangle_H. \quad (2.17)$$

Inserting (2.17) into (2.15) yields for  $i = 1, \dots, \ell$

$$(\mathcal{R}(\bar{t}^*) - \lambda_i^*)\mu_i^* = \mathcal{G}_i(\mathfrak{h}^*, \psi^*, \mathfrak{p}^*) - \langle \mathcal{G}_i(\mathfrak{h}^*, \psi^*, \mathfrak{p}^*), \psi_i^* \rangle_H \psi_i^* \in \text{span}\{\psi_i^*\}^\perp \quad (2.18)$$

implying that (2.15) admits a unique solution. We summarize these computations in the following theorem.

**Theorem 2.5.** *With (H1) holding first-order necessary optimality conditions for (P) are given by (2.11), (2.14), (2.16) and (2.18).*

Defining the reduced cost functional

$$\hat{J}(\bar{t}) = J(\mathfrak{h}(\bar{t}), \bar{t}, \psi(\bar{t}), \lambda(\bar{t})) \quad (2.19)$$

we consider the reduced problem

$$\min \hat{J}(\bar{t}) \quad \text{s.t.} \quad \bar{t} = (\bar{t}_1, \dots, \bar{t}_{\bar{k}}) \in \mathbb{R}^{\bar{k}} \text{ with } 0 \leq \bar{t}_i \leq T \text{ for } i = 1, \dots, \bar{k}, \quad (\hat{\mathbf{P}})$$

which is equivalent with (P). Using (2.14) the gradient of  $\hat{J}$  is given by

$$\nabla \hat{J}(\bar{t}) = (v_1, \dots, v_{\bar{k}}) \in \mathbb{R}^{\bar{k}}, \quad (2.20)$$

where the  $v_k$ 's are given as

$$v_k = \sum_{i=1}^{\ell} \langle Ay(\bar{t}_k^*) + f(\bar{t}_k^*), \psi_i^* \rangle_H \langle y(\bar{t}_k^*), \mu_i^* \rangle_H + \sum_{i=1}^{\ell} \langle Ay(\bar{t}_k^*) + f(\bar{t}_k^*), \mu_i^* \rangle_H \langle y(\bar{t}_k^*), \psi_i^* \rangle_H \quad \text{for } k = 1, \dots, \bar{k}.$$

**Example 2.6.** This elementary example illustrates some features of the proposed methodology. In particular, the sensitivity of the spectral data with respect to the additional snapshot location  $\bar{t}$  will be shown. Moreover, the solutions to (2.1), (2.7) and the gradient of the reduced cost functional (2.20) can be computed exactly. Let us consider the dynamical system in  $\mathbb{R}^2$

$$\dot{y}(t) = f(t) \text{ for } t \in (0, T], \text{ and } y(0) = y_o = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad (2.21)$$

where

$$f(t) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ for } t \in [0, T/4) \text{ and } f(t) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ for } t \in [T/4, T].$$

The exact solution to (2.21) is given by

$$y(t) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ for } t \in [0, T/4) \text{ and } y(t) = \begin{pmatrix} t - T/4 \\ 1 \end{pmatrix} \text{ for } t \in [T/4, T].$$

Let  $t_0 = 0$  be the fixed time instance and  $\bar{k} = 1$ , *i.e.*, we are looking for one additional snapshot location  $\bar{t} = \bar{t}_1$ . The behavior of the two eigenvalues of  $\mathcal{R}$  are shown in Figure 1. Let  $\ell = 1$ . A short computation shows that the one-dimensional POD approximation in  $\text{span}\{\psi_1\}$  is given by

$$y^\ell(t) = \begin{cases} (\psi_1)_2 \psi_1 & \text{for } t \in [0, T/4) \\ ((t - T/4)(\psi_1)_1 + (\psi_1)_2) \psi_1 & \text{for } t \in [T/4, T], \end{cases}$$

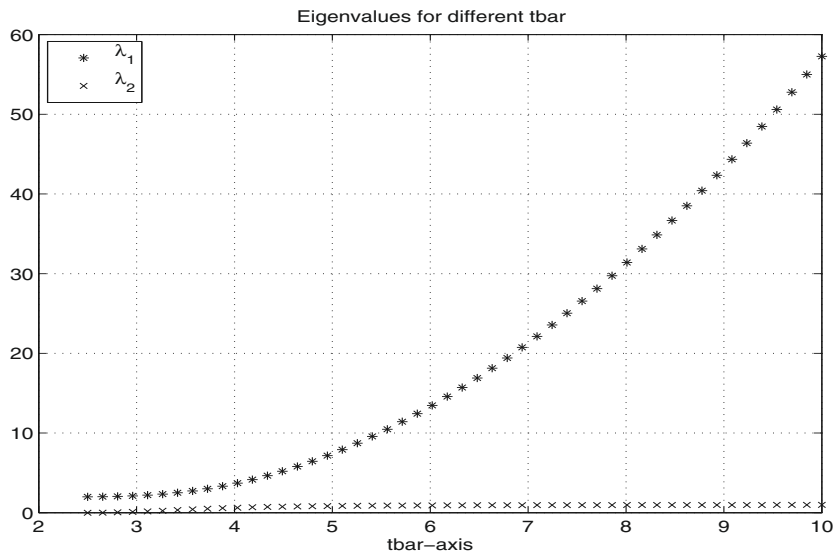


FIGURE 1. Example 2.6: eigenvalues for  $\mathcal{R}(\bar{t})$  for  $\bar{t} \in [T/4, T]$  and  $T = 4$ .

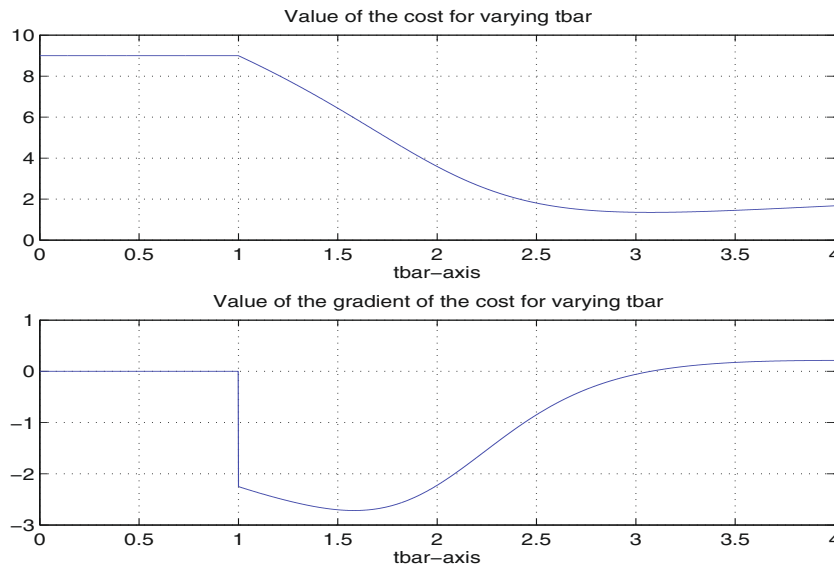


FIGURE 2. Example 2.6: values of the cost function and its gradient for  $\bar{t} \in [T/4, T]$  and  $T = 4$ .

where  $(\psi_1)_1$  and  $(\psi_1)_2$  denote the first and second component of  $\psi_1 \in \mathbb{R}^2$ , respectively. The value of the cost functional as a function of  $\bar{t}$  and its derivative are given in Figure 2. We note that  $J(\bar{t})$  has a unique global minimum in  $[0, T]$ . Moreover in the time-interval  $[0, T/4]$  of constant dynamics, the basis function captures the trajectory and hence the cost functional is insensitive to local changes of the additional snapshot location  $\bar{t}$ . As soon as  $\bar{t}$  is large enough (*i.e.*,  $\bar{t} > T/4$ ) information is added and the POD approximation is improved.

