Mathematical Modelling and Numerical Analysis Will be set by the publisher
Modélisation Mathématique et Analyse Numérique

A-POSTERIORI ERROR ESTIMATION FOR SEMILINEAR PARABOLIC OPTIMAL CONTROL PROBLEMS WITH APPLICATION TO MODEL REDUCTION BY POD *

Eileen Kammann[1], Fredi Tröltzsch[1] and Stefan Volkwein[2]

Abstract. We consider the following problem of error estimation for the optimal control of nonlinear parabolic partial differential equations: Let an arbitrary admissible control function be given. How far is it from the next locally optimal control? Under natural assumptions including a second-order sufficient optimality condition for the (unknown) locally optimal control, we estimate the distance between the two controls. To do this, we need some information on the lowest eigenvalue of the reduced Hessian. We apply this technique to a model reduced optimal control problem obtained by proper orthogonal decomposition (POD). The distance between a local solution of the reduced problem to a local solution of the original problem is estimated.

1991 Mathematics Subject Classification. 49K20, 35J61, 35K58.

The dates will be set by the publisher.

Introduction

In this paper, we focus on the following question for a class of optimal control problems for semilinear parabolic equations: Suppose that a numerical approximation $\tilde{u}$ for a locally optimal control is given. For instance, it might have been obtained by a numerical optimization method or as the solution to some reduced order optimization model. How far is this control from the nearest locally optimal control $\bar{u}$? Of course, we have to assume that such a solution $\bar{u}$ exists in a neighborhood of $\tilde{u}$. Moreover, $\tilde{u}$ should already be sufficiently close to $\bar{u}$. The question is to quantify the error $\|\tilde{u} - \bar{u}\|$ in an appropriate norm.

In principle, this is not a paper on proper orthogonal decomposition (POD). Our estimation method is not restricted to any specific method of numerical approximation for $\tilde{u}$. Our primary goal is a numerical application of a perturbation method used by Arada et al. [2] in the context of FEM approximations of elliptic optimal control problems. The main idea of this method goes back to Dontchev et al. [8] and Malanowski et al. [21],

Keywords and phrases: Optimal control, semilinear partial differential equations, error estimation, proper orthogonal decomposition

* *The first and the second author are supported by* SFB 910 Control of self-organizing nonlinear systems: Theoretical methods and concepts of application *in Berlin. The second and the third author gratefully acknowledge the support by the German Science Fund DFG under project* A-posteriori-POD-Fehlerschätzer für nichtlineare Optimalsteuerprobleme bei partiellen Differentialgleichungen.

[1] Institut für Mathematik, Technische Universität Berlin, D-10623 Berlin, Germany; e-mail: kammann@math.tu-berlin.de & troeltzsch@math.tu-berlin.de

[2] Institut für Mathematik und Statistik, Universität Konstanz, D-78457 Konstanz; e-mail: Stefan.Volkwein@uni-konstanz.de

who introduced it for the a priori error analysis of optimal control problems governed by ordinary differential equations.

However, we will apply our method to suboptimal controls $\tilde{u}$ obtained by POD. Though POD is an excellent method of model reduction for many time-varying or nonlinear differential equations, it lacks an a priori error analysis that is in some sense uniform in the right-hand side of the underlying partial differential equation, say with respect to the control function. Estimates of this type are known for the method of balanced truncation (see, e.g. Benner et al. [4]) that is limited to linear and autonomous PDEs. There are results on a priori estimates for POD that depend on certain assumptions on the orthogonal basis generated by the selected snapshots. We refer to Kunisch and Volkwein [16], Sachs and Schu [25], or Tröltzsch and Volkwein [30]. However, such estimates will in general depend on the control used for generating the snapshots. In Hinze and Volkwein [11] an a priori error analysis is presented for linear-quadratic optimal control problems. If the POD basis is computed utilizing the optimal state and associated adjoint variable, a convergence rate can be shown. But in real computation we do not know the optimal solution in advance. In view of this, we are interested in a posteriori estimates for assessing the precision of optimal controls for reduced control problems set up by POD. For the reduced-basis method a posteriori error estimates for linear-quadratic optimal control problems we refer to Grepl and Kärcher [9].

We extend a method suggested in [30] for linear equations to the case of semilinear equations. For this purpose, we have to assume that $\bar{u}$ satisfies a standard second-order sufficient optimality condition. The associated coercivity constant of the reduced Hessian operator will be estimated numerically.

Let us explain our idea for the following two optimal control problems in a bounded Lipschitz domain $\Omega \subset \mathbb{R}^n$ with boundary Γ:

We consider the *distributed optimal control problem*

$$\min J(y,u) := \frac{1}{2}\int_0^T \int_\Omega \left\{ (y(x,t) - y_d(x,t))^2 + \lambda\, u(x,t)^2 \right\} dxdt \tag{PD}$$

subject to the semilinear parabolic state equation

$$\begin{aligned} \frac{\partial y}{\partial t} - \Delta y(x,t) + d(x,t,y(x,t)) &= u(x,t) && \text{in } \Omega \times (0,T) \\ y(x,t) &= 0 && \text{in } \Gamma \times (0,T) \\ y(x,0) &= y_0(x) && \text{in } \Omega \end{aligned} \tag{0.1}$$

and to the pointwise control constraints

$$\alpha \le u(x,t) \le \beta \quad \text{ a.e. in } \Omega \times (0,T).$$

Moreover, we deal with the *boundary control problem*

$$\min J(y,u) := \frac{1}{2}\int_0^T \int_\Omega (y(x,t) - y_d(x,t))^2\, dxdt + \frac{\lambda}{2}\int_0^T \int_\Gamma u(x,t)^2\, ds(x)dt \tag{PB}$$

subject to the parabolic state equation

$$\begin{aligned} \frac{\partial y}{\partial t} - \Delta y(x,t) &= 0 && \text{in } \Omega \times (0,T) \\ \frac{\partial y}{\partial \nu}(x,t) + b(x,t,y(x,t)) &= u(x,t) && \text{in } \Gamma \times (0,T) \\ y(x,0) &= y_0(x) && \text{in } \Omega \end{aligned} \tag{0.2}$$

and to the pointwise control constraints

$$\alpha \le u(x,t) \le \beta \quad \text{ a.e. in } \Gamma \times (0,T).$$

Here, and throughout the paper, $\nu \in \mathbb{R}^n$ denotes the outward unit normal on Γ.

Our numerical analysis is based on the following assumptions:

(A1) $\Omega \subset \mathbb{R}^n$ is a bounded domain with Lipschitz boundary Γ.

(A2) The functions $d : \Omega \times [0,T] \times \mathbb{R} \to \mathbb{R}$ and $b : \Gamma \times [0,T] \times \mathbb{R} \to \mathbb{R}$ are measurable with respect to the first two variables (x,t) and twice continuously differentiable with respect to the third variable $y \in \mathbb{R}$.

The functions b and d are monotone non-decreasing with respect to the third variable $y \in \mathbb{R}$ for all fixed $(x,t) \in \Gamma \times [0,T]$ and $\Omega \times [0,T]$, respectively.

(A3) There is a constant $C_\infty > 0$ such that

$$|d(x,t,0)| + \left|\frac{\partial d}{\partial y}(x,t,0)\right| + \left|\frac{\partial^2 d}{\partial y^2}(x,t,0)\right| \le C_\infty \quad \text{for a.a. } (x,t) \in \Omega \times [0,T]$$

and

$$|b(x,t,0| + \left|\frac{\partial b}{\partial y}(x,t,0)\right| + \left|\frac{\partial^2 b}{\partial y^2}(x,t,0)\right| \le C_\infty \quad \text{for a.a. } (x,t) \in \Gamma \times [0,T].$$

Moreover, for all $M > 0$ there exists a constant $L_M > 0$ such that

$$\left|\frac{\partial^2 d}{\partial y^2}(x,t,y_1) - \frac{\partial^2 d}{\partial y^2}(x,t,y_2)\right| \le L_M \, |y_1 - y_2|$$

holds for a.a. $(x,t) \in \Omega \times [0,T]$ and all $y_1, y_2 \in \mathbb{R}$ satisfying $|y_1 - y_2| < M$. We require the same condition for b for a.a. $(x,t) \in \Gamma \times [0,T]$.

(A4) A desired state function $y_d \in \mathrm{L}^2(\Omega \times (0,T))$, an initial function $y_0 \in \mathrm{L}^2(\Omega)$, and real constants $\lambda > 0$, $\alpha < \beta$ are given.

The paper is organized as follows: After this introduction, we explain the perturbation method and its use for the simpler case of a linear equation in Section 1. In Section 2, we extend the perturbation idea to semilinear equations. To make the paper also readable for readers being not familiar with POD, we briefly introduce the main concept of POD in Section 3, and Section 4 is devoted to the numerical application of our method to different nonlinear state equations.

1. A survey on the linear-quadratic case

1.1. The perturbation method for minimizing quadratic functionals

The perturbation method was introduced in [8, 21] in the context of optimal control of ordinary differential equations, where it was used to derive a priori error estimates for associated numerical approximations. In a different way, it was adapted in [2] to elliptic control problems. Let us explain this method for the following situation:

We consider the quadratic optimization problem in Hilbert space

$$\min_{u \in \mathcal{C}} f(u) := \frac{1}{2} \|Su - y_H\|_H^2 + \frac{\lambda}{2} \|u\|_{\mathrm{L}^2(D)}^2, \tag{PQ}$$

where H is a real Hilbert space, $D \subset \mathbb{R}^m$ is a measurable bounded set, $\mathcal{C} \subset \mathrm{L}^2(D)$ is a nonempty, convex, closed and bounded set, $S : \mathrm{L}^2(D) \to H$ a continuous linear operator, $\lambda > 0$ a fixed number, and $y_H \in H$ a fixed element. In this form, optimal control problems with partial differential equations are very frequently discussed. We refer only to the monography by Lions [18] or to the more recent books by Ito and Kunisch [14], Hinze et al. [10] or to the textbook by Tröltzsch [29]. We will mainly refer to the last reference, since the discussion of optimal control problems for semilinear parabolic equations in this book is very close to our notation.

Theorem 1.1. *Under the assumptions formulated above,* (**PQ**) *has a unique optimal solution* $\bar{u} \in \mathcal{C}$*. It satisfies the variational inequality*

$$\int_D \{(S^*(S\bar{u} - y_H))(x) + \lambda\, \bar{u}(x)\}\, (u(x) - \bar{u}(x))\, dx \geq 0 \quad \forall\ u \in \mathcal{C}, \tag{1.3}$$

where, $S^* : H \to \mathrm{L}^2(D)$ *denotes the adjoint operator to* S*.*

This is a standard result of the optimization theory, we refer e.g. to [18] or to [29, Thm. 2.22]. Let now $\mathcal{C}$ be the specific set $\mathcal{C} = \{u \in \mathrm{L}^2(D) : \alpha \leq u(x) \leq \beta \text{ for a.a. } x \in D\}$ with given real numbers $\alpha < \beta$. We introduce the adjoint state $\bar{p}$ associated with $\bar{u}$ by

$$\bar{p}(x) := (S^*(S\bar{u} - y_H))(x).$$

The variational inequality can be expressed equivalently in a pointwise way, namely

$$(\bar{p}(x) + \lambda\bar{u}(x))\bar{u}(x) = \min_{v \in [\alpha,\beta]} (\bar{p}(x) + \lambda\bar{u}(x))\, v \tag{1.4}$$

for almost all $x \in D$. This has important consequences. We have for almost all $x \in D$

$$\bar{u}(x) = \begin{cases} \alpha, & \text{if } \bar{p}(x) + \lambda\bar{u}(x) > 0 \\ -\bar{p}(x)/\lambda, & \text{if } \bar{p}(x) + \lambda\bar{u}(x) = 0 \\ \beta, & \text{if } \bar{p}(x) + \lambda\bar{u}(x) < 0. \end{cases}$$

While these inequalities show, how $\bar{u}$ is determined by $\bar{p} + \lambda\bar{u}$, the implications below concern the opposite direction; they follow directly from the discussion above. It must hold

$$\begin{cases} \bar{p}(x) + \lambda\bar{u}(x) \geq 0, & \text{if } \bar{u}(x) = \alpha \\ \bar{p}(x) + \lambda\bar{u}(x) = 0, & \text{if } \alpha < \bar{u}(x) < \beta \\ \bar{p}(x) + \lambda\bar{u}(x) \leq 0, & \text{if } \bar{u}(x) = \beta. \end{cases} \tag{1.5}$$

The last implications are essential for understanding the perturbation method. It is helpful to answer the following question: Let $\tilde{u} \in \mathcal{C}$ be a function that need not be optimal for (**PQ**). Let $\tilde{p}$ be the associated adjoint state.

If $\tilde{u}$ were optimal, then $\tilde{p}(x) + \lambda\tilde{u}(x) = 0$ should be satisfied in almost all points $x \in \Omega$, where $\alpha < \tilde{u}(x) < \beta$ holds. If not, then $\tilde{p}(x) + \lambda\tilde{u}(x) + \zeta(x) = 0$, if we define a *perturbation function* ζ in these points by

$$\zeta(x) = -[\tilde{p}(x) + \lambda\tilde{u}(x)].$$

In the points, where $\tilde{u}(x) = \alpha$ holds, the inequality $\tilde{p}(x) + \lambda\tilde{u}(x) \geq 0$ should be satisfied for optimality. If not, then $\tilde{p}(x) + \lambda\tilde{u}(x) + \zeta(x) \geq 0$ is fulfilled with

$$\zeta(x) = [\tilde{p}(x) + \lambda\tilde{u}(x)]_-,$$

where $[a]_-$ is defined for a real number a by $[a]_- = (|a| - a)/2$. Analogously, we define ζ in the points, where $\tilde{u}(x) = \beta$. In this way, we obtain the following definition of ζ that is adopted from Arada et al. [2],

$$\zeta(x) := \begin{cases} [\tilde{p}(x) + \lambda\tilde{u}(x)]_-, & \text{if } \tilde{u}(x) = \alpha, \\ -[\tilde{p}(x) + \lambda\tilde{u}(x)], & \text{if } \alpha < \tilde{u}(x) < \beta, \\ -[\tilde{p}(x) + \lambda\tilde{u}(x)]_+, & \text{if } \tilde{u}(x) = \beta. \end{cases} \tag{1.6}$$

The main idea behind the definition of ζ is the following: Although $\tilde{u}$ will possibly not be optimal for (**PQ**), it is optimal for the *perturbed optimization problem*

$$\min_{u\in\mathcal{C}} f(u) := \frac{1}{2}\,\|Su - y_H\|_H^2 + \frac{\lambda}{2}\,\|u\|^2_{\mathrm{L}^2(D)} + \int_D \zeta(x)\,u(x)\,dx. \tag{$\mathbf{PQ}_\zeta$}$$

This follows from our construction, because $\tilde{u}$ satisfies the associated variational inequality

$$\int_D (\tilde{p}(x) + \lambda\tilde{u}(x) + \zeta(x))(u(x) - \tilde{u}(x))\,dx \geq 0 \quad \forall u \in \mathcal{C}.$$

By convexity, $\tilde{u}$ is optimal for ($\mathbf{PQ}_\zeta$). Now, it is easy to estimate the distance between $\tilde{u}$ and $\bar{u}$.

