Universität Konstanz

POD a-posteriori error analysis for optimal control problems with mixed control-state constraints

Martin Gubisch
Stefan Volkwein

Konstanzer Schriften in Mathematik

Nr. 314, Juni 2013

ISSN 1430-3558

Noname manuscript No.
(will be inserted by the editor)

POD a-posteriori error analysis for optimal control problems with mixed control-state constraints

Martin Gubisch · Stefan Volkwein

Received: April 21, 2013/ Accepted: date

Abstract In this work linear-quadratic optimal control problems for parabolic equations with mixed control-state constraints are considered. These problems arise when a Lavrentiev regularization is utilized for state constrained linear-quadratic optimal control problems. For the numerical solution a Galerkin discretization is applied utilizing proper orthogonal decomposition (POD). Based on a perturbation method it is determined how far the suboptimal control, computed on the basis of the POD method, is from the (unknown) exact one. Numerical examples illustrate the theoretical results. In particular, the POD Galerkin scheme is applied to a problem with state constraints.

1 Introduction

In this paper we consider a certain class of linear-quadratic optimal control problems governed by a linear evolution problem and mixed control-state constraints. Due to the following reasons mixed control-state constraints are of

The authors gratefully acknowledge support by the German Science Fund DFG grant VO 1658/2-1 *A-posteriori-POD Error Estimators for Nonlinear Optimal Control Problems governed by Partial Differential Equations*. The first author is further supported by the Landesgraduiertenförderung of Baden-Wurttemberg.

Martin Gubisch
University of Konstanz
Department of Mathematics and Statistics
Universitätsstraße 10
D-78457 Konstanz, Germany
E-mail: Martin.Gubisch@uni-konstanz.de

Stefan Volkwein
University of Konstanz
Department of Mathematics and Statistics
Universitätsstraße 10
D-78457 Konstanz, Germany
E-mail: Stefan.Volkwein@uni-konstanz.de

interest: (1) They arise in Lavrentiev-type regularizations of state constrained problems and (2) they may appear in their own rights (e.g., if the control is resticted by a multiple of the state). For the numerical solution we apply a Galerkin approximation, which is based on proper orthogonal decomposition (POD). Recall that POD is a method for deriving reduced-order models of dynamical systems; see [11], for instance. In order to ensure that the POD suboptimal solutions are sufficiently accurate, we derive an a-posteriori error estimate for the difference between the exact (unknown) optimal control and its POD suboptimal approximations. Moreover, it is shown that this error tends to zero if the number of POD basis functions in the Galerkin ansatz is increased. The proof relies on a perturbation argument [4] and a convergence analysis for the POD Galerkin scheme, where we make use of a modified POD approximation [6]. Although we can transform the optimal control problem with mixed control-state constraints into a purely control constrained optimal control problem, we can not directly apply the results from [21], because the transformation itself depends on the POD discretization as well. Furthermore, we propose a new POD Galerkin ansatz for state and adjoint equations which avoids discretization errors coming from approxiamtions of the initial values. In the numerical examples we combine the a-posteriori error estimator with an adaptive basis update strategy; see [1]. Of course, other strategies can be applied as well; see [2], for instance. Although linear-quadratic optimal control problems with mixed-control constraints can be cast into linear-quadratic optimal control problems with purely control constraints, we can not simply apply the results from [21]. This is due to the fact, that the transformation itself depends on the POD discretization as well. Let us mention that the a-posteriori analysis can also be utilized for nonlinear problems in an inexact sequential quadratic programming (SQP) approach, where in each level of the SQP iteration a linear-quadratic optimal control problem has to be solved. For instance, this is done in [12] utilizing the a-posteriori analysis from [21].

The paper is organized as follows: In Section 2 we introduce our linear-quadratic optimal control problems and review first-order optimality conditions. The a-posteriori error analysis is carried out in Section 3. Section 4 is devoted to the POD approximation and the POD convergence analysis. Finally, numerical test examples are studied in Section 5.

2 The optimal control problem

In this section we introduce a class of linear-quadratic optimal control problems. We recall the associated first-order optimality conditions and formulate the optimization problem as a reduced problem for the control variable only. For the solution of the optimal control problem we apply a primal-dual active set strategy which is equivalent to a semismooth Newton method [7]. To utilize the error analysis presented in [10, 21] we transform the reduced problem to an optimal control problem which is governed by bilateral control constraints for a transformed control variable.