## 3. SECOND-ORDER OPTIMALITY CONDITIONS

We start with the tedious characterization of the second order partial derivatives of  $\mathcal{L}$  at  $(x^*, z^*, \nu^*)$ .

- The second derivative  $\nabla_{\eta\eta}^2 \mathcal{L}(x^*, z^*, \nu^*) : H^1(0, T; \mathbb{R}^\ell) \times H^1(0, T; \mathbb{R}^\ell) \rightarrow \mathbb{R}$  of  $\mathcal{L}$  with respect to  $\eta$  is given as

$$\nabla_{\eta\eta}^2 \mathcal{L}(x^*, z^*, \nu^*)(\partial\eta, \widetilde{\partial\eta}) = \int_0^T \partial\eta(t)^T \widetilde{\partial\eta}(t) dt \quad \text{for } \partial\eta, \widetilde{\partial\eta} \in H^1(0, T; \mathbb{R}^\ell).$$

- Note that  $\nabla_{\bar{t}\partial\eta}^2 \mathcal{L}(x^*, z^*, \nu^*) \equiv 0$  holds.
- We find

$$\nabla_{\eta\psi}^2 \mathcal{L}(x^*, z^*, \nu^*)(\delta\eta, \delta\psi) = - \int_0^T (\delta A^\ell \delta\eta)^T \mathbf{p}^*(t) dt + \sum_{i=1}^{\ell} \left\langle \int_0^T \delta\eta_i(t) (y^{\ell,*}(t) - y(t)) dt, \delta\psi_i \right\rangle_H$$

and hence the operator representation is given by

$$\left( \nabla_{\eta\psi}^2 \mathcal{L}(x^*, z^*, \nu^*) \right)_{i,j} = (y^{\ell,*} - y) \delta_{i,j} + \left( \sum_k p_k A^* \psi_k^* \right) \delta_{i,j} + p_j(t) A \psi_j^*.$$

Henceforth  $C_1 = C_1(\|A\|_{\mathcal{L}(\text{dom } A, H)}, \|\psi_i^*\|_{\text{dom } A})$  denotes a generic constant. We have

$$\|\nabla_{\eta\psi}^2 \mathcal{L}(x^*, z^*, \nu^*)\| \leq C_1 (\|y^{\ell,*} - y\|_{L^2(0, T; H)} + \|\mathbf{p}^*\|_{L^2(0, T; \mathbb{R}^\ell)}),$$

and by (2.13)

$$\|\nabla_{\eta\psi}^2 \mathcal{L}(x^*, z^*, \nu^*)\| \leq C_1 \|y^{\ell,*} - y\|_{L^2(0, T; H)}. \quad (3.1)$$

- Notice that  $\nabla_{\lambda\eta}^2 \mathcal{L}(x^*, z^*, \nu^*) = 0$  holds.
- We assume in addition that

$$y \in C^2([0, T], H) \cap C^1([0, T], \text{dom } A), \quad f \in C^1([0, T], H). \quad (\text{H2})$$

Then  $\nabla_{\bar{t}\bar{t}}^2 \mathcal{L}(x^*, z^*, \nu^*)$  is a diagonal matrix with its elements given by

$$\begin{aligned} \left( \nabla_{\bar{t}\bar{t}}^2 \mathcal{L}(x^*, z^*, \nu^*) \right)_{k,k} &= \sum_{i=1}^{\ell} \left\langle Ay(\bar{t}_k^*) + \dot{f}(\bar{t}_k^*), \psi_i^* \right\rangle_H \left\langle y(\bar{t}_k^*), \mu_i^* \right\rangle_H + \sum_{i=1}^{\ell} \left\langle Ay(\bar{t}_k^*) + \dot{f}(\bar{t}_k^*), \mu_i^* \right\rangle_H \left\langle y(\bar{t}_k^*), \psi_i^* \right\rangle_H \\ &\quad + 2 \sum_{i=1}^{\ell} \left\langle Ay(\bar{t}_k^*) + \dot{f}(\bar{t}_k^*), \psi_i^* \right\rangle_H \left\langle Ay(\bar{t}_k^*) + \dot{f}(\bar{t}_k^*), \mu_i^* \right\rangle_H. \end{aligned}$$

- We find that

$$\left( \nabla_{\psi\bar{t}}^2 \mathcal{L}(x^*, z^*, \nu^*) \delta\psi \right)_k = \sum_{i=1}^{\ell} \left\langle \left\langle y(\bar{t}_k^*), \mu_j^* \right\rangle_H (Ay(\bar{t}_k^*) + \dot{f}(\bar{t}_k^*)) + \left\langle Ay(\bar{t}_k^*) + \dot{f}(\bar{t}_k^*), \mu_j^* \right\rangle_H y(\bar{t}_k^*), \delta\psi_i \right\rangle_H \quad (3.2)$$

for all  $\delta\psi = (\delta\psi_1, \dots, \delta\psi_\ell) \in H^\ell$  and for  $k = 1, \dots, \bar{k}$ . From the definition of  $\mathcal{G}_i$ , we have

$$\sup_{1 \leq i \leq \ell} \|\mathcal{G}_i\|_H \leq C_1 \|y^{\ell,*} - y\|_{L^2(0, T; H)},$$

and by (2.18) there exists a constant  $C_2 = C_2(C_1, (\mathcal{R}(\bar{t}^*) - \lambda_i^*)^{-1}|_{\text{span}\{\psi_i\}^\perp})$  such that

$$\sup_{1 \leq i \leq \ell} \|\mu_i\|_H \leq C_2 \|y^{\ell,*} - y\|_{L^2(0, T; H)}. \quad (3.3)$$

Thus, there exists  $C_3 = C_3(C_2, \|Ay + f\|_{C(0,T,H)})$  such that

$$\|\nabla_{\psi\bar{t}}^2 \mathcal{L}(x^*, z^*, \nu^*)\| \leq C_3 \|y^{\ell,*} - y\|_{L^2(0,T;H)}. \quad (3.4)$$