Lemma 1.2. *Suppose that $\bar{u}$ is optimal for* (**PQ**)*, $\tilde{u} \in \mathcal{C}$ is given arbitrarily, and ζ is defined by* (1.6)*. Then there holds*

$$\|\tilde{u} - \bar{u}\|_{\mathrm{L}^2(D)} \leq \frac{1}{\lambda}\,\|\zeta\|_{\mathrm{L}^2(D)}.$$

Proof. We just write down the variational inequalities satisfied by $\bar{u}$ and $\tilde{u}$, and insert the other function, respectively. This yields

$$\begin{aligned} (S^*(S\bar{u} - y_H) + \lambda\bar{u}, \tilde{u} - \bar{u})_{\mathrm{L}^2(D)} &\geq 0 \\ (S^*(S\tilde{u} - y_H) + \lambda\tilde{u}, \bar{u} - \tilde{u})_{\mathrm{L}^2(D)} + (\zeta, \bar{u} - \tilde{u})_{\mathrm{L}^2(D)} &\geq 0. \end{aligned}$$

Adding both inequalities, we arrive at

$$-\|S(\bar{u} - \tilde{u})\|^2_{\mathrm{L}^2(D)} - \lambda\,\|\bar{u} - \tilde{u}\|^2_{\mathrm{L}^2(D)} + (\zeta, \bar{u} - \tilde{u})_{\mathrm{L}^2(D)} \geq 0,$$

and hence, since $\|S(\bar{u} - \tilde{u})\|^2_{\mathrm{L}^2(D)} \geq 0$,

$$\lambda\,\|\bar{u} - \tilde{u}\|^2_{\mathrm{L}^2(D)} \leq \|\zeta\|_{\mathrm{L}^2(D)}\,\|\tilde{u} - \bar{u}\|_{\mathrm{L}^2(D)}.$$

Now the claimed result follows immediately. □

Notice that, given $\tilde{u}$, ζ can be computed. It is determined by the adjoint state $\tilde{p}$ associated with $\tilde{u}$.

1.2. **Application to a linear-quadratic control problem**

As a simple application, we consider the following linear-quadratic version of (**PD**),

$$\min J(y,u) := \frac{1}{2}\int_0^T \int_\Omega \left\{(y(x,t) - y_d(x,t))^2 + \lambda\,u(x,t)^2\right\}\,dxdt \tag{**PL**}$$

subject to

$$\begin{aligned} \frac{\partial y}{\partial t} - \Delta y &= u && \text{in } \Omega\times(0,T) \\ y &= 0 && \text{in } \Gamma\times(0,T) \\ y(\cdot,0) &= 0 && \text{in } \Omega \end{aligned} \tag{1.7}$$

and to the pointwise control constraints

$$\alpha \leq u(x,t) \leq \beta \quad \text{a.e. in } \Omega\times(0,T).$$

We consider y in the space

$$W(0,T) = \{y \in \mathrm{L}^2(0,T;H^1(\Omega)) \,:\, \frac{dy}{dt} \in \mathrm{L}^2(0,T;H^1(\Omega)')\},$$

cf. [18]. For simplicity, we have assumed $y_0 = y(\cdot,0) = 0$. The case $y_0 \neq 0$ is reduced to that with homogeneous initial condition in a standard way: We shift the solution $\hat{y}$ associated with homogeneous right-hand side u but inhomogeneous initial data y_0 to y_d. Then we consider $\hat{y}_d := y_d - \hat{y}$ instead of $\hat{y}$.

We introduce the notation $Q := \Omega \times (0,T)$ and $\Sigma := \Gamma \times (0,T)$ and define the set of admissible controls by

$$U_{ad} = \{u \in \mathrm{L}^2(Q) : \alpha \leq u(x,t) \leq \beta \quad \text{a.e. in } Q\}.$$

We are going to apply Lemma 1.2 with the choice $D := Q$, $\mathcal{C} := U_{ad}$, $y_H := y_d$.

The adjoint state $p \in W(0,T)$ associated with a control u is obtained from the *adjoint equation*

$$\begin{aligned} -\frac{\partial p}{\partial t} - \Delta\, p &= y_u - y_d && \text{in } \Omega \times (0,T) \\ p &= 0 && \text{in } \Gamma \times (0,T) \\ p(\cdot,T) &= 0 && \text{in } \Omega, \end{aligned} \tag{1.8}$$

where y_u denotes the state associated with u. It is a standard result that (**PL**) has a unique optimal control $\bar{u}$.

Let now $\tilde{u} \in U_{ad}$ be given. For instance, we might think of a control obtained as the solution to a model reduced optimal control problem. In our numerical test case, the model reduced problem is defined upon POD.

To estimate the distance between $\tilde{u}$ and the unknown exactly optimal control $\bar{u}$, we proceed as follows: First, we determine the exact state $\tilde{y} := y_{\tilde{u}}$ by solving the state equation of (**PL**) with control $u := \tilde{u}$. Next, we insert $\tilde{y}$ for y_u in the adjoint equation (1.8) to obtain the associated adjoint state $\tilde{p}$ as solution.

Now we are able to determine the perturbation $\zeta \in \mathrm{L}^2(Q)$ by (1.6) (take $x := (x,t)$ and $D := Q$). Finally, we arrive at the estimate

$$\|\tilde{u} - \bar{u}\|_{\mathrm{L}^2(Q)} \leq \frac{1}{\lambda} \|\zeta\|_{\mathrm{L}^2(Q)}.$$

In Studinger and Volkwein [27] and [30], this method of estimation was successfully applied to different optimal control problems with quadratic objective functional and linear state equation. It is also successfully applied for other reduced-basis approximations, see Tonn et al. [28]. For recent extension to nonlinear problems we refer to Kahlbacher and Volkwein [15], where the presented error estimates are utilized in a multilevel SQP algorithm. Notice that we tacitly assume to solve the state and adjoint equation for $\tilde{y}$ and $\tilde{p}$ exactly, i.e. we neglect associated discretization errors. The inclusion of such errors is subject of ongoing research.

2. The nonlinear case

2.1. The perturbation method for nonconvex functionals

We consider now the nonconvex but smooth optimization problem

$$\min_{u \in \mathcal{C}} f(u) := \frac{1}{2} \|G(u) - y_H\|_H^2 + \frac{\lambda}{2} \|u\|_{\mathrm{L}^2(D)}^2, \tag{P}$$

where $G : \mathrm{L}^\infty(D) \to H$ is a twice continuously Fréchet differentiable operator and all other quantities are defined as in (**PQ**). We assume that, for all $u \in \mathcal{C}$, the first- and second-order derivatives $G'(u) : \mathrm{L}^\infty(D) \to H$ and $G''(u) : \mathrm{L}^\infty(D) \times \mathrm{L}^\infty(D) \to H$ can be continuously extended to $\mathrm{L}^2(D)$: There exists some constant $c > 0$

not depending on u and v such that

$$\|G'(u)v\|_H \le c\,\|v\|_{\mathrm{L}^2(D)} \qquad \forall u \in \mathcal{C},\ \forall v \in \mathrm{L}^\infty(D) \tag{2.9}$$

$$\|G''(u)(v_1,v_2)\|_H \le c\,\|v_1\|_{\mathrm{L}^2(D)}\|v_2\|_{\mathrm{L}^2(D)} \qquad \forall u \in \mathcal{C},\ \forall v_1, v_2 \in \mathrm{L}^\infty(D) \tag{2.10}$$

with some constant $c > 0$. Then the operators $G'(u)$ and $G''(u)$ can also be applied to increments v, v_1, v_2 in $\mathrm{L}^2(D)$. Therefore, we can view $G'(u)$ as continuous linear operator from $\mathrm{L}^2(D)$ to H so that its adjoint operator maps continuously H to $\mathrm{L}^2(D)$. We should mention that the uniformity with respect to u in (2.9), (2.10) requires usually the boundedness of $\mathcal{C}$ in $\mathrm{L}^\infty(D)$.

Remark 2.1. This is the typical way to deal with the well-known two-norm discrepancy that usually occurs in nonlinear parabolic control problems. Here, G is not differentiable in $\mathrm{L}^2(D)$, but in $\mathrm{L}^\infty(D)$. On the other hand, $f''(u)$ is coercive in $\mathrm{L}^2(D)^2$ but not in $\mathrm{L}^\infty(D)^2$. This genuine difficulty was addressed first by Ioffe [13] and discussed later by Malanowski [20] in the context of optimal control of ordinary differential equations. The treatment of this problem in second-order sufficient optimality conditions for the control of semilinear parabolic PDEs is explained in [29, Sections 4.10 and 5.7].

The derivative $f'(u)$ is given by

$$\begin{aligned} f'(u)\,v &= (G(u) - y_H, G'(u)\,v)_H + \lambda(u,v)_{\mathrm{L}^2(D)} \\ &= (G'(u)^*(G(u) - y_H) + \lambda u, v)_{\mathrm{L}^2(D)}. \end{aligned} \tag{2.11}$$

The $\mathrm{L}^2(D)$-function

$$p_u := G'(u)^*(G(u) - y_H) \tag{2.12}$$

is the *adjoint state* associated with u.

Let us determine for convenience also the second-order derivative of $f : \mathrm{L}^\infty(D) \to \mathbb{R}$. Thanks to the assumptions on G, this derivative exists. To determine it, we consider the expression for f' with fixed increment $v := v_1 \in \mathrm{L}^\infty(D)$ and differentiate again in the direction v_2. We find by the chain and product rule

$$\begin{aligned} f''(u)(v_1,v_2) = {} & (G'(u)\,v_2, G'(u)\,v_1)_H \\ & + (G(u) - y_H, G''(u)(v_2,v_1))_H + \lambda(v_2,v_1)_{\mathrm{L}^2(D)}. \end{aligned} \tag{2.13}$$

By our assumptions on G, also $f''(u)$ can be continuously extended to a bilinear form on $\mathrm{L}^2(D) \times \mathrm{L}^2(D)$ and

$$|f''(u)(v_1,v_2)| \le c\,\|v_1\|_{\mathrm{L}^2(D)}\|v_2\|_{\mathrm{L}^2(D)} \qquad \forall u \in \mathcal{C},\ \forall v_1, v_2 \in \mathrm{L}^\infty(D).$$

Let us now assume that $\bar u \in \mathcal{C}$ is a locally optimal solution to $(\mathbf{P})$ in the sense of $\mathrm{L}^\infty(D)$, i.e. there is some $\rho > 0$ such that

$$f(u) \ge f(\bar u) \quad \forall u \in \mathcal{C} \ \text{ with } \|u - \bar u\|_{\mathrm{L}^\infty(D)} \le \rho.$$

We obtain the following extension of Theorem 1.1:

Theorem 2.2. *If $\bar u \in \mathcal{C}$ is a locally optimal solution of $(\mathbf{P})$ in the sense of $\mathrm{L}^\infty(D)$, then it obeys the variational inequality*

$$\int_D \{(G'(\bar u)^*(G(\bar u) - y_H))(x) + \lambda\,\bar u(x)\}\,(u(x) - \bar u(x))\,dx \ge 0 \quad \forall\ u \in \mathcal{C}. \tag{2.14}$$

This result follows from the variational inequality

$$f'(\bar u)(u - \bar u) \ge 0 \quad \forall u \in \mathcal{C},$$

cf. e.g. Lemma 4.18 in [29]. Here, we use the representation (2.11) of $f'(\bar u)$.

Let now $\bar{p} := p_{\bar{u}}$ be the adjoint state defined in (2.12) and again $\mathcal{C} = \{u \in \mathrm{L}^2(D) : \alpha \le u(x) \le \beta \text{ a.e. in } D\}$ with real numbers $\alpha < \beta$. Analogously to the quadratic problem (**PQ**), the variational inequality can be expressed in a pointwise form,

$$(\bar{p}(x) + \lambda\bar{u}(x))\bar{u}(x) = \min_{v\in[\alpha,\beta]} (\bar{p}(x) + \lambda\bar{u}(x))\, v \qquad \text{for a.a. } x \in D. \tag{2.15}$$

This is identical with (1.4), hence we can draw the same conclusions from (2.15) as in the linear case. Therefore, the perturbation ζ is defined exactly as in (1.6).

Let now be $\tilde{u} \in \mathcal{C}$ sufficiently close to $\bar{u}$. We pose the same question as for (**PQ**): Can we quantify the distance $\|\tilde{u} - \bar{u}\|$? Now, however, the situation is more complicated. We need some second-order information on $\bar{u}$. Assume that there exists some $\delta > 0$ such that the coercivity condition

$$f''(\bar{u})(v, v) \ge \delta \|v\|^2_{\mathrm{L}^2(D)} \qquad \forall v \in \mathrm{L}^2(D) \tag{2.16}$$

is satisfied. Then for any $0 < \delta' < \delta$ there exists a radius $r(\delta') > 0$ such that

$$f''(u)(v, v) \ge \delta' \|v\|^2_{\mathrm{L}^2(D)} \quad \forall\, u \text{ with } \|u - \bar{u}\|_{\mathrm{L}^\infty(D)} \le r(\delta'),\ \forall v \in \mathrm{L}^2(D). \tag{2.17}$$

If $\tilde{u}$ belongs to this neighborhood, then we are able to estimate the distance.

Remark 2.3. To be on the safe side, we might select $\delta' := \delta/2$ and set $r := r(\delta/2)$. This can be too pessimistic. In the application to POD we are mainly interested in the order of the error so that the factor $1/2$ is not that important. We use $\delta' := \delta$, although this can be slightly too optimistic. ◊

Theorem 2.4. *Let $\bar{u}$ be locally optimal for* (**P**) *and assume that $\bar{u}$ satisfies the second-order condition* (2.16). *If $\tilde{u} \in \mathcal{C}$ is given such that $\|\tilde{u} - \bar{u}\|_{\mathrm{L}^\infty(D)} \le r$, where r is the radius introduced in* (2.17) *for $\delta' := \delta/2$, then it holds*

$$\|\tilde{u} - \bar{u}\|_{\mathrm{L}^2(D)} \le \frac{2}{\delta} \|\zeta\|_{\mathrm{L}^2(D)},$$

where ζ is defined as in (1.6).