2.1 Problem formulation

Let V and H be real, separable Hilbert spaces and suppose that V is dense in H with compact embedding. In particular, there exists a constant $C_V > 0$ such that

$$\|\varphi\|_H \leq C_V \|\varphi\|_V \quad \forall \varphi \in V. \tag{2.1}$$

By $\langle \cdot\,,\cdot\rangle_H$ and $\langle \cdot\,,\cdot\rangle_V$ we denote the inner products in H and V, respectively. Let $T > 0$ be the fixed final time. For $t \in [0,T]$ we define a time-dependent symmetric bilinear form $a(t;\cdot\,,\cdot) : V \times V \to \mathbb{R}$ satisfying

$$\begin{aligned}
\big|a(t;\varphi,\psi)\big| &\leq \alpha \|\varphi\|_V \|\psi\|_V && \forall \varphi \in V \text{ a.e. in } [0,T], && \text{(2.2a)}\\
a(t;\varphi,\varphi) &\geq \alpha_1 \|\varphi\|_V^2 - \alpha_2 \|\varphi\|_H^2 && \forall \varphi \in V \text{ a.e. in } [0,T] && \text{(2.2b)}
\end{aligned}$$

for constants α, $\alpha_1 > 0$ and $\alpha_2 \geq 0$ which do not depend on t. In (2.2), the abbreviation "a.e." stands for "almost everywhere". By identifying H with its dual H' it follows that

$$V \hookrightarrow H = H' \hookrightarrow V',$$

each embedding being continuous and dense. Recall that the space $W(0,T)$

$$W(0,T) = \big\{\varphi \in L^2(0,T;V)\,\big|\,\varphi_t \in L^2(0,T;V')\big\}$$

is a Hilbert space endowed with the common inner product [3, pp. 472-479]. The control space is given by $U = L^2(0,T;\mathbb{R}^m)$ with $m \in \mathbb{N}$. In particular, we identify U with its dual space U'. For $u \in U$, $y_\circ \in H$ and $f \in L^2(0,T;V')$ we consider the linear evolution problem

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\langle y(t),\varphi\rangle_H + a(t;y(t),\varphi) &= \langle (f+\mathcal{B}u)(t),\varphi\rangle_{V',V}\ \forall \varphi \in V \text{ a.e. in } [0,T],\\
y(0) &= y_\circ \qquad\qquad\qquad\qquad \text{in } H,
\end{aligned} \tag{2.3}$$

where $\langle \cdot\,,\cdot\rangle_{V',V}$ stands for the dual pairing between V and its dual space V' and $\mathcal{B} : U \to L^2(0,T;V')$ is a continuous, linear operator.

It is known that for every $f \in L^2(0,T;V')$, $u \in U$ and $y_\circ \in H$ there is a unique weak solution $y \in W(0,T)$ satisfying (2.3) and

$$\|y\|_{W(0,T)} \leq C\Big(\|y_\circ\|_H + \|f\|_{L^2(0,T;V')} + \|u\|_U\Big) \tag{2.4}$$

for a constant $C > 0$ which is independent of $y_\circ$, f and u. For a proof of the existence of a unique solution we refer to [3, pp. 512-520]. The a-priori error estimate follows from standard variational techniques and energy estimates. If $f+\mathcal{B}u \in L^2(0,T;H)$, $a(t;\cdot\,,\cdot) = a(\cdot\,,\cdot)$ (independent of t) and $y_\circ \in V$ hold, we have $y \in L^\infty(0,T;V) \cap H^1(0,T;H)$; see [3, pp. 532-533] and [5, pp. 360-364].

Remark 2.1 We split the solution to (2.3) in one part, which depends on the fixed initial condition $y_\circ$ and right-hand side f, and another part depending linearly on the control variable. Let $\hat{y} \in W(0,T)$ be the unique solution to the problem

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\langle \hat{y}(t),\varphi\rangle_H + a(t;\hat{y}(t),\varphi) &= \langle f(t),\varphi\rangle_{V',V} && \forall\varphi\in V \text{ a.e. in } [0,T],\\ \hat{y}(0) &= y_\circ && \text{in } H. \end{aligned}$$

We define the subspace

$$W_0(0,T) = \big\{\varphi\in W(0,T)\,\big|\,\varphi(0)=0 \text{ in } H\big\}$$

endowed with the topology of $W(0,T)$. Let us now introduce the linear solution operator $\mathcal{S}: U\to W_0(0,T)$: for $u\in U$ the function $y=\mathcal{S}u\in W_0(0,T)$ is the unique solution to

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle y(t),\varphi\rangle_H + a(t;y(t),\varphi) = \langle (\mathcal{B}u)(t),\varphi\rangle_{V',V} \quad \forall\varphi\in V \text{ a.e. in } [0,T].$$

From $y\in W_0(0,T)$ it follows that $y(0)=0$ in H. The boundedness of $\mathcal{S}$ follows from (2.4). Now, the solution to (2.3) can be expressed as $y=\hat{y}+\mathcal{S}u$. $\Diamond$