- It follows that  $\nabla_{\lambda\bar{t}}^2 \mathcal{L}(x^*, z^*, \nu^*) = 0$  holds.
- Moreover,  $\nabla_{\psi\psi}^2 \mathcal{L}(x^*, z^*, \nu^*) \in \mathcal{L}(H^\ell)$  and for  $\partial\psi \in H^\ell$  we find

$$\begin{aligned} \nabla_{\psi\psi}^2 \mathcal{L}(x^*, z^*, \nu^*)(\delta\psi, \widetilde{\delta\psi}) &= \int_0^T \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} \eta_i^*(t) \eta_j^*(t) \langle \widetilde{\delta\psi}_i, \delta\psi_j \rangle_H dt \\ &\quad - \int_0^T \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} \eta_i^*(t) \mathfrak{p}_j^*(t) \left( \langle A\delta\psi_j, \widetilde{\delta\psi}_i \rangle_H + \langle A\widetilde{\delta\psi}_j, \delta\psi_i \rangle_H \right) dt - 2 \sum_{i=1}^{\ell} \eta_i^* \langle \widetilde{\delta\psi}_i, \delta\psi_i \rangle_H, \end{aligned}$$

which implies the operator representation

$$\left( \nabla_{\psi\psi}^2 \mathcal{L}(x^*, z^*, \nu^*) \right)_{i,j} = \int_0^T \eta_i^*(t) \eta_j^*(t) dt - 2\eta_i^* \delta_{ij} - \left( \int_0^T \eta_i^*(t) \mathfrak{p}_j^*(t) dt \right) A^* - \left( \int_0^T \sum_{k=1}^{\ell} \eta_k^*(t) \mathfrak{p}_k^*(t) dt \right) A$$

for  $1 \leq i, j \leq \ell$ . By (2.17) we have

$$\sup_{1 \leq i \leq \ell} |\eta_i^*| \leq C_1 \|y^{\ell,*} - y\|_{L^2(0,T;H)}. \quad (3.5)$$

Therefore  $\nabla_{\psi\psi}^2 \mathcal{L}(x^*, z^*, \nu^*)$  consists of a non-negative term and terms which behave like  $\|y^{\ell,*} - y\|_{L^2(0,T;H)}$ .

- The operator  $\nabla_{\lambda\psi}^2 \mathcal{L}(x^*, z^*, \nu^*) \in \mathcal{L}(H^\ell, \mathbb{R}^\ell)$  is a diagonal operator, with elements given by

$$\left( \nabla_{\lambda\psi}^2 \mathcal{L}(x^*, z^*, \nu^*) \right)_{i,i} = -\mu_i^*.$$

By (2.17) these elements behave like  $\|y^{\ell,*} - y\|_{L^2(0,T;H)}$ .

- Finally,  $\frac{\partial^2 \mathcal{L}}{\partial \lambda^2}(x^*, z^*, \nu^*) = 0$ .

The structure of zero entries into  $\nabla_*^2 \mathcal{L}(x^*, z^*, \nu^*)$  is depicted in the following matrix, where the variables are ordered as  $(\mathfrak{h}, \psi, \lambda, \bar{t})$ :

$$\nabla_{xx}^2 \mathcal{L}(x^*, z^*, \nu^*) = \left( \begin{array}{c|c|c|c} I & * & 0 & 0 \\ * & * & * & * \\ \hline 0 & * & 0 & 0 \\ 0 & * & 0 & * \end{array} \right).$$

Assuming that  $y^{\ell,*} = y$  its structure is given by

$$\nabla_{xx}^2 \mathcal{L}(x^*, z^*, \nu^*) \sim \left( \begin{array}{c|c|c|c} I & 0 & 0 & 0 \\ \hline 0 & M^\ell & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right), \quad (3.6)$$

where the nonnegative matrix  $M^\ell \in \mathbb{R}^{\ell \times \ell}$  is given by

$$M_{ij}^\ell = \int_0^T \eta_i^*(t) \eta_j^*(t) dt.$$

These quantities allow to specify the Hessian  $\nabla^2 \hat{J}(\bar{t})$  of the reduced function with respect to  $\bar{t}$ . We set

$$\hat{x} = (\eta, \psi, \lambda)$$

and obtain:

$$\nabla^2 \hat{J}(\bar{t}) = \nabla_{\hat{x}\bar{t}}^2 \mathcal{L}(\hat{x}, \bar{t}, z, \nu) w_1 + \nabla_{\bar{t}} e(\hat{x}, \bar{t})^* w_2 + \nabla_{\bar{t}\bar{t}}^2 \mathcal{L}(\hat{x}, \bar{t}, z, \nu), \quad (3.7)$$

where  $(\hat{x}, z, \nu) = (\hat{x}(\bar{t}), z(\bar{t}), \nu(\bar{t}))$ ,  $e = (e_1, e_2, e_3, e_4)$  and

$$(w_1, w_2) \in \mathcal{L}(\mathbb{R}^{\bar{k}}, H^1(0, T; \mathbb{R}^\ell) \times H^\ell \times \mathbb{R}^\ell) \times \mathcal{L}(\mathbb{R}^{\bar{k}}, L^2(0, T; \mathbb{R}^\ell) \times \mathbb{R}^\ell \times H^\ell \times \mathbb{R}^\ell)$$

is the solution to

$$\begin{pmatrix} \nabla_{xx}^2 \mathcal{L}(\hat{x}, \bar{t}, z, \nu) & \nabla_{\hat{x}} e(\hat{x}, \bar{t})^* \\ \nabla_{\hat{x}} e(\hat{x}, \bar{t}) & 0 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} + \begin{pmatrix} \nabla_{\hat{x}\bar{t}}^2 \mathcal{L}(\hat{x}, \bar{t}, z, \nu) \\ \nabla_{\bar{t}} e(\hat{x}, \bar{t}) \end{pmatrix} = 0. \quad (3.8)$$

This expression for the Hessian can be derived analogously as in pde-constrained optimization, see [8,9], for example. Note that solving (3.8) requires to solve the linearized equation

$$\nabla_{\hat{x}} e(\hat{x}, \bar{t}) w_1 = -\nabla_{\bar{t}} e(\hat{x}, \bar{t}), \quad (3.9)$$

and the adjoint equation

$$(\nabla_{\hat{x}} e(\hat{x}, \bar{t}))^* w_2 = -\nabla_{\hat{x}\hat{x}}^2 \mathcal{L}(\hat{x}, \bar{t}, z, \nu) w_1 - \nabla_{\hat{x}\bar{t}}^2 \mathcal{L}(\hat{x}, \bar{t}, z, \nu). \quad (3.10)$$

Unique solvability of (3.9) follows with the same arguments as used in the proof of Proposition 2.4 given in the Appendix. The  $\bar{t}$  component is left out and  $\delta \bar{t}$  is set equal to 0. Moreover, since  $\nabla_{\bar{t}} e_i(x) = 0$  for  $i = 1, 2$ , the first and second components of  $w_1$  equal 0, and only (A.2) needs to be solved with  $(\tilde{c}_3^i, c_4^i) = -(\nabla_{\bar{t}} e_3(\hat{x}, \bar{t}), 0)$ . Unique solvability of (3.10) can again be argued in a manner analogous to that in the Appendix. Indeed, the only non selfadjoint term is  $A^\ell$ .