Proof. As in the convex quadratic case, $\tilde{u}$ satisfies the first-order necessary optimality conditions for the perturbed optimization problem

$$\min_{u\in\mathcal{C}} f(u) + (\zeta, u)_{\mathrm{L}^2(D)}. \tag{$\mathbf{P}_\zeta$}$$

We insert $\bar{u}$ in the variational inequality for $\tilde{u}$ and $\tilde{u}$ in the one for $\bar{u}$ and obtain

$$\begin{aligned} (f'(\tilde{u}) + \zeta, \bar{u} - \tilde{u})_{\mathrm{L}^2(D)} &\ge 0 \\ (f'(\bar{u}), \tilde{u} - \bar{u})_{\mathrm{L}^2(D)} &\ge 0. \end{aligned}$$

We add both inequalities and find

$$(f'(\tilde{u}) - f'(\bar{u}), \bar{u} - \tilde{u})_{\mathrm{L}^2(D)} + (\zeta, \bar{u} - \tilde{u})_{\mathrm{L}^2(D)} \ge 0.$$

The mean value theorem implies

$$-f''(\hat{u})(\bar{u} - \tilde{u})^2 + (\zeta, \bar{u} - \tilde{u})_{\mathrm{L}^2(D)} \ge 0$$

with some $\hat{u} \in \{v \in \mathrm{L}^2(D) \,|\, v = s\bar{u} + (1 - s)\tilde{u} \text{ with } s \in (0, 1)\}$. Now we apply (2.17) and the Cauchy-Schwarz inequality to deduce

$$\frac{\delta}{2} \|\tilde{u} - \bar{u}\|^2_{\mathrm{L}^2(D)} \le \|\zeta\|_{\mathrm{L}^2(D)} \|\tilde{u} - \bar{u}\|_{\mathrm{L}^2(D)}.$$

¿From this inequality, the assertion of the theorem follows in turn. □

2.2. **Application to the parabolic boundary control problem (PB)**

Our general result can be applied in various ways. Let us explain its use for the parabolic boundary control problem (**PB**). From the theoretical point of view, (**PB**) is more difficult than the distributed problem (**PD**). For (**PB**), the control-to-state mapping G is not differentiable from $\mathrm{L}^2(\Sigma)$ to $C(\bar{Q})$. Here, the two-norm discrepancy, mentioned in Section 2.1, is a genuine issue. We refer, for instance, to the extensive discussion of this issue in [29] and to the paper by Casas et al. [5] that shows in particular, how the two-norm discrepancy can be overcome in problems with control and state constraints. In (**PD**), this difficulty does not occur for the spatial dimension $n = 1$, while it also arises for all $n \geq 2$. Therefore, we selected (**PB**) for our analysis.

We define

$$U_{ad} = \{u \in \mathrm{L}^\infty(\Sigma) : \alpha \leq u(x,t) \leq \beta \ \text{ for a.a. } (x,t) \in \Sigma\}. \tag{2.18}$$

The following result is known on the solvability of the state equation:

Theorem 2.5. *For all $u \in U_{ad}$, the state equation (0.2) has a unique solution $y_u \in Y := W(0,T) \cap C(\bar{Q})$. The mapping $G : u \mapsto y_u$ is twice continuously differentiable from $\mathrm{L}^\infty(\Sigma)$ to Y. In particular, it is twice continuously differentiable from $\mathrm{L}^\infty(\Sigma)$ to $H := \mathrm{L}^2(Q)$. The first- and second-order derivatives G' and G'' obey the extension conditions (2.9), (2.10) for the choice $D := Q$, $\mathcal{C} = U_{ad}$.*

We refer to [29, Theorems 5.9 and 5.16]. The extension conditions follow immediately from the representation formulas for these derivatives stated in these theorems. The same theorems include also the result for the problem (**PD**). Here, the reduced functional is given by

$$f(u) := \frac{1}{2}\int_0^T\int_\Omega (y_u(x,t) - y_d(x,t))^2 \, dxdt + \frac{\lambda}{2}\int_0^T\int_\Gamma u(x,t)^2 \, ds(x)dt.$$

Analogously to (2.11) and (2.12), the first derivative of the reduced functional can be expressed in the form

$$f'(u)\, v = \int_0^T\int_\Gamma (p_u(x,t) + \lambda u(x,t))v(x,t) \, ds(x)dt, \tag{2.19}$$

where p_u is the adjoint state associated with u. It is the unique solution to the adjoint equation

$$\begin{aligned} -\frac{\partial p}{\partial t} - \Delta p &= y_u - y_d && \text{in } \Omega \times (0,T) \\ \frac{\partial p}{\partial \nu} + \frac{\partial b}{\partial y}(x,t,y_u)\, p &= 0 && \text{in } \Gamma \times (0,T) \\ p(x,T) &= 0 && \text{in } \Omega. \end{aligned} \tag{2.20}$$

In view of our Theorem 2.2, any locally optimal control $\bar{u}$ of (**PB**) has to obey the following variational inequality:

$$\int_0^T\int_\Gamma (\bar{p}(x,t) + \lambda\bar{u}(x,t))(u(x,t) - \bar{u}(x,t)) \, ds(x)dt \geq 0 \quad \forall\, u \in U_{ad}, \tag{2.21}$$

where $\bar{p} := p_{\bar{u}}$.

The general form of the second derivative of the reduced functional was determined in (2.13). In this representation, terms of the form $G'(u)v$ appeared. For given $v \in \mathrm{L}^2(\Sigma)$, the function $y = G'(u)v$ is the unique solution of the *linearized parabolic equation*

$$\begin{aligned} \frac{\partial y}{\partial t} - \Delta y &= 0 && \text{in } \Omega \times (0,T) \\ \frac{\partial y}{\partial \nu} + \frac{\partial b}{\partial y}(x,t,y_u)\, y &= v && \text{in } \Gamma \times (0,T) \\ y(x,0) &= 0 && \text{in } \Omega, \end{aligned} \tag{2.22}$$

where $y_u = G(u)$ is the state associated with u, cf. [29, Theorem 5.9]. Therefore, the mapping $v \mapsto y$ is continuous from $\mathrm{L}^2(\Sigma)$ to $W(0,T)$, in particular to $\mathrm{L}^2(Q)$; this yields the first condition of continuous extension (2.9).

The second derivative $z = G''(u)(v_1, v_2)$ is the unique solution to the same equation with v substituted by $v := -\dfrac{\partial^2 b}{\partial y^2}(x,t,y_u)y_1y_2$, where $y_i = G'(u)v_i$, $i = 1,2$, and

$$\begin{aligned} \frac{\partial z}{\partial t} - \Delta z &= 0 && \text{in } \Omega \times (0,T) \\ \frac{\partial z}{\partial \nu} + \frac{\partial b}{\partial z}(x,t,y_u)\, z &= -\frac{\partial^2 b}{\partial y^2}(x,t,y_u)y_1y_2 && \text{in } \Gamma \times (0,T) \\ z(x,0) &= 0 && \text{in } \Omega. \end{aligned} \tag{2.23}$$

This is a linear equation with "control" $-\dfrac{\partial^2 b}{\partial y^2}(x,t,y_u)y_1y_2$. Therefore, it is not difficult to see that the second term in the representation (2.13) for $f''(u)$ can be re-written in terms of the adjoint state. Namely, it holds

$$\begin{aligned} f''(u)v^2 &= J''(y_u,u)(y,v)^2 - \int_0^T\!\!\int_\Omega p_u \frac{\partial^2 b}{\partial y^2}(x,t,y_u)\, y^2\, ds(x)dt, \\ &= \int_0^T\!\!\int_\Omega y^2\, dxdt + \lambda \int_0^T\!\!\int_\Gamma v^2\, ds(x)dt - \int_0^T\!\!\int_\Gamma p_u \frac{\partial^2 b}{\partial y^2}(x,t,y_u)\, y^2\, ds(x)dt. \end{aligned} \tag{2.24}$$

We refer, for instance to [29, Section 5.7] or to the recent paper by Casas and Tröltzsch [6], where second-order derivatives are determined for more general quasilinear equations. The second extension condition (2.10) follows immediately from (2.23).

In (2.24), y is the solution to the linearized equation (2.22). The expression above is just the second-order derivative of the Lagrangian function defined upon the adjoint state p_u associated with u. Therefore, the second-order sufficient optimality condition (2.16) means that

$$\begin{aligned} \int_0^T\!\!\int_\Omega y^2\, dxdt + \int_0^T\!\!\int_\Gamma \left\{ \lambda\, v^2 - p_u \frac{\partial^2 b}{\partial y^2}(x,t,y_u)\, y^2 \right\} ds(x)dt \\ \geq \delta \int_0^T\!\!\int_\Gamma v^2 ds(x)dt \quad \forall\, v \in \mathrm{L}^2(\Sigma). \end{aligned} \tag{2.25}$$

Since y is the solution of the linearized equation that is associated to v and the left-hand side of this equation is the second derivative of the Lagrangian function, this expresses a well known formulation of the (strong) second-order condition: The second derivative of the Lagrangian function is assumed to be coercive on the subspace defined by the linearized equation.

2.3. **Application of the perturbation method**

Let now $\tilde{u} \in U_{ad}$ be an arbitrary control that is close to a locally optimal control $\bar{u}$ of (**PB**). We assume that $\bar{u}$ satisfies the sufficient second-order optimality condition with some $\delta > 0$. In other words, the coercivity condition (2.25) is fulfilled for $u := \bar{u}$. Since the second-order condition is stable with respect to L^∞-perturbations of $\bar{u}$, we are justified to assume that there exists some $\rho > 0$ such that

$$f''(u)v^2 \geq \frac{\delta}{2} \int_0^T\!\!\int_\Gamma v^2 ds(x)dt \quad \forall\, v \in \mathrm{L}^2(\Sigma), \forall u \in B_\rho(\bar{u}), \tag{2.26}$$

where $B_\rho(\bar{u}) = \big\{ v \in \mathrm{L}^2(\Sigma) : \|\bar{u} - v\|_{\mathrm{L}^\infty(\Sigma)} < \rho \big\}$. We assume that $\tilde{u}$ belongs to $B_\rho(\bar{u})$. Now we define the perturbation function ζ as in (1.6), where $(x,t) \in \Omega \times [0,T]$ is substituted for x. Thanks to Theorem 2.4, we

obtain

$$\|\tilde{u} - \bar{u}\|_{\mathrm{L}^2(\Sigma)} \leq \frac{2}{\delta}\|\zeta\|_{\mathrm{L}^2(\Sigma)}. \tag{2.27}$$

We should mention already here a serious theoretical obstacle that can hardly be rigorously overcome. To apply our method of a posteriori estimation, we need the numbers ρ and δ. As ρ is concerned, we can only assume that the method of determining the (suboptimal) control $\tilde{u}$ was sufficiently precise so that $\|\tilde{u} - \bar{u}\| < \rho$.

There exists a method by Rösch and Wachsmuth [23, 24] to verify if there exists a locally optimal control $\bar{u}$ in a certain neighborhood of a given $\tilde{u}$. However, this tool is difficult to apply. It will work only for fairly special problems.

In addition, we must determine the coercivity constant δ. In computations, we deal with a finite-dimensional approximation of the control problem and determine the smallest eigenvalue of the reduced Hessian. In general, this way is not reliable in estimating the coercivity constant δ for the infinite-dimensional undiscretized optimal control problem: In [24], a discouraging example is constructed, where computations with very small mesh size indicate a sufficient second-order condition for a point, which is a saddle point. In special cases, the ideas of [23, 24] might be applied that rely on deep analytical estimations.

Summarizing this issue, we must confess that there is no reliable and at the same time practicable method to verify our assumptions absolutely certain. We have to trust that our problem behaves well in a neighborhood of the unknown solution $\bar{u}$. This is similar to problems of nonlinear programming, where optimization routines are started without evidence whether a constraint qualification is satisfied at the unknown solution or not.

We assume in the sequel that our computed suboptimal control belongs to a neighborhood of a locally optimal solution $\bar{u}$ that satisfies a second-order sufficient optimality condition. Moreover, we assume that we are able to determine the order of the coercivity constant δ by the lowest eigenvalue of the Hessian matrix associated with the suboptimal control.

3. Model reduction by POD

In this section, we give a very brief survey on how to establish a reduced order model of the parabolic boundary control problem (**PB**) by applying standard POD. For more information and proofs we refer the reader to [16] or Volkwein [32], for instance. A survey on methods of model reduction and a comparison of different reduction methods is contained in Antoulas [1].

3.1. The discrete POD method

Define $H = \mathrm{L}^2(\Omega)$ and let $u_s \in \mathrm{L}^2(\Sigma)$ be an arbitrary control with associated state $y_s = G(u_s) \in Y \subset \mathrm{L}^2(0,T;H)$. Since we cannot compute the whole trajectory $y_s(t)$ for all $t \in [0,T]$, we define a partition $0 = t_0 < t_1 < \ldots < t_{n_t} = T$ of the time interval $[0,T]$ and take snapshots $y_i^s := y_s(t_i) \in H$, $i = 0, \ldots, n_t$, of the state y_s at the given time instances $t_0, \ldots, t_{n_t}$.