We introduce the Hilbert space

$$X = W(0,T)\times U$$

endowed with the natural product topology. Let $\mathcal{I}: L^2(0,T;V)\to U$ be a bounded linear operator. By X_{ad} we denote the closed, convex and bounded subset

$$X_{\mathsf{ad}} = \big\{(\hat{y}+\mathcal{S}u,u)\in X\,\big|\,u_a\le \varepsilon u+\mathcal{I}(\hat{y}+\mathcal{S}u)\le u_b \text{ in } \mathbb{R}^m \text{ a.e. in } [0,T]\big\},$$

where $u_a,u_b\in U$ satisfy $u_a\le u_b$ componentwise in $\mathbb{R}^m$ a.e. and $\varepsilon>0$ holds. The cost function $J: X\to\mathbb{R}$ is given by

$$J(x) = \frac{\sigma_Q}{2}\int_0^T \|y(t)-y_Q(t)\|_H^2\,\mathrm{d}t + \frac{\sigma_\Omega}{2}\|y(T)-y_\Omega\|_H^2 + \frac{\sigma}{2}\|u\|_U^2 \tag{2.5}$$

for $x=(y,u)\in X$, where $(y_Q,y_\Omega)\in L^2(0,T;H)\times H$ are desired states. Furthermore, $\sigma_Q,\sigma_\Omega\ge 0$ and $\sigma>0$. The optimal control problem is given by

$$\min J(x) \quad \text{subject to (s.t.)} \quad x\in X_{\mathsf{ad}}. \tag{$\mathbf{P}$}$$

Applying standard arguments [15] one can prove that there exists a unique optimal solution $\bar{x}=(\bar{y},\bar{u})$ to (**P**). The uniqueness follows from the strict convexity properties of the objective functional on X_{ad}. Throughout this paper, a bar indicates optimality. Next we formulate the first-order sufficient optimality conditions of (**P**) (see [20], for instance):

Theorem 2.1 *Suppose that the feasible set X_{ad} is nonempty and that $(\bar y, \bar u) \in X_{\mathsf{ad}}$ is the solution to* (**P**). *Then there exists a unique Lagrange multiplier pair $(\bar p, \bar\lambda) \in X$ satisfying together with $(\bar y, \bar u)$ the primal-dual system*

$$\bar y = \hat y + \mathcal{S}\bar u, \tag{2.6a}$$

$$\begin{aligned}
-\frac{\mathrm{d}}{\mathrm{d}t}\langle \bar p(t), \varphi\rangle_H + a(t;\bar p(t),\varphi) + \langle (\mathcal{I}^\star\bar\lambda)(t),\varphi\rangle_{V',V} & \\
+\sigma_Q\,\langle \bar y(t) - y_Q(t), \varphi\rangle_H = 0 & \quad \forall \varphi \in V \text{ a.e.}, \quad (2.6\text{b})\\
\bar p(T) + \sigma_\Omega\big(\bar y(T) - y_\Omega\big) = 0 & \quad \text{in } H,\\
\sigma \bar u(t) - (\mathcal{B}^\star \bar p)(t) + \varepsilon\bar\lambda(t) = 0 & \quad \text{in } \mathbb{R}^m \text{ a.e.}, \quad (2.6\text{c})
\end{aligned}$$

$$\begin{aligned}
\bar\lambda(t) = \max\big(0, \bar\lambda(t) + \gamma((\mathcal{I}\bar y)(t) + \varepsilon\bar u(t) - u_b(t))\big) & \quad (2.6\text{d})\\
+ \min\big(0, \bar\lambda(t) + \gamma((\mathcal{I}\bar y)(t) + \varepsilon\bar u(t) - u_a(t))\big) & \quad \text{in } \mathbb{R}^m \text{ a.e.},
\end{aligned}$$

where $\mathcal{B}^\star : L^2(0,T;V) \to U$ and $\mathcal{I}^\star : U \to L^2(0,T;V')$ denote the adjoints of $\mathcal{B}$ and $\mathcal{I}$, respectively. Furthermore, $\gamma \neq 0$ is an arbitrary real number. In (2.6d) *the max- and min-operations are interpreted componentwise in the pointwise everywhere sense.*