We end this section by giving a second order sufficient condition for local optimality of  $x^*$ . For later use we introduce for  $i = 1, \dots, \ell$

$$(c_3)_i = \sum_{k=1}^{\bar{k}} \langle \dot{y}(t_k^*), \psi_i \rangle y(t_k^*) + \langle y(\bar{t}^*), \psi_i \rangle \dot{y}(t_k^*). \quad (3.11)$$

The Hessian  $\nabla^2 \hat{J}(\bar{t}^*)$  can also be expressed as

$$\nabla^2 \hat{J}(\bar{t}^*) = \mathcal{T}(x^*)^* \frac{\partial^2}{\partial x^2} \mathcal{L}(x^*, z^*, \lambda^*) \mathcal{T}(x^*),$$

where

$$\mathcal{T}(x) \in \mathcal{L}(\mathbb{R}^{\bar{k}}, H^1(0, T; \mathbb{R}^\ell) \times \mathbb{R}^{\bar{k}} \times H^\ell \times \mathbb{R}^\ell)$$

is the null-space representation of  $\nabla e(x^*)$  given by

$$\mathcal{T}(x) = \begin{pmatrix} -(\frac{\partial}{\partial \hat{x}} e(x))^{-1} \frac{\partial}{\partial \bar{t}} e(x) \\ I \end{pmatrix}, \quad (3.12)$$

see, e.g., [8,9]. In view of the computations in the Appendix we have

$$(\nabla_{\hat{x}} e(x))^{-1} \nabla_{\bar{t}} e(x) = (0, \delta \psi, \delta \lambda),$$

where  $(\delta\psi, \delta\lambda)$  is the solution to

$$\begin{pmatrix} (\mathcal{R}(\bar{t}^*) - \lambda_1) \delta\psi_1 - \delta\lambda_1 \psi_1 \\ \vdots \\ (\mathcal{R}(\bar{t}^*) - \lambda_\ell) \delta\psi_\ell - \delta\lambda_\ell \psi_\ell \end{pmatrix} = \begin{pmatrix} c_{3,1} \\ \vdots \\ c_{3,\ell} \end{pmatrix}, \quad \begin{pmatrix} \langle \psi_1, \delta\psi_1 \rangle_H \\ \vdots \\ \langle \psi_\ell, \delta\psi_\ell \rangle_H \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

If  $\text{col}\{c_{3,i} - \langle c_{3,i}, \psi_i \rangle \psi_i\}_{i=1}^\ell$  is not the zero-vector, then  $\text{col}\{\delta\psi\}_{i=1}^\ell$  is nontrivial. In view of (3.6) we have for  $y^{\ell,*} = y$  that

$$\mathcal{T}(x^*)^* \nabla_{xx}^2 \mathcal{L}(x^*, z^*, \lambda^*) \mathcal{T}(x^*) = \delta\psi^T M^\ell \delta\psi \geq 0, \quad (3.13)$$

and the finite dimensional operator on the left hand side of (3.13) is positive definite if  $M^\ell$  is positive definite. A perturbation argument implies that this remains valid if  $\|y^{\ell,*} - y\|_{L^2(0,T;\mathbb{R}^\ell)}$  is sufficiently small. We summarize these arguments as a theorem.

**Theorem 3.1.** *If  $M^\ell$  is positive definite,  $\text{col}\{c_{3,i} - \langle c_{3,i}, \psi_i \rangle \psi_i\}_{i=1}^\ell$  is not the zero-vector, and  $\|y^{\ell,*} - y\|_{L^2(0,T;\mathbb{R}^\ell)}$  is sufficiently small, then  $\nabla^2 \hat{J}(\bar{t}^*)$  is positive definite.*

Under the assumptions of Theorem 3.1 therefore,  $x^*$  is a strict local minimum.

#### 4. NUMERICAL EXAMPLES

This section is devoted to illustrate the presented theoretical results for some applications. The existence of first- and second-order derivatives justifies the use of a second-order numerical method for solving  $(\hat{\mathbf{P}})$ . Since the second derivative is rather involved we utilize a quasi-Newton method, which for the readers convenience is repeated in Algorithm 1. For a convergence analysis we refer, *e.g.*, to [16].

---

##### Algorithm 1 (Quasi-Newton method)

---

- 1: Choose a starting value  $\bar{t}^0 \in [0, T]$  and an initial quasi-Newton matrix  $H^0 \in \mathbb{R}^{\bar{k} \times \bar{k}}$ , fix the number of POD basis functions  $\ell$ , set  $i = 0$ .
- 2: **repeat**
- 3:   Compute the operator  $\mathcal{R}(\bar{t}^i)$  and solve the eigenvalue problem (2.2).
- 4:   Solve for  $\eta^i$  the state system (2.7).
- 5:   Compute the dual  $\mathbf{p}^i$  from (2.11).
- 6:   Set  $\eta^i$  by (2.17) and solve (2.18) for  $\mu^i$ .
- 7:   Determine the reduced gradient  $\nabla \hat{J}(\bar{t}^i)$  from (2.20) and solve the quasi-Newton system

$$H^i d^i = -\nabla \hat{J}(\bar{t}^i)$$

for  $d^i \in \mathbb{R}^{\bar{k}}$ .

- 8:   Apply a linesearch strategy to obtain a next iterate  $\bar{t}^{i+1} = \bar{t}^i + s_i \Delta \bar{t}^i \in [0, T]$ .
  - 9:   Compute a positive definite quasi-Newton matrix  $H^{i+1} \in \mathbb{R}^{\bar{k} \times \bar{k}}$  using the BFGS formula and set  $i = i + 1$ .
  - 10: **until** some stopping criterium is satisfied.
- 

Next we turn to a numerical examples, where the dynamical systems (2.1) are given by a linear, parabolic problem. The codes are written in MATLAB utilizing routines from the FEMLAB package for the FE discretization of the spatial domain. Algorithm 1 is realized by calling the routine `fmincon` from the MATLAB OPTIMIZATION TOOLBOX. The options are chosen as follows

```
optimset('Diagnostics','on','Display','iter','LargeScale','off',...
        'GradObj','on','TolFun',1e-6,'TolX',1e-4,'MaxIter',100);
```

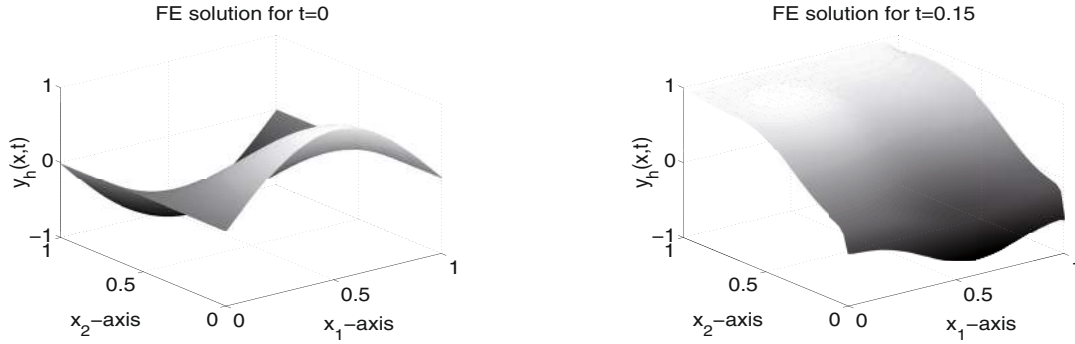


FIGURE 3. Run 1: initial condition  $y_0$  (left plot) and FE solution  $y_h$  for  $t = 0.15$  (right plot).

For given time vector  $\bar{t}^k \in \mathbb{R}^{\bar{k}}$  our user supplied function computes both the reduced cost (2.19) and its gradient (2.20). Our numerical experience indicates that the accurate fulfilment of (2.16) is significant.