Our goal is to find a small Galerkin basis that well expresses the main properties of the PDE in the finite dimensional subspace H^{n_t} defined by

$$H^{n_t} = \operatorname{span}\{y_i^s \,|\, i = 0, \ldots, n_t\} \subset H$$

with dimension $d = \dim H^{n_t}$. Let $\{\phi_1, \ldots, \phi_d\}$ be a basis of H^{n_t} that is orthonormal with respect to the inner product of $\mathrm{L}^2(\Omega)$. Then every $y_i^s = y_s(t_i)$, $i = 0, \ldots, n_t$, is a linear combination of $\phi_1, \ldots, \phi_d$,

$$y_i^s = \sum_{j=1}^{d} (y_i^s, \phi_j)_H \, \phi_j.$$

Now choose a fixed number $\ell \in \{1, \dots, d\}$. In the application to model reduction, ℓ is small compared with d. We intend to find an orthonormal set of functions $\{\varphi_j\}_{j=1}^{\ell} \subset H$ such that the sum of squared errors

$$\varepsilon_i = \left\| y_i^s - \sum_{j=1}^{\ell} (y_i^s, \varphi_j)_H \varphi_j \right\|_H^2$$

is minimized. In other words, we consider the optimization problem

$$\begin{aligned} \min_{\varphi_1,\dots,\varphi_\ell \in H} \quad & \sum_{i=0}^{n_t} \alpha_i \left\| y_i^s - \sum_{j=1}^{\ell} (y_i^s, \varphi_j)_H \, \varphi_j \right\|_H^2 =: \mathcal{J}^{n_t}(\varphi_1, \dots, \varphi_\ell), \\ \text{s.t.} \quad & (\varphi_i, \varphi_j)_H = \delta_{ij} \quad \forall\, i, j \in \{1, \dots, \ell\}, \end{aligned} \tag{3.28}$$

where $\alpha_0, \dots, \alpha_{n_t}$ are the trapezoidal weights

$$\alpha_0 = (t_1 - t_0)/2, \quad \alpha_i = (t_{i+1} - t_{i-1})/2, \quad \text{for} \quad i = 1, \dots, n_t - 1, \quad \alpha_{n_t} = (t_{n_t} - t_{n_t-1})/2.$$

The solution to (3.28) is obtained by solving a certain eigenvalue problem, described by the linear operator $\mathcal{R}^{n_t} : H \to H^{n_t}$ that maps $z \in H$ to

$$\mathcal{R}^{n_t} z = \sum_{i=0}^{n_t} \alpha_i \, (z, y_i^s)_H \, y_i^s. \tag{3.29}$$

The following proposition characterizes the operator $\mathcal{R}^{n_t}$ and shows that the eigenfunctions to the ℓ largest eigenvalues of $\mathcal{R}^{n_t}$ solve the problem (3.28).

Proposition 3.1. *Let H be a real Hilbert space, $y \in C([0,T], H)$, $H^{n_t} = \operatorname{span}\{y(t_i) \,|\, i = 0, \dots, n_t\} \subset H$ with dimension $d = \dim H^{n_t}$ and $\mathcal{R}^{n_t} : H \to H^{n_t}$ defined in* (3.29)*. Then the following statements hold:*

(i) *The operator $\mathcal{R}^{n_t}$ is bounded, self-adjoint, compact and non-negative.*

(ii) *There exist a sequence of eigenvalues $\{\lambda_i^{n_t}\}_{i=1}^{d}$ with $\lambda_1^{n_t} \geq \lambda_2^{n_t} \geq \dots \geq \lambda_d^{n_t} \geq 0$ and an associated orthonormal set $\{\psi_i^{n_t}\}_{i=1}^{d} \subset H$ of eigenelements that solve the eigenvalue problem*

$$\mathcal{R}^{n_t} \psi_i^{n_t} = \lambda_i^{n_t} \psi_i^{n_t}, \quad i = 1, \dots, d.$$

(iii) *For any fixed number $\ell \in \{1, \dots, d\}$, problem* (3.28) *is solved by the eigenfunctions $\{\psi_1^{n_t}, \dots, \psi_\ell^{n_t}\}$.*

(iv) *The corresponding minimal value of $\mathcal{J}^{n_t}$ is given by*

$$\inf_{\varphi_1,\dots,\varphi_\ell \in H} \mathcal{J}^{n_t}(\varphi_1, \dots, \varphi_\ell) = \mathcal{J}^{n_t}(\psi_1^{n_t}, \dots, \psi_\ell^{n_t}) = \sum_{i=\ell+1}^{d} \lambda_i^{n_t}.$$

For a proof we refer to e.g. [32], [31, Section 1.3] and Holmes et al. [12, Section 3]. Note that we have $\mathcal{R}^{n_t} = \mathcal{Y}\mathcal{Y}^*$, where the bounded linear operator $\mathcal{Y} : \mathbb{R}^{n_t+1} \to H^{n_t}$ and its Hilbert space adjoint operator $\mathcal{Y}^* : H \to \mathbb{R}^{n_t+1}$ are defined by

$$\mathcal{Y} q = \sum_{i=0}^{n_t} \alpha_i q_i \, y_i^s, \qquad (\mathcal{Y}^* z) = \Big((z, y_0^s)_H, \dots, \big(z, y_{n_t}^s\big)_H \Big)^\top.$$

Here, we used the notation $q = (q_0, \dots, q_{n_t})^\top$ for every $q \in \mathbb{R}^{n_t+1}$. Analogously to the theory of singular value decomposition for matrices, the linear, bounded, compact, and self-adjoint operator $\mathcal{K}^{n_t} = \mathcal{Y}^*\mathcal{Y} : \mathbb{R}^{n_t+1} \to \mathbb{R}^{n_t+1}$ defined by

$$(\mathcal{K}^{n_t} q)_i = \sum_{j=0}^{n_t} \alpha_j \left(y_j^s, y_i^s\right)_H q_j, \quad i = 0, \dots, n_t,$$

has the same eigenvalues $\{\lambda_i^{n_t}\}_{i=1}^d$ as $\mathcal{R}^{n_t}$. The corresponding eigenvectors of $\mathcal{K}^{n_t}$ are obtained by

$$(\Phi_i^{n_t})_j = \frac{1}{\sqrt{\lambda_i^{n_t}}}(\mathcal{Y}^*\psi_i^{n_t})_j = \frac{1}{\sqrt{\lambda_i^{n_t}}}(\psi_i^{n_t}, y_j^s)_H, \quad j \in \{0,\dots,n_t\},$$

for all $i \in \{1,\dots,\ell\}$. In many cases, the number (n_t+1) of time instances is much smaller than the number m_x of spatial ansatz functions. Then it is convenient to solve the eigenvalue problem

$$\mathcal{K}^{n_t}\Phi_i^{n_t} = \lambda_i^{n_t}\Phi_i^{n_t}, \quad i = 1,\dots,d, \tag{3.30}$$

instead of using the operator $\mathcal{R}^{n_t}$.

So far, we considered y_i^s as exact solution to the parabolic PDE at the time instances t_i. Now we apply a standard finite element approximation with respect to the spatial variable x. For this purpose, let $\mathcal{T}_h$ be a family of regular triangulations of Ω with mesh size $h > 0$. For each $\mathcal{T}_h$, let $V_h \subset \mathrm{L}^2(\Omega)$ be the space of piecewise linear and continuous functions spanned by the finite element basis functions $\{\varphi_1,\dots,\varphi_{m_x}\}$, $m_x \in \mathbb{N}$.

Then for every function $z \in V_h$ there exist $z_1,\dots,z_{m_x} \in \mathbb{R}$ such that

$$z(x) = \sum_{j=1}^{m_x} z_j\varphi_j(x), \quad \text{for all } x \in \Omega. \tag{3.31}$$

Thanks to our numerical approximation of the PDE, we can assume that our snapshots y_i^s belong to V_h. Therefore, for every $i \in \{0,\dots,n_t\}$ there exist $\gamma_1^i,\dots,\gamma_{m_x}^i \in \mathbb{R}$, such that

$$y_i^s = \sum_{j=1}^{m_x} \gamma_j^i\varphi_j, \tag{3.32}$$

where now y_i^s denotes the semidiscrete FEM-approximation of y_i^s. In order to avoid too many indices, we do not write $y_{i,h}^s$. Set

$$\mathcal{H}^{n_t} = \operatorname{span}\{y_i^s \mid i = 0,\dots,n_t\} \subset V_h.$$

For convenience, we introduce the vectors $\vec{y}_i^s = (\gamma_1^i,\dots,\gamma_{m_x}^i)^\top$. In this way, we have the correspondence $V_h \ni y_i^s \sim \vec{y}_i^s \in \mathbb{R}^{m_x}$. Let $z \in V_h$ be given. Then

$$(\mathcal{R}^{n_t}z)(x) \;=\; \sum_{i=0}^{n_t} \alpha_i\,(y_i^s, z)_{\mathrm{L}^2(\Omega)}\,y_i^s(x).$$

By definition, we have $\mathcal{R}^{n_t} : H \to \mathcal{H}^{n_t} \subset V_h$. Let us denote the restriction of $\mathcal{R}^{n_t}$ to V_h with range in V_h by the same symbol so that $\mathcal{R}^{n_t} : V_h \to V_h$. Applying equation (3.31) for every node $x_1,\dots,x_{m_x}$ of the triangulation of Ω, we infer for $V_h \ni z \sim \vec{z} \in \mathbb{R}^{m_x}$ and $V_h \ni y_i^s \sim \vec{y}_i^s \in \mathbb{R}^{m_x}$

$$\begin{aligned}
(\mathcal{R}^{n_t}z)(x_j) &= \sum_{i=0}^{n_t}\sum_{k=1}^{m_x} \alpha_i\,\gamma_k^i\,\varphi_k(x_j)\sum_{l,\nu=1}^{m_x}\gamma_l^i\,z_\nu\,(\varphi_l,\varphi_\nu)_{L^2(\Omega)} \\
&= \sum_{i=0}^{n_t}\sum_{l,\nu=1}^{m_x} \alpha_i\,\gamma_j^i\,\gamma_l^i\,(\varphi_l,\varphi_\nu)_{L^2(\Omega)}\,z_\nu \\
&= \sum_{i=0}^{n_t}\sum_{l,\nu=1}^{m_x} \alpha_i\,Y_{ji}\,Y_{li}(M_h)_{l\nu}\,z_\nu \;=\; \sum_{i=0}^{n_t}\sum_{l,\nu=1}^{m_x} \alpha_i\,Y_{ji}\,(Y^\top)_{il}(M_h)_{l\nu}\,z_\nu \\
&= \sum_{i=0}^{n_t} Y_{ji}\,D_{ii}\,(Y^\top M_h\,z)_i = \left(Y\,D\,Y^\top M_h\vec{z}\right)_j,
\end{aligned} \tag{3.33}$$

where $Y = \left[\vec{y}_0^s | \dots | \vec{y}_{n_t}^s\right]$, $D = \mathrm{diag}(\alpha_0, \dots, \alpha_{\mathrm{n_t}})$. In the first line, we used $\varphi_k(x_j) = \delta_{kj}$. M_h with entries

$$(M_h)_{kl} = (\varphi_k, \varphi_l)_{L^2(\Omega)}$$

is the mass matrix associated with the finite element basis functions. Define $\bar{\mathcal{R}}^{n_t} : \mathbb{R}^{m_x} \to \mathbb{R}^{m_x}$ by

$$\bar{\mathcal{R}}^{n_t} \vec{z} := \sum_{i=0}^{n_t} \alpha_i \left((\vec{y}_i^s)^\top M_h \, \vec{z} \right) \vec{y}_i^s.$$

Since

$$\begin{aligned} (\mathcal{R}^{n_t} z)(x_j) &= \sum_{i=0}^{n_t} \alpha_i \left(y_i^s, z\right)_{\mathrm{L}^2(\Omega)} y_i^s(x_j) = \sum_{i=0}^{n_t} \alpha_i (\vec{y}_i^s)^\top M_h \vec{z} \gamma_j^i \\ &= \left(\sum_{i=0}^{n_t} \alpha_i ((\vec{y}_i^s)^\top M_h \vec{z}) \, \vec{y}_i^s \right)_j = \left(\bar{\mathcal{R}}^{n_t} \vec{z} \right)_j, \end{aligned}$$

the operator $\mathcal{R}^{n_t}$ corresponds to the matrix $\bar{\mathcal{R}}^{n_t}$ in the numerical approximation. Moreover, by (3.33) we have

$$\bar{\mathcal{R}}^{n_t} \vec{z} = Y D \, Y^\top M_h \vec{z}.$$

To obtain an optimal solution of (3.28), we solve the eigenvalue problem

$$\bar{\mathcal{R}}^{n_t} \vec{\psi}_i^{n_t} = Y D Y^\top M_h \, \vec{\psi}_i^{n_t} = \lambda_i^{n_t} \vec{\psi}_i^{n_t}, \quad i = 1, \dots, \ell. \tag{3.34}$$

Defining $\hat{Y} = M_h^{\frac{1}{2}} Y D^{\frac{1}{2}}$, $\hat{\psi}_i^{n_t} = M_h^{\frac{1}{2}} \vec{\psi}_i^{n_t}$ and multiplying (3.34) by $M_h^{\frac{1}{2}}$ from the left, we arrive at the symmetric eigenvalue problem

$$\hat{Y} \hat{Y}^\top \hat{\psi}_i^{n_t} = \lambda_i^{n_t} \hat{\psi}_i^{n_t}, \quad i = 1, \dots, \ell.$$

Using similar arguments for the operator $\mathcal{K}^{n_t}$, we also can solve the discretized eigenvalue problem

$$\bar{\mathcal{K}}^{n_t} \vec{\Phi}_i^{n_t} = Y^\top M_h Y D \, \vec{\Phi}_i^{n_t} = \lambda_i^{n_t} \vec{\Phi}_i^{n_t}, \quad i = 1, \dots, \ell. \tag{3.35}$$

Defining $\hat{\Phi}_i^{n_t} = D^{\frac{1}{2}} \vec{\Phi}_i^{n_t}$ and multiplying (3.35) by $D^{\frac{1}{2}}$ from the left, we infer

$$\hat{Y}^\top \hat{Y} \hat{\Phi}_i^{n_t} = \lambda_i^{n_t} \hat{\Phi}_i^{n_t}, \quad i = 1, \dots, \ell.$$

Summarizing, we compute the *POD basis of rank* ℓ by the *method of snapshots* as follows:

Algorithm 3.2 (Method of Snapshots)**.** *Let* $m_x \gg n_t$.