We present two numerical examples, focusing on two different aspects:

- Run 1. In this case a relatively large convection in the  $x_2$ -direction implies that the effect of the forcing function and initial condition are transported rather rapidly during the early stages of the evolution. At about  $t = 0.3$  the effect of the boundary condition at  $x_2 = 0$  counteracts this effect and brings the system to a quasi steady state, compare Figures 3 and 4. Due to this dynamical behavior it can therefore be expected that the majority of a “good” snapshot set should be at the beginning of the time interval. To test our approach we initialize our snapshot grid to be uniform within  $[0, T]$ . Where should we put  $\bar{k}$  additional (here  $\bar{k} = 4$ ) time instances to improve the reduced-order solution?
- Run 2. Differently from the previous run we now choose a discontinuous initial condition and reduce the convection coefficient in the  $x_2$ -direction. Starting from a fine uniform grid having fixed snapshots only on the subinterval  $[T/2, T]$ , we determine only one additional snapshot by solving  $(\hat{\mathbf{P}})$ . Since the solution is regularized as time increases, we would expect that a “good” snapshot set should contain information for small  $t$ -values. Does our framework coincide with this expectation and place the additional snapshot in the subinterval  $[0, T/2)$ ?

**Run 1.** In this numerical example we consider the heat equation on the unit square. For  $T = 1$  let  $Q = (0, T) \times \Omega$  and  $\Omega = (0, 1) \times (0, 1) \subset \mathbb{R}^2$ . The boundary  $\partial\Omega$  of  $\Omega$  is denoted by  $\Gamma$  and we set  $\Sigma = (0, T) \times \Gamma$ . For the finite element (FE) triangulation we choose a uniform grid with mesh size  $h = 1/40$ , *i.e.*, we have 900 degrees of freedoms for the spatial discretization. Then, we consider

$$\begin{aligned} y_t(t, \mathbf{x}) - c\Delta y(t, \mathbf{x}) + \beta \cdot \nabla y(t, \mathbf{x}) + y(t, \mathbf{x}) &= f(\mathbf{x}) && \text{for all } (t, \mathbf{x}) \in Q, \\ c \frac{\partial y}{\partial n}(t, \mathbf{x}) + q(\mathbf{x})y(t, \mathbf{x}) &= g(\mathbf{x}) && \text{for all } (t, \mathbf{x}) \in \Sigma, \\ y(0, \mathbf{x}) &= y_0(\mathbf{x}) && \text{for all } \mathbf{x} \in \Omega, \end{aligned}$$

where  $c = 0.1$ ,  $\beta = (0.1, -10)^T \in \mathbb{R}^2$ ,

$$f(\mathbf{x}) = \begin{cases} 4 & \text{for all } \mathbf{x} = (x_1, x_2) \text{ with } (x_1 - 0.25)^2 + (x_2 - 0.65)^2 \leq 0.05, \\ 0 & \text{otherwise,} \end{cases}$$

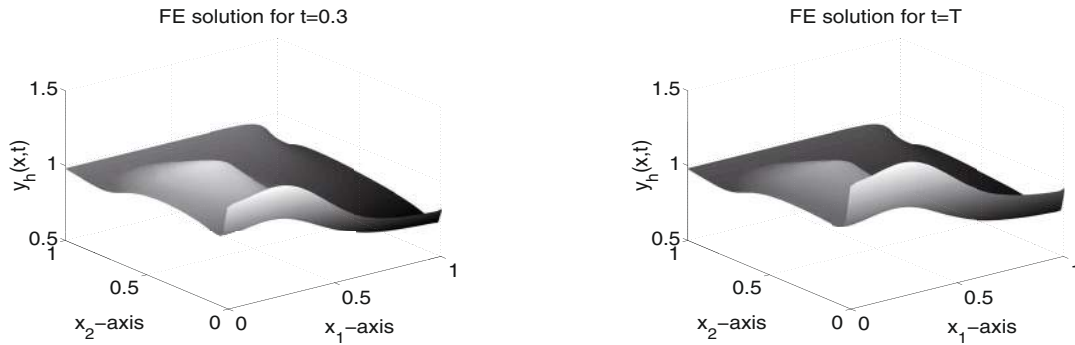


FIGURE 4. Run 1: FE solution  $y_h$  for  $t = 0.3$  (left plot) and for  $t = T$  (right plot).

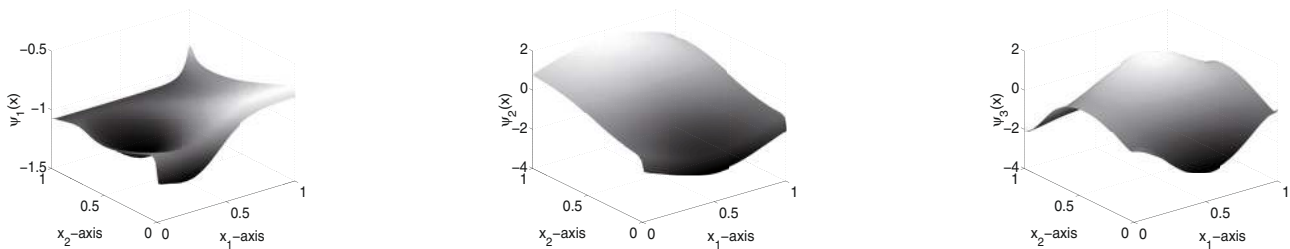


FIGURE 5. Run 1: POD basis functions for the initial time grid.

and  $y_o(\mathbf{x}) = \sin(\pi x_1) \cos(\pi x_2)$  for  $\mathbf{x} = (x_1, x_2) \in \Omega$  (see Fig. 3, left plot). Furthermore, we have

$$q(\mathbf{x}) = \begin{cases} 1 & \text{for } \mathbf{x} = (x_1, 1) \text{ with } 0 < x_1 < 1, \\ x_2 & \text{for } \mathbf{x} = (1, x_2) \text{ with } 0 < x_2 < 1, \\ -2 & \text{for } \mathbf{x} = (x_1, 0) \text{ with } 0 < x_1 < 1, \\ 0 & \text{for } \mathbf{x} = (0, x_2) \text{ with } 0 < x_2 < 1, \end{cases}$$

$$g(\mathbf{x}) = \begin{cases} 1 & \text{for } \mathbf{x} = (x_1, 1) \text{ with } 0 < x_1 < 1, \\ 0 & \text{for } \mathbf{x} = (1, x_2) \text{ with } 0 < x_2 < 1, \text{ for } \mathbf{x} = (0, x_2) \text{ with } 0 < x_2 < 1, \\ -1 & \text{for } \mathbf{x} = (x_1, 0) \text{ with } 0 < x_1 < 1. \end{cases}$$

We utilize piecewise linear FE ansatz functions. The FE solution  $y_h = y_h(t, \mathbf{x})$  for different time instances  $t$  are shown in Figures 3 and 4. We observe that the finite element solution does not change significantly from  $t = 0.3$  to  $t = T$ . Next we take snapshots on the fixed uniform time grid  $t_j = j\Delta t$ ,  $0 \leq j \leq m$ , with  $m = 10$  and  $\Delta t = T/m = 0.1$ . The goal is to determine four additional time instances  $\bar{t} = (\bar{t}_1, \dots, \bar{t}_4) \in [0, T]$  based on  $(\hat{\mathbf{P}})$ . Since the behavior of the solution exhibits more change during the initial time interval  $[0, 0.3]$  than later on, we initialize Algorithm 1 by  $\bar{t}^0 = (0.05, 0.15, 0.25, 0.35) \in [0, T]$ . The number of POD ansatz functions is chosen to be  $\ell = 3$ . The corresponding value of the cost is  $\hat{J}(\bar{t}^0) \approx 0.1093$ . Utilizing the FE snapshots  $\{y_h(t_0, \cdot), \dots, y_h(t_m, \cdot)\} \cup \{y_h(\bar{t}_1^0, \cdot), \dots, y_h(\bar{t}_4^0, \cdot)\}$  we compute the three POD basis functions corresponding to the largest eigenvalues of  $\mathcal{R}$  (see Fig. 5). After 14 iterations and the routine `fmincon` stops with the numerical solution  $\bar{t}^* = (0.0092, 0.0076, 0.1336, 0.2882) \in [0, T]$ . The fixed uniform time grid together with the optimal locations of the new four time instances are plotted in Figure 6. The associated value of the cost function is  $\hat{J}(\bar{t}^*) \approx 0.0165$ . Thus, the reduction of the cost is about 85%. The associated three POD basis functions



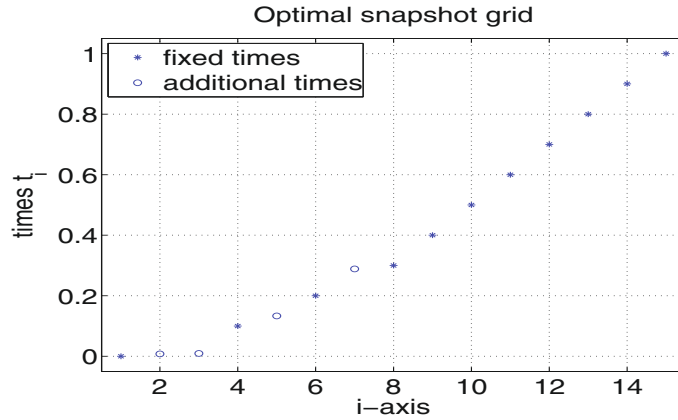


FIGURE 6. Run 1: fixed time grid (\*) and the four additional optimal time instances (o).

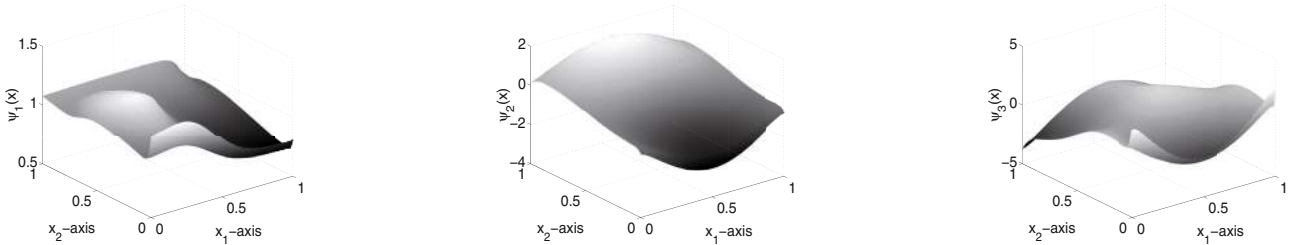


FIGURE 7. Run 1: POD basis functions for the optimized time grid.

TABLE 1. Run 1: optimal value of the cost  $\hat{J}$  for different number  $\ell$  of POD basis functions.

	$\ell = 2$	$\ell = 3$	$\ell = 4$	$\ell = 6$
$\hat{J}(\bar{t}^*)$	0.0676	0.0165	0.0015	0.0013

are plotted in Figure 7. Compared to Figure 7 the shapes of the second and third POD basis functions are significantly different from the ones associated with the starting time vector  $\bar{t}^0$ . In Table 1 we present the value of the cost functional evaluated at the optimal solution  $\bar{t}^*$  to  $(\hat{\mathbf{P}})$  for different values of  $\ell$ .

**Run 2.** In the second numerical example we consider the heat equation on the same domain and with the same FE discretization as in Run 1. The coefficients are  $T = 1$ ,  $c = 0.1$ ,  $\beta = (0.1, -0.1)$ ,  $a = 1$ ,  $f = 0$ ,  $q = 0$ ,  $g = 0$  and

$$y_o(\mathbf{x}) = \begin{cases} 4 & \text{for all } \mathbf{x} = (x_1, x_2) \text{ with } (x_1 - 0.25)^2 + (x_2 - 0.65)^2 \leq 0.05, \\ 0 & \text{otherwise.} \end{cases}$$

The FE solution  $y_h = y_h(t, \mathbf{x})$  for different time instances  $t$  are shown in Figures 8 and 9. In contrast to Run 1 we compute the POD basis using FE snapshots

- only on the subinterval  $[T/2, T]$  using a fine uniform grid, and
- have at our disposal only one additional snapshot determined by  $(\hat{\mathbf{P}})$ .

More precisely, we take snapshots on the fixed uniform time grid  $t_j = T/2 + j\Delta t$ ,  $0 \leq j \leq m$ , with  $m = 150$  and  $\Delta t = T/(2m) \approx 0.0033$ . The goal is to determine one additional time instance  $\bar{t} \in [0, T]$  based on  $(\hat{\mathbf{P}})$ . The number of POD ansatz functions is chosen to be  $\ell = 3$ . Due to the chosen snapshot grid, the initial

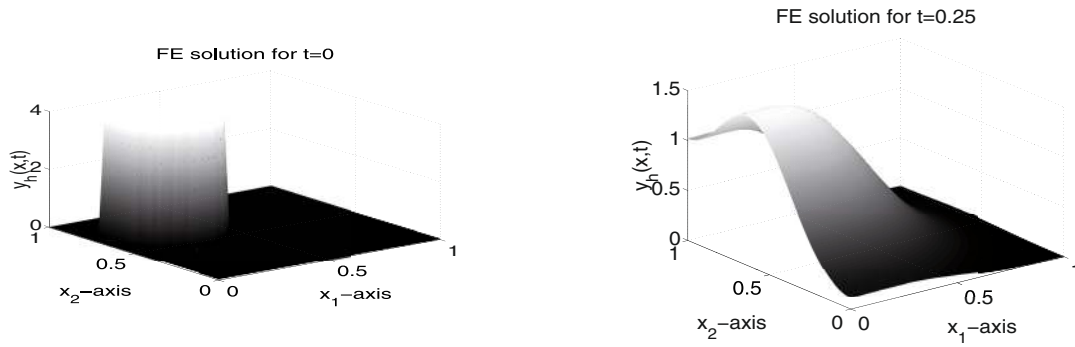


FIGURE 8. Run 2: initial condition  $y_0$  (left plot) and FE solution  $y_h$  for  $t = 0.15$  (right plot).

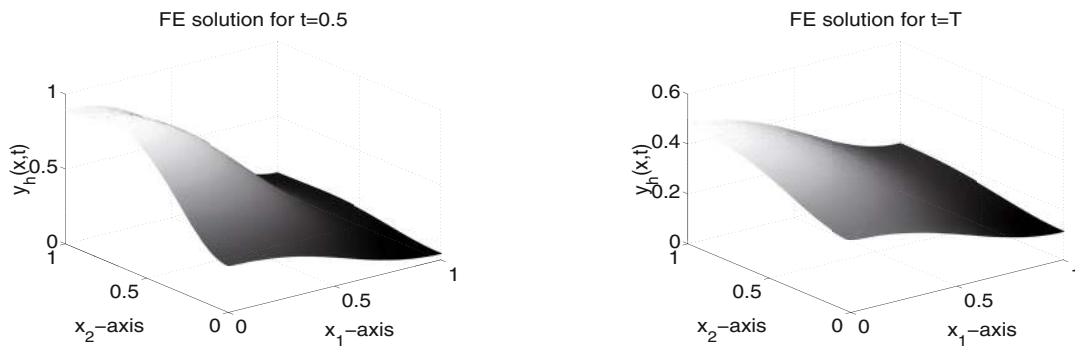


FIGURE 9. Run 2: FE solution  $y_h$  for  $t = 0.3$  (left plot) and for  $t = T$  (right plot).