(1) *Solve the symmetric eigenvalue problem of dimension* $(n_t + 1) \times (n_t + 1)$

$$\hat{Y}^\top \hat{Y} \hat{\Phi}_i^{n_t} = \lambda_i^{n_t} \hat{\Phi}_i^{n_t}, \quad i = 1, \dots, \ell.$$

(2) *Transform the eigenvectors* $\hat{\Phi}_1^{n_t}, \dots, \hat{\Phi}_\ell^{n_t}$ *by singular value decomposition to the POD basis vectors* $\vec{\psi}_1^{n_t}, \dots, \vec{\psi}_\ell^{n_t}$,

$$\vec{\psi}_i^{n_t} = M_h^{-\frac{1}{2}} \hat{\psi}_i^{n_t} = \frac{1}{\sqrt{\lambda_i^{n_t}}} M_h^{-\frac{1}{2}} \hat{Y} \hat{\Phi}_i^{n_t} = \frac{1}{\sqrt{\lambda_i^{n_t}}} Y D^{\frac{1}{2}} \hat{\Phi}_i^{n_t}, \quad i \in \{1, \dots, \ell\}.$$

Then the POD basis functions $V_h \ni \psi_i^{n_t} \sim \vec{\psi}_i^{n_t} \in \mathbb{R}^{m_x}, i = 1, \dots, \ell$, are given by

$$\psi_i^{n_t} = \sum_{j=1}^{m_x} (\vec{\psi}_i^{n_t})_j \varphi_j(x).$$

Remark 3.3. (1) In the method of snapshots, we do not have to set up the matrix $M_h^{\frac{1}{2}}$, since

$$\hat{Y}^\top \hat{Y} = D^{\frac{1}{2}} Y^\top M_h Y D^{\frac{1}{2}}.$$

(2) Note that (3.34) is an $m_x \times m_x$ eigenvalue problem, whereas (3.35) has dimension $(n_t+1)\times(n_t+1)$. In many applications we have $m_x \gg n_t$. Hence, the operator $\mathcal{K}^{n_t}$ is preferred for generating the POD basis. In the case, $m_x \leq n_t$, one should solve the problem (3.34) instead. Moreover, the singular value decomposition is much more stable in computing the small eigenvalues, see e.g. Lass and Volkwein [17] and [27] for this aspect in the context of POD. ◊

3.2. **POD Galerkin projection**

Let us consider exemplarily the state equation (0.2) of our boundary control problem (**PB**). For convenience, we use the short notation $(\cdot,\cdot)_\Omega := (\cdot,\cdot)_{\mathrm{L}^2(\Omega)}$ and $(\cdot,\cdot)_\Gamma := (\cdot,\cdot)_{\mathrm{L}^2(\Gamma)}$ for inner products. Then the corresponding variational formulation of (0.2) is given by

$$\begin{aligned} \frac{d}{dt}(y(t),\varphi)_\Omega + (\nabla y(t),\nabla\varphi)_\Omega + (b(\cdot,t,y(t)),\varphi)_\Gamma &= (u(t),\varphi)_\Gamma \\ (y(0),\varphi)_\Omega &= (y_0,\varphi)_\Omega, \end{aligned}$$

for all $\varphi \in H^1(\Omega)$ and for almost all $t \in [0,T]$. Under our assumptions on the given data, to each $u \in L^r(\Gamma\times(0,T))$ with $r>2$, there exists a unique solution $y \in W(0,T)\cap C(\bar{Q})$ of (0.2).

Let now $\{t_i\}_{i=0}^{n_t}$ be an equidistant partition of $[0,T]$. As above, we denote by V_h the space of piecewise linear functions spanned by the FE basis functions $\{\varphi_1,\ldots,\varphi_{m_x}\}$ defined on a regular triangulation $\mathcal{T}_h$ of Ω. To generate the snapshots, we select a fixed reference control $u_s \in L^r(\Sigma)$ and solve the semidiscrete equation for $y:[0,T]\to V_h$,

$$\begin{aligned} \frac{d}{dt}(y(t),\varphi_j)_\Omega + (\nabla y(t),\nabla\varphi_j)_\Omega + (b(\cdot,t,y(t)),\varphi_j)_\Gamma &= (u(t),\varphi_j)_\Gamma \\ (y(0),\varphi_j)_\Omega &= (y_0,\varphi_j)_\Omega, \end{aligned} \tag{3.36}$$

for all $j\in\{1,\ldots,m_x\}$ and for the given time instances t_i. As before, we denote the snapshots by $V_h \ni y_s(t_i) = y_i^s = \sum_{j=1}^{m_x}\gamma_i^j\varphi_j, i=0,\ldots,n_t$ and

$$\mathcal{H}^{n_t} = \operatorname{span}\{y_i^s \mid i\in\{0,\ldots,n_t\}\}$$

with dimension $d\in\mathbb{N}$. Next, we fix a small natural number $\ell \leq d$ and compute the POD basis functions $\{\psi_1^{n_t},\ldots,\psi_\ell^{n_t}\}$ as described in Section 3.1. Notice that all $\psi_i^{n_t}, i=1,\ldots,\ell$, are linear combinations of the FE basis functions $\varphi_1,\ldots,\varphi_{m_x}$.

Applying a Galerkin scheme based on the (small) POD basis, we obtain a nonlinear initial value problem of finding a function y^ℓ with

$$y^\ell(t) = \sum_{i=1}^{\ell}\eta_i(t)\psi_i^{n_t},$$

such that

$$\begin{aligned} \frac{d}{dt}(y^\ell(t),\psi_i^{n_t})_\Omega + (\nabla y^\ell(t),\nabla\psi_i^{n_t})_\Omega + (b(\cdot,t,y^\ell(t)),\psi_i^{n_t})_\Gamma &= (u(t),\psi_i^{n_t})_\Gamma \\ (y^\ell(0),\psi_i^{n_t})_\Omega &= (y_0,\psi_i^{n_t})_\Omega \end{aligned} \tag{3.37}$$

is satisfied for almost all $t\in[0,T]$ and all $i\in\{1,\ldots,\ell\}$. This is the state equation of a low size optimal control problem with state space of dimension ℓ, see problem (**PB.1**$^{\boldsymbol{\ell}}$) below. This small optimal control problem is then solved to obtain a suboptimal control $\bar{u}^\ell$.

Remark 3.4. It is not obvious that the reduced equation (3.37) has a unique solution, because the reduced Galerkin basis might destroy the monotonicity of the nonlinearity. To avoid this theoretical difficulty, we might truncate b as follows:

$$b_c(x,t,y) = \begin{cases} b(x,t,c) & \text{if } y \geq c \\ b(x,t,y) & \text{if } |y| \leq c/2 \\ b(x,t,-c) & \text{if } y \leq -c, \end{cases}$$

where b_c is defined in $(-c,-c/2) \cup (c/2,c)$ such that b_c is C^2 with respect to y. Then b_c is uniformly Lipschitz with respect to y so that the theorem by Picard and Lindelöf permits to show existence and uniqueness of a solution of (3.37) for all $T > 0$. In our computations, the function y^ℓ was uniformly bounded so that we did not follow this idea. ◊

4. A posteriori error estimation for POD solutions

To apply our a posteriori error estimation technique described in Section 2, we select a solution of some POD reduced optimal control problem. This is convenient, since model reduction by POD does not provide a rigorous a priori error analysis. Nevertheless, it often gives excellent results. This will be illustrated in the following two numerical examples.

4.1. **Example 1**

We first discuss a one-dimensional boundary control problem that is governed by a semilinear parabolic equation. The boundary control u is to drive a spatio-temporal temperature distribution to a predefined, desired temperature distribution at the final time T.

Let $\Omega = (0,1)$, $T = 1.58$ and $Q := \Omega \times (0,T)$. The optimal control problem is given by

$$\min \frac{1}{2}\int_0^1 \Big(y(x,T) - \frac{1}{2}(1-x^2)\Big)^2\,dx + \frac{1}{20}\int_0^T u(t)^2\,dt =: \mathcal{J}(y,u) \qquad \textbf{(PB.1)}$$

subject to the heat equation with Stefan-Boltzmann type boundary condition

$$\begin{aligned} y_t(x,t) - y_{xx}(x,t) &= 0 && \text{in } Q \\ y_x(0,t) &= 0 && \text{in } (0,T) \\ y_x(1,t) + y(1,t) + y^4(1,t) &= u(t) && \text{in } (0,T) \\ y(x,0) &= 0 && \text{in } \Omega \end{aligned} \tag{4.38}$$

and to the bilateral control constraints

$$-1 \leq u(t) \leq 1, \quad t \in (0,T).$$

This is a well-known test example that goes back to Schittkowski [26].

Remark 4.1. Formally, the function $y \mapsto y^4$ does not fulfill our assumption of monotonicity. Therefore, often $y \mapsto |y|y^3$ is considered instead. In our numerical tests, all occuring state functions y were non-negative. Therefore, we are justified to write y^4. ◊

The corresponding adjoint equation is defined by

$$\begin{aligned} -p_t(x,t) - p_{xx}(x,t) &= 0 && \text{in } Q \\ p_x(0,t) &= 0 && \text{in } (0,T) \\ p_x(1,t) + p(1,t) + 4\,y^3(1,t)p(1,t) &= 0 && \text{in } (0,T) \\ p(x,T) &= y(x,T) - y_d(x) && \text{in } \Omega. \end{aligned} \tag{4.39}$$

To reduce the state equation by POD, we selected an equidistant partition of the time interval, $0 = t_0 < t_1 < \ldots < t_{n_t} = T$, and a piecewise linear finite element discretization on the interval Ω with standard FE basis ("hat") functions $\{\varphi_1, \ldots, \varphi_{m_x}\}$. In all what follows, this time partition is only used to define the snapshots $y_s(t_i)$. For the ease of presentation, we tacitly assume to solve the occuring differential equations without time discretization. In our concrete computations, however, we used a semi-implicit Euler scheme in time to solve the resulting semidiscrete equation (3.36). Nevertheless, in the optimality conditions, in particular in the formulas for the computation of ζ, we ignore this time discretization.

Associated with the fixed control $u_s \equiv 0.5$, the snapshots

$$\left\{ y_i^s = y_s(t_i) = \sum_{j=1}^{m_x} \gamma_j^i \varphi_j \right\}_{i=0}^{n_t} \subset V_h \subset \mathrm{L}^2(\Omega)$$

were generated. Now we collect the snapshot vectors $\vec{y}_i^s = (\gamma_1^i, \ldots, \gamma_{m_x}^i)^\top$, $i = 0, \ldots, n_t$, in the matrix $Y = \left[\vec{y}_0^s | \ldots | \vec{y}_{n_t}^s\right] \in \mathbb{R}^{m_x \times (n_t+1)}$. As above, M denotes the mass matrix associated with $\{\varphi_1, \ldots, \varphi_{m_x}\}$ with entries

$$M_{ij} = (\varphi_i, \varphi_j)_{\mathrm{L}^2(\Omega)}, \quad i, j \in \{1, \ldots, m_x\}.$$

For convenience, we ignore the dependence on h. Furthermore, let $\tau = T/n_t$ and $\alpha_0, \ldots, \alpha_{n_t}$ be the trapezoidal weights

$$\alpha_0 = \alpha_{n_t} = \tau/2, \quad \alpha_i = \tau, \quad \text{for} \quad i = 1, \ldots, n_t - 1.$$

Define $D = \mathrm{diag}(\alpha_0, \ldots, \alpha_{n_t})$. Then the small orthonormal system $\{\psi_1^{n_t}, \ldots, \psi_\ell^{n_t}\}$ is obtained from the eigenvalue decomposition of $Y^\top MYD$, see Section 3.1, Algorithm 3.2.

We use the POD Galerkin ansatz for both the state y from (4.38) and the corresponding adjoint state p that solves the adjoint equation (4.39),

$$y_\ell(x,t) = \sum_{j=1}^{\ell} \eta_j(t) \psi_j^{n_t}(x), \qquad p_\ell(x,t) = \sum_{j=1}^{\ell} \rho_j(t) \psi_j^{n_t}(x).$$

The associated POD reduced optimal control problem is

$$\min \frac{1}{2} \int_0^1 \left(y_\ell(x,T) - \frac{1}{2}(1-x^2) \right)^2 dx + \frac{1}{20} \int_0^T u(t)^2 \, dt \qquad (\mathbf{PB.1^\ell})$$

subject to

$$\begin{aligned} M^\ell \frac{\partial}{\partial t} \vec{y}^\ell(t) + (K^\ell + F^\ell)\vec{y}^\ell(t) + E^\ell(\vec{y}^\ell(t)) &= B^\ell u(t) \\ \vec{y}^\ell(0) &= 0, \end{aligned} \qquad (4.40)$$

and

$$-1 \le u(t) \le 1,$$

for almost all $t \in (0,T)$. Here, $\vec{y}^\ell : [0,T] \to \mathbb{R}^\ell$ denotes the vector function $\vec{y}^\ell(t) = (\eta_1(t), \ldots, \eta_\ell(t))^\top$. The matrices $M^\ell, K^\ell, F^\ell \in \mathbb{R}^{\ell \times \ell}$, $B^\ell \in \mathbb{R}^\ell$ are given by

$$M_{ij}^\ell = (\psi_i^{n_t}, \psi_j^{n_t})_{\mathrm{L}^2(\Omega)}, \quad K_{ij}^\ell = ((\psi_i^{n_t})', (\psi_j^{n_t})')_{\mathrm{L}^2(\Omega)}, \quad F_{ij}^\ell = \psi_i^{n_t}(1)\psi_j^{n_t}(1), \quad B_i^\ell = \psi_i^{n_t}(1),$$

and $E^\ell : \mathbb{R}^\ell \to \mathbb{R}^\ell$ is defined by $y \mapsto (y^\top B^\ell)^4 B^\ell$.

The adjoint state $\vec{p}^\ell : [0,T] \to \mathbb{R}^\ell$ associated with $\vec{y}^\ell$ is the unique solution of the equation

$$\begin{aligned} -M^\ell \frac{\partial}{\partial t} \vec{p}^\ell(t) + (K^\ell + F^\ell)\vec{p}^\ell(t) + 4\tilde{E}^\ell(\vec{y}^\ell(t))F^\ell \vec{p}^\ell(t) &= 0 \\ \vec{p}^\ell(T) &= \vec{y}^\ell(T) - \vec{y}_d. \end{aligned} \qquad (4.41)$$

with $\tilde{E}^\ell : \mathbb{R}^\ell \to \mathbb{R}^\ell$ defined by $y \mapsto (y^\top B^\ell)^3 B^\ell$.

Remark 4.2. In general, the use of the same POD space for both the state and the adjoint state is not necessarily the best option. The adjoint state might be better approximated by an own POD basis. Nevertheless, we used only the snapshots of the state function y to set up the POD basis, because an extension by snapshots of p did not really improve the result.

The state equation of the associated discretized optimal control problem is solved again by a semi-implicit Euler scheme and the controls u are chosen as step functions according to the given partition $\{t_0, \ldots, t_{n_t}\}$ by

$$u(t) = u_i, \quad t_{i-1} \le t < t_i, \ i = 1, \ldots, n_t.$$

Therefore, we have the equivalence $u(\cdot) \sim (u_1, \ldots, u_{n_t})^\top = \vec{u}^\top \in \mathbb{R}^{n_t}$. To explain our method, for simplicity we only discretize the control but not the state and adjoint state functions.