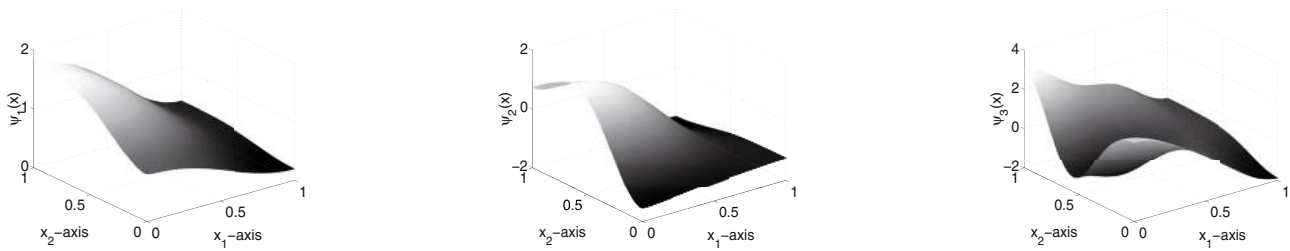


FIGURE 10. Run 2: POD basis functions for the initial time grid.

condition  $y_0$  is not included in the POD modelling. We initialize Algorithm 1 by  $\bar{t}^0 = 0.4 \in [0, T]$  and expect that the optimal  $\bar{t}^*$  is more close to 0. The corresponding value of the cost is  $\hat{J}(\bar{t}^0) \approx 0.0128$ . Utilizing the FE snapshots  $\{y_h(t_0, \cdot), \dots, y_h(t_m, \cdot)\} \cup \{y_h(\bar{t}^0, \cdot)\}$  we compute the three POD basis functions corresponding to the largest eigenvalues of  $\mathcal{R}$  (see Fig. 10). After 8 iterations and the routine `fmincon` stops with the numerical solution  $\bar{t}^* = 0.0358 \in [0, T]$ . The associated value of the cost function is  $\hat{J}(\bar{t}^*) \approx 0.0012$ . Thus, the reduction of the cost is about 90%. The associated three POD basis functions are plotted in Figure 11. Again the second and third basis functions differ significantly from those associated to  $\bar{t}^0$ , compare Figure 10. It turns out

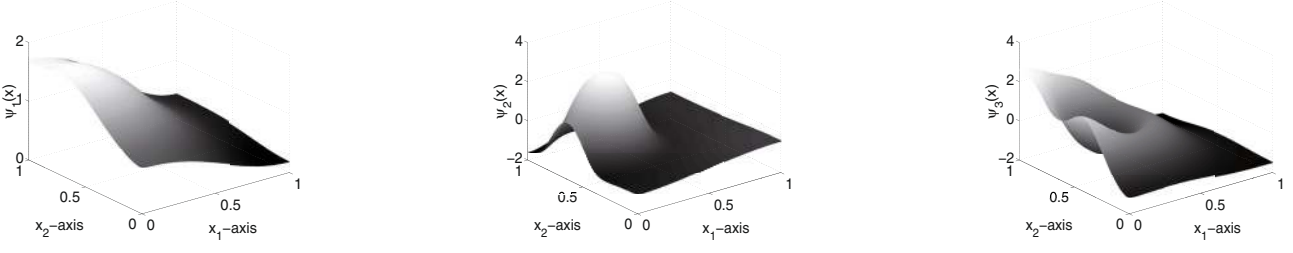


FIGURE 11. Run 2: POD basis functions for the optimized time grid.

TABLE 2. Run 2: optimal time instance  $\bar{t}^*$  and optimal cost value  $\hat{J}$  for different number  $\ell$  of POD basis functions.

	$\ell = 1$	$\ell = 2$	$\ell = 3$	$\ell = 4$	$\ell = 5$	$\ell = 6$
$\bar{t}^*$	0.0000	0.0468	0.0358	0.0273	0.0221	0.0177
$\hat{J}(\bar{t}^*)$	0.0395	0.0036	0.0012	0.0007	0.0005	0.0004

that the relative error

$$\mathcal{E}_o(\psi_1, \psi_2, \psi_3) = \left\| y_o \sum_{i=1}^3 \langle y_o, \psi_i \rangle_{L^2(\Omega)} \psi_i \right\|_{L^2(\Omega)}^2$$

is about 38% using the POD basis associated to  $\bar{t}^0$ , whereas  $\mathcal{E}_o$  is about 11% for the POD basis when we use the optimal time instance  $\bar{t}^*$ . In Table 2 the optimal time instance  $\bar{t}^*$  and optimal cost value  $\hat{J}$  are shown for different number  $\ell$  of POD basis functions.

## A. APPENDIX

*Proof of Proposition 2.4.* Let  $(c, d) = (c_1, c_2, c_3, c_4, d_1, d_2) \in \mathcal{Y} \times \mathbb{R}^{\sharp(\underline{A})} \times \mathbb{R}^{\sharp(\bar{A})}$  be arbitrary and consider the equations

$$\nabla e(x)\delta x = c, \quad \nabla g_{1, \underline{A}}(x^*)\delta x = d_1, \quad \nabla g_{2, \bar{A}}(x^*)\delta x = d_2, \quad (\text{A.1})$$

where we write  $x$  for  $x^*$  and  $\delta x \in \mathcal{X}$ . We set

$$(\delta \bar{t})_{\underline{A}} = -d_1, \quad (\delta \bar{t})_{\bar{A}} = d_2, \quad \text{and } \delta \bar{t}_i = 0 \text{ for } i \notin \underline{A} \cup \bar{A}.$$

Turning to the first, note that

$$\begin{aligned} \nabla e_1(x)\delta x = c_1 \\ \nabla e_2(x)\delta x = c_2 \end{aligned} \Leftrightarrow \begin{pmatrix} \delta \dot{\eta} - A^\ell \delta \eta - \delta A^\ell \eta - \delta f \\ \delta \eta(0) - \delta \eta_o \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix},$$

where  $\delta f$  and  $\delta \eta_o$  have been introduced in (2.8). Clearly, for any  $\delta \psi \in H^\ell$  there exists a unique  $\delta \eta \in W^{1,2}(0, T; \mathbb{R}^\ell)$  satisfying

$$\begin{pmatrix} \delta \dot{\eta} - A^\ell \delta \eta \\ \delta \eta(0) \end{pmatrix} = \begin{pmatrix} c_1 + \delta f \\ c_2 + \delta \eta_o \end{pmatrix}.$$