This discretized optimal control problem was solved by an SQP method, where the associated sequence of linear-quadratic control systems was treated by a primal-dual active set strategy. Let us denote by $\vec{\bar{u}}^\ell = (\bar{u}_1^\ell, \ldots, \bar{u}_{n_t}^\ell)^\top$ the obtained locally suboptimal control vector, and by $\bar{u}_\tau^\ell(\cdot) \sim \vec{\bar{u}}^\ell$ the corresponding step function.

We are interested in estimating the distance between $\bar{u}_\tau^\ell$ and the next locally optimal control $\bar{u}$. For this purpose, we apply our perturbation method that was formulated in Section 2.3. This needs two full PDE solves. First, we have to determine the state $\bar{y}_\ell = G(\bar{u}_\tau^\ell)$, i.e. the solution of the parabolic boundary value problem (4.38) associated with $\bar{u}_\tau^\ell$. Second, we must compute the solution $\bar{p}_\ell$ of the full adjoint equation (4.39). Then the time-dependent perturbation ζ^ℓ is also a step function on $[0, T]$. It is equivalent to the vector $(\zeta_1, \ldots, \zeta_{n_t})^\top$ defined by

$$\zeta_i = \begin{cases} \left[\dfrac{1}{\tau}\displaystyle\int_{t_{i-1}}^{t_i} \bar{p}_\ell(1,t)dt + \lambda\,\bar{u}_i^\ell\right]_-, & \text{if} \quad \bar{u}_i^\ell = -1, \\ -\left[\dfrac{1}{\tau}\displaystyle\int_{t_{i-1}}^{t_i} \bar{p}_\ell(1,t)dt + \lambda\,\bar{u}_i^\ell\right], & \text{if} \quad -1 < \bar{u}_i^\ell < 1, \\ -\left[\dfrac{1}{\tau}\displaystyle\int_{t_{i-1}}^{t_i} \bar{p}_\ell(1,t)dt + \lambda\,\bar{u}_i^\ell\right]_+, & \text{if} \quad \bar{u}_i^\ell = 1. \end{cases}$$

for all $i \in \{1, \ldots, n_t\}$ and

$$\zeta^\ell(t) = \zeta_i^\ell, \quad t_{i-1} \le t < t_i.$$

In real computation, we are not able to determine the exact state $\bar{p}_\ell$. Therefore, we applied the mentioned semi-implicit Euler method to solve a fully discrete version of equation (4.41). Then, the associated adjoint state will correspond to a step function $\bar{p}_\tau^\ell$ in time with

$$\bar{p}_\tau^\ell(\cdot, t) = \bar{p}_i^\ell(\cdot), t \in [t_{i-1}, t_i), i = 1, \ldots, n_t,$$

and one can use $\frac{1}{\tau}\int_{t_{i-1}}^{t_i} \bar{p}_\ell(1,t)dt = \bar{p}_i^\ell(1)$.

We assume that the suboptimal control $\bar{u}_\tau^\ell$ belongs to an L^∞-ball around a locally optimal control $\bar{u}$, where the coercivity condition (2.26) is satisfied. In other words, we assume that $\bar{u}_\tau^\ell$ is sufficiently precise and the optimal control problem behaves well around $\bar{u}$.

Then we need an approximation of the coercivity constant δ. This will be accomplished by that of the reduced Hessian matrix Ψ^ℓ, associated with the suboptimal control $\vec{\bar{u}}^\ell$. We have to assume that Ψ^ℓ is positive definite. Let be σ_{min}^ℓ the smallest eigenvalue of Ψ^ℓ. Then there holds for all vectors $\vec{u} \in \mathbb{R}^\ell$ and their associated step functions u_τ that

$$\vec{u}^\top \Psi^\ell \vec{u} \ge \sigma_{min}^\ell |\vec{u}|_2^2 = \frac{\sigma_{min}^\ell}{\tau} \|u_\tau\|_{\mathrm{L}^2(0,T)}^2.$$

Therefore, if the control problem behaves well around $\bar{u}$, the coercivity constant δ' can be approximated by σ^ℓ_{min}/τ. Assuming that

$$\frac{\sigma^\ell_{min}}{\tau} \leq \delta'$$

holds, we can deduce that the distance of $\bar{u}^\ell_t au$ to the unknown locally optimal control $\bar{u}$ can be estimated by

$$\|\bar{u}^\ell_\tau - \bar{u}\|_{\mathrm{L}^2(0,T)} \leq \frac{\tau}{\sigma^\ell_{min}} \|\zeta^\ell\|_{\mathrm{L}^2(0,T)}.$$

In our numerical test, we computed a sufficiently precise approximation of $\bar{u}$, since this control was not exactly known. For this purpose, the interval $\Omega = (0,1)$ was split in 400 intervals $[x_j, x_{j+1}), j = 1, \ldots, m_x - 1$, (step size $h = \frac{1}{400}$), and we selected a time step size $\tau = \frac{T}{n_t} = \frac{T}{200}$. The solution of the accordingly discretized optimal control problem (**PB.1**) was the function $\bar{u}_\tau$ shown in Figure 1.

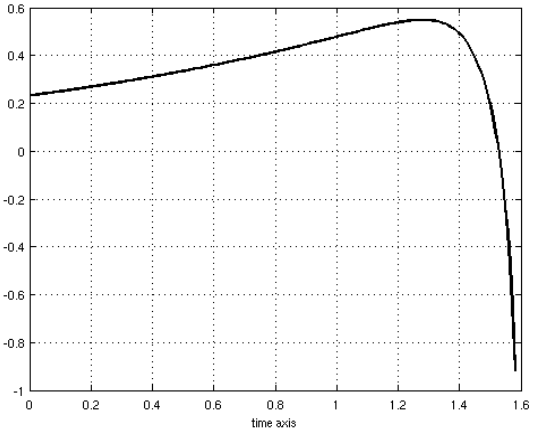

FIGURE 1. Example 1: Optimal FE control $\bar{u}_\tau$.

Table 1 compares the a posteriori error estimate for $\|\bar{u} - \bar{u}^\ell_\tau\|_{\mathrm{L}^2(0,T)}$ determined by our method with the "exact" numerical error $\|\bar{u}_\tau - \bar{u}^\ell_\tau\|_{\mathrm{L}^2(0,T)}$.

Remark 4.3. The left column of Table 1 shows some convergence for increasing ℓ, because both $\bar{u}_\tau$ and $\bar{u}^\ell_\tau$ are based on the same discretized model with discretization parameters τ and h. In contrast to this, the error estimator in the middle column stagnates at $3 \cdot 10^{-4}$. This behavior can be explained as follows: Even for the "exact" $\bar{u}_\tau$ the estimator delivers a value of 0.00030587 although one might expect a very small number. This value complies with the error estimate for $\|\bar{u} - \bar{u}_{\tau,h}\|$ explained in the next remark. Moreover, the discretized adjoint equation is not the one associated with the discretized model. The latter one would not be useful for our purpose, since the associated ζ would ignore the discretization error. We discretized the continuous adjoint equation by a semi-implicit Euler scheme. This causes another error of the discretization order. The estimator cannot be better than the discretization error; compare [27].

Remark 4.4. So far, we have tacitly assumed in our error estimates that, to determine a perturbation ζ, we are able to solve the state equation and the adjoint equation exactly. This is not possible in numerical applications. In solving these equations, numerical errors are unavoidable. There are two sources of errors, which are related to the approximation in time and the approximation in space. It is known from an early paper by Malanowski [19] and a recent publication by Neitzel and Vexler [22] that the error between a locally optimal control $\bar{u}$ and the corresponding discretized locally optimal control $\bar{u}_{\tau,h}$ satisfies the a priori error estimate $\|\bar{u} - \bar{u}_{\tau,h}\|_{\mathrm{L}^2(Q)} \leq C\,(\tau + h^2)$ with some constant C.

ℓ	$\|\bar{u}_\tau - \bar{u}_\tau^\ell\|_{\mathrm{L}^2(0,T)}$	$\frac{\tau}{\sigma_{min}^\ell}\|\zeta^\ell\|_{\mathrm{L}^2(0,T)}$	σ_{min}^ℓ
1	3.625e-1	6.441e-1	7.9724e-4
2	5.805e-2	6.498e-2	8.0598e-4
3	3.552e-3	4.191e-3	8.0686e-4
4	2.737e-4	5.043e-4	8.0686e-4
5	3.621e-5	2.964e-4	8.0686e-4
6	5.338e-6	3.081e-4	8.0686e-4
7	6.749e-7	3.056e-4	8.0686e-4
8	7.108e-8	3.059e-4	8.0686e-4
9	7.643e-9	3.059e-4	8.0686e-4

TABLE 1. Example 1: Numerical errors $\|\bar{u}_\tau - \bar{u}_\tau^\ell\|_{\mathrm{L}^2(0,T)}$ and error estimates for $\|\bar{u} - \bar{u}_\tau^\ell\|_{\mathrm{L}^2(0,T)}$ with corresponding eigenvalues σ_{min}^ℓ, $\ell = 1\ldots, 6$.

In our method, we want to estimate the error $\|\bar{u} - \bar{u}_\tau^\ell\|_{\mathrm{L}^2(0,T)}$, but actually we estimate the norm of $\bar{u}_{\tau,h} - \bar{u}_\tau^\ell$. By the triangle inequality and the a priori error estimate mentioned above, we obtain

$$\begin{aligned}\|\bar{u} - \bar{u}_\tau^\ell\|_{\mathrm{L}^2(0,T)} &\le \|\bar{u} - \bar{u}_{\tau,h}\|_{\mathrm{L}^2(0,T)} + \|\bar{u}_{\tau,h} - \bar{u}_\tau^\ell\|_{\mathrm{L}^2(0,T)} \\ &\le C\,(\tau + h^2) + \|\bar{u}_{\tau,h} - \bar{u}_\tau^\ell\|_{\mathrm{L}^2(0,T)} \sim C\,\tau + \|\bar{u}_{\tau,h} - \bar{u}_\tau^\ell\|_{\mathrm{L}^2(0,T)}\end{aligned} \tag{4.42}$$

since in our computations we have $\tau \gg h^2$. To make this estimate work, we need an estimation for C that is described below. ◊

Let us tacitly assume that the discretization in space can be neglected and denote by $\bar{u}_\tau$ the solution of the full system discretized only in time. To determine C in (4.42), we solve the (full size) discretized optimal control problem with mesh sizes τ and $\tau/2$, respectively. Then it holds

$$\begin{aligned}\|\bar{u}_\tau - \bar{u}_{\tau/2}\|_{\mathrm{L}^2(0,T)} &\le \|\bar{u}_\tau - \bar{u}\|_{\mathrm{L}^2(0,T)} + \|\bar{u} - \bar{u}_{\tau/2}\|_{\mathrm{L}^2(0,T)} \\ &\le C\,\tau + C\,\frac{\tau}{2} = \frac{3\tau}{2}\,C\end{aligned}$$

so that we can expect that

$$C \approx \frac{2}{3\tau}\,\|\bar{u}_\tau - \bar{u}_{\tau/2}\|_{\mathrm{L}^2(0,T)}.$$

In our example above, we applied this rough estimation technique by solving the problem with mesh sizes $\tau = T/200$ and $\tau/2 = T/400$, respectively. We obtained $C \approx 0.2741$ and therefore the discretization error can be estimated by

$$\|\bar{u} - \bar{u}_\tau\|_{\mathrm{L}^2(0,T)} \le 2.1654 \cdot 10^{-3}.$$

According to Table 1, this is of lower order than the error $\|\bar{u}_\tau - \bar{u}_\tau^\ell\|_{\mathrm{L}^2(0,T)}$. The computations show that a more detailed analysis on the influence of the discretization errors is desirable. Here, we do not consider this issue.

By the same SQP type optimization algorithm as for the full discretized problem, the POD-reduced optimal control problem (**PB.1**$^{\boldsymbol{\ell}}$) with $\ell = 4$ was solved within 2.3 seconds, whereas the optimization of the full problem (**PB.1**) needed 43 seconds, see Table 2. It turns out that the difference between the minimal values of the objective function $\mathcal{J}$ for the full FE solution and the POD solutions with $\ell = 4$ only amounts to the order $3 \cdot 10^{-6}$.

To obtain the smallest eigenvalue of the reduced Hessian that is needed for the a-posteriori estimate, we first constructed the reduced Hessian matrix Ψ^ℓ by applying the null space method. This is fairly expensive (11.24 seconds of the 13.24 seconds in Table 3). Next we used the Lanczos algorithm for computing its smallest eigenvalue.

	FEM	POD ($\ell = 4$)
CPU time for optimization	43 s	2.3 s
Minimal value of $\mathcal{J}$	0.023238	0.023240

TABLE 2. Example 1: CPU times for POD optimization and corresponding minimal values of the objective function in comparison with FE method.

($\ell = 4$)	CPU time
POD basis	1.50 s
Optimization	0.76 s
Computation of y_τ^ℓ	1.35 s
Computation of p_τ^ℓ	2.44 s
Error estimation	13.24 s

TABLE 3. Example 1: CPU times for POD optimization and a-posteriori error estimation.

4.2. **Example 2: A distributed control problem**

In our second example, we consider a distributed 2D optimal control problem, which is essentially more demanding, since distributed controls have more direct influence on the solution of the parabolic equation than boundary controls. Moreover, the dimension two of Ω increases the computational effort.

The system dynamics of the problem is given by a semilinear parabolic partial differential equation in $\Omega = (0,\pi)^2$. We constructed the example in such a way that we know the exact optimal solution. Here, a desired spatio-temporal heat distribution y_Q is to be pursued as close as possible.

Let be $T = 1$ and $Q := \Omega \times (0,T)$. Four time dependent controls $u^1, \dots, u^4 \in \mathrm{L}^2(0,T)$ are given that are distributed to Ω by some weight functions $w_i \in \mathrm{L}^2(\Omega)$, $i = 1, \dots, 4$.