From (A.1) further

$$\nabla e_3(x)\delta x = c_3 \quad \text{and} \quad \nabla e_4(x)\delta x = c_4,$$

or equivalently,

$$\begin{pmatrix} (\mathcal{R}(\bar{t}) - \lambda_1) \delta\psi_1 - \delta\lambda_1 \psi_1 \\ \vdots \\ (\mathcal{R}(\bar{t}) - \lambda_\ell) \delta\psi_\ell - \delta\lambda_\ell \psi_\ell \end{pmatrix} = \tilde{c}_3$$

$$\begin{pmatrix} 2\langle \psi_1, \delta\psi_1 \rangle_H \\ \vdots \\ 2\langle \psi_\ell, \delta\psi_\ell \rangle_H \end{pmatrix} = c,$$

where

$$\tilde{c}_3 = c_3 - \begin{pmatrix} \sum_{k=1}^{\bar{k}} (\delta\bar{t})_k (\langle \dot{y}(\bar{t}_k), \psi_1 \rangle_H y(\bar{t}_k) + \langle y(\bar{t}_k), \psi_1 \rangle_H \dot{y}(\bar{t}_k)) \\ \vdots \\ \sum_{k=1}^{\bar{k}} (\delta\bar{t})_k (\langle \dot{y}(\bar{t}_k), \psi_\ell \rangle_H y(\bar{t}_k) + \langle y(\bar{t}_k), \psi_\ell \rangle_H \dot{y}(\bar{t}_k)) \end{pmatrix}.$$

We consider the last set of equations coordinate-wise, *i.e.*

$$(\mathcal{R}(\bar{t}) - \lambda_i) \delta\psi_i - \delta\lambda_i \psi_i = \tilde{c}_3^i \tag{A.2a}$$

$$2\langle \psi_i, \delta\psi_i \rangle_H = c_4^i \tag{A.2b}$$

for  $1 \leq i \leq \ell$ ,  $\tilde{g}_3^i \in H$ ,  $g_4^i \in \mathbb{R}$ . Each  $\delta\psi_i$  is decomposed according to

$$\delta\psi_i = (\delta\psi_i)_1 + (\delta\psi_i)_2 \in \ker(\mathcal{R}(\bar{t}) - \lambda_i) \oplus \ker(\mathcal{R}(\bar{t}) - \lambda_i)^\perp,$$

where  $\ker(\mathcal{R}(\bar{t}) - \lambda_i) = \text{span}\{\psi_i\}$ . From (A.2b) we have

$$(\delta\psi_i)_1 = c_4^i \psi_i / 2,$$

and from (A.2a),

$$\begin{aligned} \langle \tilde{c}_3^i, \psi_i \rangle_H &= \langle (\mathcal{R} - \lambda_i) \delta\psi_i, \psi_i \rangle_H - \delta\lambda_i \|\psi_i\|_H^2 \\ &= \langle \delta\psi_i, (\mathcal{R} - \lambda_i) \psi_i \rangle_H - \delta\lambda_i = -\delta\lambda_i. \end{aligned}$$

This implies that

$$\delta\lambda_i = -\langle \tilde{c}_3^i, \psi_i \rangle_H \quad \text{and} \quad \tilde{c}_3^i + \delta\lambda_i \psi_i \in \ker(\mathcal{R} - \lambda_i)^\perp = \text{range}(\mathcal{R} - \lambda_i).$$

Hence there exists a unique  $(\delta\psi_i)_2 \in \ker(\mathcal{R} - \lambda_i)^\perp$  such that

$$(\mathcal{R}(\bar{t}) - \lambda_i)(\delta\psi_i)_2 = g_3^1 + \delta\lambda_i \psi_i$$

and solvability of (A.1) follows. These computations also show that  $\nabla_{(\mathfrak{y}, \psi, \lambda)} e(x) \delta = r$  has a unique solution for any  $r \in \mathcal{Y}$ .  $\square$

## REFERENCES

- [1] G. Berkooz, P. Holmes and J.L. Lumley, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry – Cambridge Monographs in Mechanics*. Cambridge University Press, UK (1996).
- [2] T. Bui-Thanh, *Model-constrained optimization methods for reduction of parameterized systems*. Ph.D. Thesis, MIT, USA (2007).
- [3] T. Bui-Thanh, M. Damodoran and K. Willcox, Aerodynamic data reconstruction and inverse design using proper orthogonal decomposition. *AIAA Journal* **42** (2004) 1505–1516.

- [4] T. Bui-Thanh, K. Willcox, O. Ghattas and B. van Bloemen Wanders, Goal-oriented, model-constrained optimization for reduction of large-scale systems. *J. Comput Phys.* **224** (2007) 880–896.
- [5] R. Everson and L. Sirovich, The Karhunen-Loeve procedure for gappy data. *J. Opt. Soc. Am.* **12** (1995) 1657–1664.
- [6] K. Fukunaga, *Introduction to Statistical Recognition*. Academic Press, New York, USA (1990).
- [7] M.A. Grepl, Y. Maday, N.C. Nguyen and A.T. Patera, Efficient reduced-basis treatment of affine and nonlinear partial differential equations. *ESAIM: M2AN* **41** (2007) 575–605.
- [8] M. Heinkenschloss, *Formulation and Analysis of a Sequential Quadratic Programming Method for the Optimal Dirichlet Boundary Control of Navier Stokes Flow – Optimal Control: Theory, Methods and Applications*. Kluwer Academic Publisher, B.V. (1998) 178–203.
- [9] M. Hinze and K. Kunisch, Second order methods for optimal control of time – Dependent fluid flow. *SIAM J. Contr. Optim.* **40** (2001) 925–946.
- [10] K. Ito and S.S. Ravindran, A reduced-order method for simulation and control of fluid flows. *J. Comput. Phys.* **143** (1998) 403–425.
- [11] T. Kato, *Perturbation Theory for Linear Operators*. Springer Verlag, Germany (1980).
- [12] K. Kunisch and S. Volkwein, Control of Burgers’ equation by reduced order approach using proper orthogonal decomposition. *J. Optim. Theory Appl.* **102** (1999) 345–371.
- [13] K. Kunisch and S. Volkwein, Galerkin proper orthogonal decomposition methods for parabolic problems. *Numer. Math.* **90** (2001) 117–148.
- [14] S. Lall, J.E. Marsden and S. Glavaski, Empirical model reduction of controlled nonlinear systems, in *Proceedings of the IFAC Congress*, Vol. **F** (1999) 473–478.
- [15] H.V. Ly and H.T. Tran, Proper orthogonal decomposition for flow calculations and optimal control in a horizontal CVD reactor. *Quarterly Appl. Math.* **60** (2002) 631–656.
- [16] J. Nocedal and S.J. Wright, *Numerical Optimization, Springer Series in Operation Research*. Second Edition, Springer Verlag, New York, USA (2006).
- [17] M. Rathinam and L.R. Petzold, A new look at proper orthogonal decomposition. *SIAM J. Numer. Anal.* **41** (2003) 1893–1925.
- [18] S.S. Ravindran, Adaptive reduced-order controllers for a thermal flow system using proper orthogonal decomposition. *SIAM J. Sci. Comput.* **23** (2002) 1924–1942.
- [19] C.W. Rowley, Model reduction for fluids using balanced proper orthogonal decomposition. *Int. J. Bifur. Chaos* **15** (2005) 997–1013.
- [20] G. Rozza, D.B.P. Huynh and A.T. Patera, Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations: Application to transport and continuum mechanics. *Arch. Comput. Method. E.* **15** (2008) 229–275.
- [21] R. Temam, *Infinite Dimensional Dynamical Systems in Mechanics and Physics*. Second edition, Springer, Berlin, Germany (1997).
- [22] K. Willcox, O. Ghattas, B. von Bloemen Wanders and W. Bader, An optimization framework for goal-oriented, model-based reduction of large-scale systems, in *44th IEEE Conference on Decision and Control*, Sevilla, Spain (2005).