Thus, we consider a distributed control $\hat{u}$ of the form

$$\hat{u}(x,t) = \sum_{i=1}^{4} w_i(x) u^i(t). \tag{4.43}$$

Furthermore, we define some auxiliary functions $a_y \in \mathrm{L}^2(\Omega)$ and $d \in \mathrm{L}^2(Q)$, which help us constructing an explicitely known optimal solution. Our objective functional to be minimized is

$$\mathcal{J}(y, u^1, \dots, u^4) = \frac{1}{2} \|y - y_Q\|^2_{\mathrm{L}^2(Q)} + (a_y, y(\cdot,T))_{\mathrm{L}^2(\Omega)} + \frac{1}{200} \sum_{i=1}^{4} \|u^i\|^2_{\mathrm{L}^2(0,T)}. \tag{PD.2}$$

The state equation is the semilinear heat equation

$$\begin{aligned} \frac{\partial y}{\partial t}(x,t) - \Delta y(x,t) + y^3(x,t) + d(x,t) &= \sum_{i=1}^{4} w_i(x) u^i(t) && \text{in } Q \\ \frac{\partial y}{\partial \nu}(x,t) &= 0 && \text{in } \Sigma \\ y(x,0) &= y_0(x) && \text{in } \Omega \end{aligned} \tag{4.44}$$

with given initial function $y_0 \in \mathrm{L}^2(\Omega)$ and homogeneous Neumann boundary condition. Moreover, we impose pointwise box constraints on the four controls,

$$|u^i(t)| \le 1, \quad \text{a.e. in } (0,T), \quad i = 1, \dots, 4.$$

Remark 4.5. (1) The cost functional (**PD.2**) does not exactly fit in the setting of problem (**PD**), where the linear term $(a_y, y(\cdot,T))_{\mathrm{L}^2(\Omega)}$ is missing. However, the associated theory of optimality conditions is fairly analogous to the one of (**PD**). Here, the adjoint equation contains an inhomogeneous final condition $p(\cdot,T) = a_y$. It looks as follows:

$$\begin{aligned} -\frac{\partial p}{\partial t}(x,t) - \Delta p(x,t) + 3y^2(x,t)p(x,t) &= y(x,t) - y_Q(x,t) && \text{in } Q \\ \frac{\partial p}{\partial \nu}(x,t) &= 0 && \text{in } \Sigma \\ p(x,T) &= a_y(x) && \text{in } \Omega. \end{aligned}$$

(2) Furthermore, the equation (4.44) does not exactly fit in problem (0.1). First, we have a Neumann boundary condition instead of a Dirichlet one. Second, the control function has the form (4.43). However, the theory is similar to that explained at the beginning of the paper. Moreover, we detail the necessary changes. The treatment of homogeneous Neumann boundary conditions is completely analogous to the Dirichlet case.

$\Diamond$

All given parameters are chosen such that locally optimal solutions are known in advance. For this purpose, we define with $x = (x_1, x_2)$

$$\begin{aligned} y_0(x) &= \cos(x_1)\cos(x_2), \\ y_Q(x,t) &= \cos(x_1)\cos(x_2)(1 + 0.02(t - t^2) - 0.03\, t^2 \cos^2(x_1)\cos^2(x_2)), \\ a_y(x) &= 0.01\cos(x_1)\cos(x_2), \\ w_1(x) &= \max\{0, 10 - 50(x_1 - x_2)(x_1 + x_2 - \tfrac{\pi}{2})\}, \\ w_2(x) &= \max\{0, 10 - 50(x_1 - x_2 - \tfrac{\pi}{2})(x_1 + x_2 - \pi)\}, \\ w_3(x) &= \max\{0, 10 - 50(x_1 - x_2)(x_1 + x_2 - \tfrac{3\pi}{2})\}, \\ w_4(x) &= \max\{0, 10 - 50(x_1 - x_2 + \tfrac{\pi}{2})(x_1 + x_2 - \pi)\}, \\ \beta^i(t) &= \mathbb{P}_{[-1,1]}\left\{-t^2 \int_\Omega w_i(x)\cos(x_1)\cos(x_2)dx\right\}, \quad i = 1,\dots,4, \\ d(x,t) &= \sum_{i=1}^{4} w_i(x)\beta^i(t) - 2\cos(x_1)\cos(x_2) - \cos^3(x_1)\cos^3(x_2). \end{aligned}$$

The graphs of the four functions $w_1, \dots, w_4$ are pictured in Figure 2. Consequently, locally optimal controls can be specified exactly as

$$\bar{u}^i(t) = \mathbb{P}_{[-1,1]}\left\{-t^2 \int_\Omega w_i(x)\cos(x_1)\cos(x_2)dx\right\},\ i = 1,\dots,4.$$

They are shown in Figure 3. The associated (locally) optimal state $\bar{y}$ and the corresponding adjoint state $\bar{p}$ are given by

$$\bar{y}(x,t) = \cos(x_1)\cos(x_2), \quad \bar{p}(x,t) = \frac{1}{100}\, t^2 \cos(x_1)\cos(x_2).$$

The domain Ω is discretized by a regular triangulation so that an FE space of $m_x = 687$ piecewise linear ansatz functions is obtained. For the snapshot generation we took an equidistant partition of the time interval $[0,T]$ with step size $\tau = 1/119$, so that $n_t = 120$ snapshots were computed. Then a number ℓ was chosen, the POD basis $\{\psi_i^{n_t}\}_{i=1}^{\ell}$ was computed as explained in Section 3.1 and we applied the standard POD Galerkin

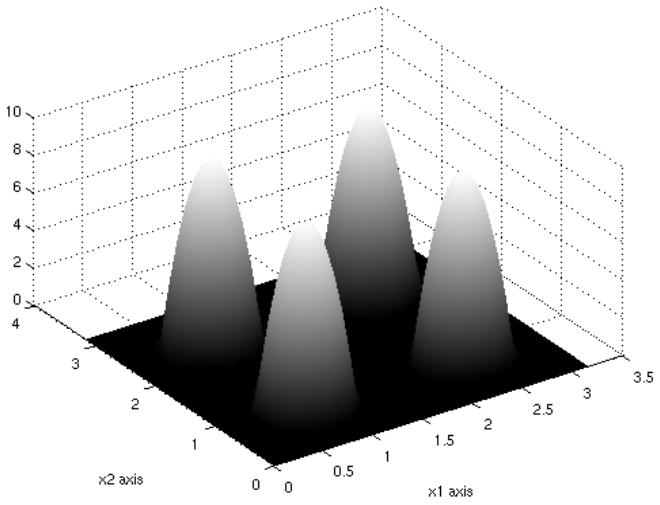

FIGURE 2. Example 2: Weight functions $w_1, \ldots, w_4$.

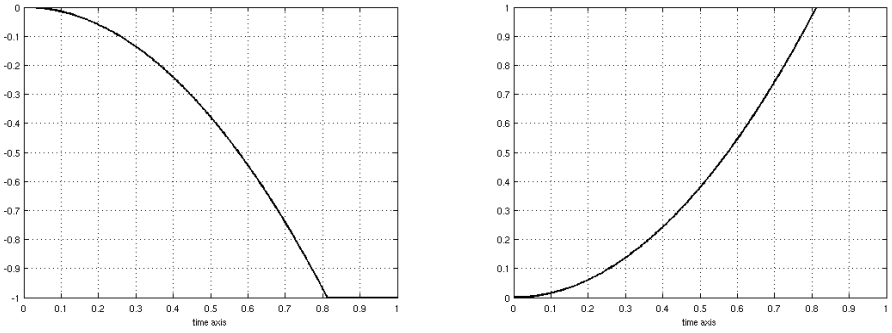

FIGURE 3. Example 2: Optimal controls $\bar{u}^1, \bar{u}^2$. For symmetry reasons it holds $\bar{u}^1 = \bar{u}^3, \bar{u}^2 = \bar{u}^4$.

ansatz for both the state y and the adjoint state p,

$$y_\ell(x,t) = \sum_{i=1}^{\ell} \eta_i(t)\psi_i^{n_t}(x), \qquad p_\ell(x,t) = \sum_{i=1}^{\ell} \rho_i(t)\psi_i^{n_t}(x),$$

as well as for the functions $w_1, \ldots, w_4 \in \mathrm{L}^2(\Omega)$ and $d \in \mathrm{L}^2(Q)$,

$$d_\ell(x,t) = \sum_{i=1}^{\ell} \delta_i(t)\psi_i^{n_t}(x), \qquad (w_k)_\ell(x) = \sum_{i=1}^{\ell} (\omega_k)_i \psi_i^{n_t}(x), \quad k = 1, \ldots, 4,$$

to obtain the reduced order model ($\mathbf{PD}.\mathbf{2}^{\boldsymbol{\ell}}$)

$$\min \frac{1}{2} \|y_\ell - y_Q\|^2_{\mathrm{L}^2(Q)} + (a_y, y_\ell(\cdot,T))_{\mathrm{L}^2(\Omega)} + \frac{1}{200} \sum_{i=1}^{4} \|u^i\|^2_{\mathrm{L}^2(0,T)} \qquad (\mathbf{PD}.\mathbf{2}^{\boldsymbol{\ell}})$$

subject to

$$\begin{aligned} M^\ell \frac{\partial}{\partial t}\vec{y}^\ell(t) + K^\ell \vec{y}^\ell(t) + E^\ell(\vec{y}^\ell(t)) + M^\ell \vec{d}^\ell(t) &= \sum_{i=1}^{4} M^\ell \vec{w}_i^\ell \, u^i(t) \\ \vec{y}^\ell(0) &= 0, \end{aligned} \tag{4.45}$$

and

$$|u_i(t)| \le 1, \quad i = 1, \dots, 4,$$

for almost all $t \in (0,T)$. As above, $\vec{y}^\ell : [0,T] \to \mathbb{R}^\ell$ denotes the vector function $\vec{y}^\ell(t) = (\eta_1(t), \dots, \eta_\ell(t))^\top$, and the matrices $M^\ell, K^\ell \in \mathbb{R}^{\ell \times \ell}$ are given by

$$M_{ij}^\ell = (\psi_i^{n_t}, \psi_j^{n_t})_{\mathrm{L}^2(\Omega)}, \quad K_{ij}^\ell = (\nabla\psi_i^{n_t}, \nabla\psi_j^{n_t})_{\mathrm{L}^2(\Omega)}, \quad i,j = 1, \dots, \ell,$$

and $E^\ell : \mathbb{R}^\ell \to \mathbb{R}^\ell$ is defined by

$$\big(E^\ell(y)\big)_i = \sum_{j,k,l=1}^{\ell} y_j y_k y_l \int_\Omega \psi_j^{n_t}(x)\psi_k^{n_t}(x)\psi_l^{n_t}(x)\psi_i^{n_t}(x)\,dx, \quad i = 1, \dots, \ell.$$

After discretizing the semidiscrete problem (**PD.2**$^{\boldsymbol{\ell}}$) by the semi-implicit Euler scheme with an equidistant partition $\{t_i\}_{i=0}^{n_t}$ of $[0,T]$, we applied an SQP method for solving the discretized reduced problem. In this context, the integration of linearized terms such as

$$\begin{aligned} \int_\Omega (y^k(x,t))^2 y(x,t)\psi_j^{n_t}(x)dx &= \int_\Omega \left(\sum_{i=1}^{\ell} \eta_i^k(t)\psi_i^{n_t}(x)\right)^2 \sum_{l=1}^{\ell} \eta_l(t)\psi_l^{n_t}(x)\psi_j^{n_t}(x)dx, \\ &= \sum_{i,\mu,l=1}^{\ell} \eta_i^k(t)\eta_\mu^k(t)\eta_l(t) \int_\Omega \psi_i^{n_t}(x)\psi_\mu^{n_t}(x)\psi_l^{n_t}(x)\psi_j^{n_t}(x)dx \end{aligned}$$

cannot be easily accomplished on using the mass matrix. We applied a prismoidal formula. Therefore, we interpolated the node data of the POD basis functions $\psi_1^{n_t}, \dots, \psi_\ell^{n_t}$ linearly to the triangle midpoints x_{τ_ν} of the finite elements τ_ν such that the cubature is given by

$$\int_\Omega \psi_i^{n_t}(x)\psi_\mu^{n_t}(x)\psi_l^{n_t}(x)\psi_j^{n_t}(x)dx \approx \sum_{\nu=1}^{\tilde{m}} A_{\tau_\nu} \psi_i^{n_t}(x_{\tau_\nu})\psi_\mu^{n_t}(x_{\tau_\nu})\psi_l^{n_t}(x_{\tau_\nu})\psi_j^{n_t}(x_{\tau_\nu}),$$

where the corresponding areas of the triangles of the finite element triangulation are denoted by A_{τ_ν} and $\tilde{m}$ is the number of triangles of the approximation of Ω. The linear-quadratic subproblems in each level of the SQP method were treated by the primal-dual active set strategy.

Remark 4.6. We should mention in this context that this ad-hoc method can certainly be improved by the application of the *discrete empirical interpolation method (DEIM)* as it was suggested by Barrault et al. [3] or Chaturantabut and Sorensen [7]. Let us also refer to the recent paper by Lass and Volkwein [17]. However, we did not concentrate on this technique, because our main concern is the problem of a posteriori error estimation. ◊

Now denote the obtained POD-optimal control vector by $\bar{u}_\tau^\ell = (\bar{u}_\tau^{\ell,1}, \dots, \bar{u}_\tau^{\ell,4})^\top$ and by $\bar{y}_\ell$ the associated state. Define the control-to-state mapping $G : \mathrm{L}^\infty(0,T;\mathbb{R}^4) \to \mathrm{L}^2(Q)$ that maps the control terms $u = (u^1, u^2, u^3, u^4)^\top \in \mathrm{L}^\infty(0,T;\mathbb{R}^4)$ to the associated state y.

We obtain the following corollary of Theorem 2.4 on error estimation:

Corollary 4.7. *Suppose there are a radius $r > 0$ and some $\delta > 0$ such that there holds*

$$f''(u)(v,v) \geq \delta \|v\|^2_{\mathrm{L}^2(0,T;\mathbb{R}^4)}$$

for all u with $\|u - \bar{u}\|_{\mathrm{L}^\infty(0,T;\mathbb{R}^4)} \leq r$ and for all $v \in \mathrm{L}^2(0,T;\mathbb{R}^4)$. Let $\ell \in \mathbb{N}$ be a fixed number and $\bar{u}^\ell_\tau \in \mathrm{L}^\infty(0,T;\mathbb{R}^4)$. If $\|\bar{u}^\ell_\tau - \bar{u}\|_{\mathrm{L}^\infty(0,T;\mathbb{R}^4)} \leq r$, then we have

$$\|\bar{u}^\ell_\tau - \bar{u}\|_{\mathrm{L}^2(0,T;\mathbb{R}^4)} \leq \frac{1}{\delta} \|\zeta\|_{\mathrm{L}^2(0,T;\mathbb{R}^4)}.$$

As before, the coercivity parameter δ is approximated by means of the smallest eigenvalue of the reduced Hessian matrix associated with the POD solution $\bar{u}^\ell_\tau$, as described in Section 4.1. Under the assumptions of Corollary 4.7, we have

$$\|\bar{u}^\ell_\tau - \bar{u}\|_{\mathrm{L}^2(0,T;\mathbb{R}^4)} \leq \frac{\tau}{\sigma^\ell_{min}} \|\zeta\|_{\mathrm{L}^2(0,T;\mathbb{R}^4)}.$$

The control functions u^1, u^2, u^3, u^4 are chosen as step functions according to the given time partition by

$$u^k_\tau(t) = u^k_i, \quad t_{i-1} \leq t < t_i, \ i = 1, \ldots, n_t,$$

for $k = 1, \ldots, 4$. Let us denote by $\vec{\bar{u}}^{\ell,k} = (\bar{u}^{\ell,k}_1, \ldots, \bar{u}^{\ell,k}_{n_t})^\top$ the obtained locally suboptimal control vector with corresponding step function $\bar{u}^{\ell,k}_\tau$. Then the time-dependent perturbation functions $\zeta^1, \ldots, \zeta^4$ are also step functions on $[0,T]$, and they are equivalent to vectors $(\zeta^k_1, \ldots, \zeta^k_{n_t})^\top, k = 1, \ldots, 4$, defined by

$$\zeta^k_i = \begin{cases} \left[\dfrac{1}{\tau}\displaystyle\int_{t_{i-1}}^{t_i}\int_\Omega (w_k)_\ell(x)\bar{p}_\ell(x,t)dxdt + \lambda\bar{u}^{\ell,k}_i\right]_-, & \text{if} \quad \bar{u}^{\ell,k}_i = -1, \\ -\left[\dfrac{1}{\tau}\displaystyle\int_{t_{i-1}}^{t_i}\int_\Omega (w_k)_\ell(x)\bar{p}_\ell(x,t)dxdt + \lambda\bar{u}^{\ell,k}_i\right], & \text{if} \quad -1 < \bar{u}^{\ell,k}_i < 1, \\ -\left[\dfrac{1}{\tau}\displaystyle\int_{t_{i-1}}^{t_i}\int_\Omega (w_k)_\ell(x)\bar{p}_\ell(x,t)dxdt + \lambda\bar{u}^{\ell,k}_i\right]_+, & \text{if} \quad \bar{u}^{\ell,k}_i = 1, \end{cases} \tag{4.46}$$

with

$$\zeta^k(t) = \zeta^k_i, \quad t_{i-1} \leq t < t_i,$$

for all $i \in \{1, \ldots, n_t\}, k = 1, \ldots, 4$. Finally, we use $\|\zeta\|^2_{\mathrm{L}^2(0,T;\mathbb{R}^4)} = \sum_{k=1}^4 \|\zeta^k\|^2_{\mathrm{L}^2(0,T)}$.

Remark 4.8. Analogous to Section 4.1, after discretization, the adjoint state $\bar{p}_\ell$ will correspond to a step function $\bar{p}^\ell_\tau$ in time. Then, the time integrals in equation (4.46) can be directly evaluated and expressed in terms of $\bar{p}^\ell_i(\cdot) = \bar{p}^\ell_\tau(\cdot,t), t \in [t_{i-1}, t_i), i = 1, \ldots, n_t$.

We computed a sufficiently precise approximation of the known optimal controls by step functions $\{\bar{u}^k_\tau\}^4_{k=1}$. Let $\bar{u}_\tau = (\bar{u}^1_\tau, \ldots, \bar{u}^4_\tau)^\top$. The comparison of the error estimators and the numerical errors between the POD-optimal control vector function $\bar{u}^\ell_\tau = (u^{\ell,1}_\tau, \ldots, u^{\ell,4}_\tau)^\top$ and $\bar{u}_\tau$ is illustrated in Table 4. In contrast to Example 1, this example does not fit that well for POD; compare Table 1. Notice that the success of the standard POD is not guaranteed and may depend on the particular type of equation.

The exact optimal solutions $\bar{u}_i$ must obey the necessary optimality conditions

$$\bar{u}_i(t) = \mathbb{P}_{[-1,1]}\left\{-\frac{1}{\lambda}\int_\Omega w_i(x)p(x,t)dx\right\} \quad , i = 1, \ldots, 4.$$

ℓ	$\|\bar{u}_\tau - \bar{u}_\tau^\ell\|_{\mathrm{L}^2(0,T;\mathbb{R}^4)}$	$\frac{\tau}{\sigma_{min}^\ell}\|\zeta\|_{\mathrm{L}^2(0,T;\mathbb{R}^4)}$	σ_{min}^ℓ
4	6.414e-3	2.467e-2	0.0015327
5	5.051e-3	1.791e-2	0.0015335
6	4.026e-3	1.001e-2	0.0015345
7	2.293e-3	3.832e-3	0.0015354
8	2.601e-3	5.714e-3	0.0015353
9	5.142e-3	1.562e-2	0.0015341
10	7.335e-3	2.646e-2	0.0015328
11	9.112e-3	3.141e-2	0.0015353
12	5.234e-3	1.567e-2	0.0015343

TABLE 4. Example 2: Numerical errors $\|\bar{u}_\tau - \bar{u}_\tau^\ell\|_{\mathrm{L}^2(0,T;\mathbb{R}^4)}$ and error estimators $\frac{\tau}{\sigma_{min}^\ell}\|\zeta\|_{\mathrm{L}^2(0,T;\mathbb{R}^4)}$ for different numbers ℓ with corresponding eigenvalues σ_{min}^ℓ.

In Figure 4 the precision of satisfying these conditions is visualized for the POD-optimal controls $\bar{u}_1^\ell$ and $\bar{u}_2^\ell$ using $\ell = 7$ POD basis functions. Here, the solid lines mark the POD controls whereas the dashed lines present the terms $-\int_\Omega w_i(x)p(x,t)dx/\lambda$. The corresponding results for the controls $\bar{u}_3^\ell$, $\bar{u}_4^\ell$ look just as well.

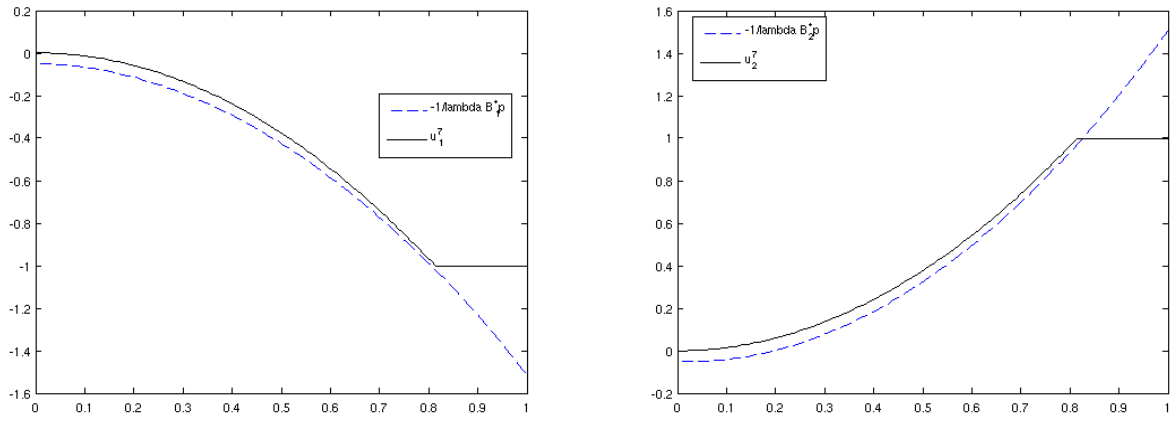

FIGURE 4. Example 2: Optimality test for the POD controls $\bar{u}_1^7, \bar{u}_2^7$.

The optimization of the full system needed more than 18 minutes, whereas the POD reduced optimal control problem ($\mathbf{PD.2}^{\boldsymbol{\ell}}$) with $\ell = 7$ was solved in about 4.5 seconds, a tremendous gain. The performance of the method is illustrated in Table 5. There is no significant difference between the computed optimal values.

	FEM	POD ($\ell = 7$)
CPU time for optimization	18 min 35 s	4.25 s
Minimal value of $\mathcal{J}$	0.031540	0.031539

TABLE 5. Example 2: CPU times of POD optimization and corresponding minimal values of the objective function in comparison with FE method.

We should remark that the computing times in the Tables 2 and 5 do not include the time for setting up the reduced Hessian and for computing its smallest eigenvalue. Again, the computation of the reduced Hessian matrix by the null space method was a time consuming task. Here it took us about 7.5 minutes. This shows

that, to speed up the process, the smallest eigenvalue of the reduced Hessian should be computed by a method that avoids the detour via the null space.

References

[1] A. C. Antoulas. *Approximation of large-scale dynamical systems*, volume 6 of *Advances in Design and Control*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2005. With a foreword by Jan C. Willems.

[2] N. Arada, E. Casas, and F. Tröltzsch. Error estimates for the numerical approximation of a semilinear elliptic control problem. *Computational Optimization and Applications*, 23:201–229, 2002.

[3] M. Barrault, Y. Maday, N. C. Nguyen, and A. T. Patera. An 'empirical interpolation' method: application to efficient reduced-basis discretization of partial differential equations. *C. R. Math. Acad. Sci. Paris*, 339(9):667–672, 2004.

[4] P. Benner and E.S. Quintana-Ortí. Model reduction based on spectral projection methods. In Reduction of Large-Scale Systems, P. Benner, V. Mehrmann, D. C. Sorensen (eds.), Lecture Notes in Computational Science and Engineering, 45, 5-48, 2005.

[5] E. Casas, J. C. De los Reyes, and F. Tröltzsch. Sufficient second-order optimality conditions for semilinear control problems with pointwise state constraints. *SIAM J. on Optimization*, 19:616–643, 2008.

[6] E. Casas and F. Tröltzsch. First- and second-order optimality conditions for a class of optimal control problems with quasilinear elliptic equations. *SIAM J. Control and Optimization*, 48:688–718, 2009.

[7] S. Chaturantabut and D. C. Sorensen. Nonlinear model reduction via discrete empirical interpolation. *SIAM J. Sci. Comput.*, 32(5):2737–2764, 2010.

[8] A. L. Dontchev, W. W. Hager, A. B. Poore, and B. Yang. Optimality, stability, and convergence in nonlinear control. *Applied Math. and Optimization*, 31:297–326, 1995.

[9] M.A. Grepl and M. Kärcher. Reduced basis a posteriori error bounds for parametrized linear-quadratic elliptic optimal control problems. *C. R. Acad. Sci. Paris, Ser. I*, 349:873-877, 2011.

[10] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE Constraints*, volume 23. Springer-Verlag, Berlin, 2009.

[11] M. Hinze and S. Volkwein. Error estimates for abstract linear-quadratic optimal control problems using proper orthogonal decomposition. *Computat. Optim. and Appl.*, 39:319-345, 2008.

[12] P. Holmes, J. L. Lumley, and G. Berkooz. *Turbulence, coherent structures, dynamical systems and symmetry*. Cambridge Monographs on Mechanics. Cambridge University Press, Cambridge, 1996.

[13] A. D. Ioffe. Necessary and sufficient conditions for a local minimum 3: Second order conditions and augmented duality. *SIAM J. Control and Optimization*, 17:266–288, 1979.

[14] K. Ito and K. Kunisch. *Lagrange Multiplier Approach to Variational Problems and Applications*. SIAM, Philadelphia, 2008.

[15] M. Kahlbacher and S. Volkwein. POD a-posteriori error based inexact SQP method for bilinear elliptic optimal control problems. *ESAIM: Mathematical Modelling and Numerical Analysis*, 46:491-511, 2012.

[16] K. Kunisch and S. Volkwein. Galerkin proper orthogonal decomposition methods for parabolic problems. *Numerische Mathematik*, 90:117–148, 2001.

[17] O. Lass and S. Volkwein. POD Galerkin schemes for nonlinear elliptic-parabolic systems. Submitted, 2011.

[18] J. L. Lions. *Optimal Control of Systems Governed by Partial Differential Equations*. Springer-Verlag, Berlin, 1971.

[19] K. Malanowski. Convergence of approximations vs. regularity of solutions for convex, control–constrained optimal control problems. *Applied Math. Optimization*, 8:69–95, 1981.

[20] K. Malanowski. Two-norm approach in stability and sensitivity analysis of optimization and optimal control problems. *Advances in Mathematical Sciences and Applications*, 2:397–443, 1993.

[21] K. Malanowski, Ch. Büskens, and H. Maurer. Convergence of approximations to nonlinear optimal control problems. In *Mathematical Programming with Data Perturbations*, volume 195 of *Lecture Notes in Pure and Applied Mathematics*, pages 253–284, New York, 1997. Marcel-Dekker, Inc.

[22] I. Neitzel and B. Vexler. A priori error estimates for space-time finite element discretization of semilinear parabolic optimal control problems. *Numerische Mathematik*, DOI: 10.1007/s00211-011-0409-9, published online 2011.

[23] A. Rösch and D. Wachsmuth. Numerical verification of optimality conditions. *SIAM J. Control Optimization*, 47(5):2557–2581, 2008.

[24] A. Rösch and D. Wachsmuth. How to check numerically the sufficient optimality conditions for infinite-dimensional optimization problems. In *Optimal control of coupled systems of partial differential equations*, volume 158 of *Internat. Ser. Numer. Math.*, pages 297–317. Birkhäuser Verlag, Basel, 2009.

[25] E. W. Sachs and M. Schu. A priori error estimates for reduced order models in finance. Submitted 2011.

[26] K. Schittkowski. Numerical solution of a time–optimal parabolic boundary-value control problem. *JOTA*, 27:271–290, 1979.

[27] A. Studinger and S. Volkwein. *Numerical analysis of POD a-posteriori error estimation for optimal control*. To appear 2012.

[28] T. Tonn, K. Urban, and S. Volkwein. Comparison of the reduced-basis and POD a-posteriori error estimators for an elliptic linear quadratic optimal control problem. *Mathematical and Computer Modelling of Dynamical Systems*, *Mathematical and Computer Modelling of Dynamical Systems*, 17:355-369, 2011.

[29] F. Tröltzsch. *Optimal Control of Partial Differential Equations. Theory, Methods and Applications*, volume 112. American Math. Society, Providence, 2010.
[30] F. Tröltzsch and S. Volkwein. POD a-posteriori error estimates for linear-quadratic optimal control problems. *Computational Optimization and Applications*, 44:83–115, 2009.
[31] S. Volkwein. Optimal control of a phase-field model using proper orthogonal decomposition. *ZAMM Z. Angew. Math. Mech.*, 81(2):83–97, 2001.
[32] S. Volkwein. *Model Reduction using Proper Orthogonal Decomposition.* Lecture notes, Institute of Mathematics and Statistics, University of Konstanz, 2011.