Remark 2.2 Analogous to Remark 2.1 we split the adjoint variable into one part depending on the fixed desired states and into two other parts, which depend linearly on the control variable and on the multiplier λ. Recall that we have defined $\hat y$ as well as the operator $\mathcal{S}$ in Remark 2.1. For given $y_Q \in L^2(0,T;H)$ and $y_\Omega \in H$ let $\hat p = \hat p \in W(0,T)$ denote the unique solution to the adjoint equation

$$\begin{aligned}
-\frac{\mathrm{d}}{\mathrm{d}t}\langle \hat p(t), \varphi\rangle_H + a(t;\hat p(t),\varphi) = \sigma_Q\,\langle (y_Q - \hat y)(t), \varphi\rangle_H & \quad \forall \varphi \in V \text{ a.e.},\\
\hat p(T) = \sigma_\Omega\big(y_\Omega - \hat y(T)\big) & \quad \text{in } H.
\end{aligned}$$

Further, we define the linear, bounded operators $\mathcal{A}_1$, $\mathcal{A}_2 : U \to W(0,T)$ as follows: for given $u \in U$ the function $p = \mathcal{A}_1 u$ is the unique solution to

$$\begin{aligned}
-\frac{\mathrm{d}}{\mathrm{d}t}\langle p(t), \varphi\rangle_H + a(t;p(t),\varphi) = -\sigma_Q\,\langle (\mathcal{S}u)(t), \varphi\rangle_H & \quad \forall \varphi \in V \text{ a.e.},\\
p(T) = -\sigma_\Omega(\mathcal{S}u)(T) & \quad \text{in } H
\end{aligned}$$

and for given $\lambda \in U$ the function $p = \mathcal{A}_2\lambda$ uniquely solves

$$\begin{aligned}
-\frac{\mathrm{d}}{\mathrm{d}t}\langle p(t), \varphi\rangle_H + a(t;p(t),\varphi) + \langle (\mathcal{I}^\star\lambda)(t),\varphi\rangle_{V',V} = 0 & \quad \forall \varphi \in V \text{ a.e.},\\
p(T) = 0 & \quad \text{in } H.
\end{aligned}$$

Then, the solution to (2.6b) can be expressed by $\bar p = \hat p + \mathcal{A}_1\bar u + \mathcal{A}_2\bar\lambda$. $\Diamond$

2.2 The reduced problem

In Remark 2.1 we have introduced the solution operator $\mathcal{S}$. The reduced cost functional $\hat{J} : U \to \mathbb{R}$ is defined as

$$\hat{J}(u) = J(\hat{y} + \mathcal{S}u, u), \quad u \in U.$$

We define the set of admissible controls by

$$U_{\mathsf{ad}} = \big\{ u \in U \,\big|\, u_a \le \varepsilon u + \mathcal{I}\big(\hat{y} + \mathcal{S}u\big) \le u_b \text{ in } \mathbb{R}^m \text{ a.e. in } [0,T] \big\},$$

which is convex, closed and bounded in U; see [8, Prop. 2.2]. Now we consider the reduced optimal control problem:

$$\min \hat{J}(u) \quad \text{s.t.} \quad u \in U_{\mathsf{ad}}. \tag{$\hat{\mathbf{P}}$}$$

Clearly, if $\bar{u}$ is the optimal solution to ($\hat{\mathbf{P}}$), then $\bar{x} = (\hat{y} + \mathcal{S}\bar{u}, \bar{u})$ is the optimal solution to ($\mathbf{P}$). On the other hand, if $\bar{x} = (\bar{y}, \bar{u})$ is the solution to ($\mathbf{P}$), then $\bar{u}$ solves ($\hat{\mathbf{P}}$). A first-order sufficient optimality condition for the convex linear-quadratic problem ($\hat{\mathbf{P}}$) is given by the variational inequality

$$\langle \nabla \hat{J}(\bar{u}), u - \bar{u} \rangle_U \ge 0 \quad \text{for all } u \in U_{\mathsf{ad}}. \tag{2.7}$$

Combining (2.6c) with (2.6d) and choosing $\gamma = \sigma\varepsilon^{-2}$ to prevent the dependency of the min- and max-terms on u, this is equivalent with

$$u = \big(\mathcal{B}^\star p - \varepsilon\mathcal{N}(y,p)\big)/\sigma \quad \text{in } U, \tag{2.8}$$

where (y,p) solves the coupled nonlinear primal-dual system

$$\begin{aligned} y &= \hat{y} + \frac{1}{\sigma}\, \mathcal{S}\big(\mathcal{B}^\star p - \varepsilon\mathcal{N}(y,p)\big), \\ p &= \hat{p} + \frac{1}{\sigma}\, \mathcal{A}_1\big(\mathcal{B}^\star p - \varepsilon\mathcal{N}(y,p)\big) + \mathcal{A}_2\mathcal{N}(y,p), \end{aligned} \tag{2.9}$$

and the nonlinearity $\mathcal{N} : W(0,T) \times W(0,T) \to U$,

$$\begin{aligned} \mathcal{N}(y,p) = {} & \max(0, \varepsilon^{-1}\mathcal{B}^\star p + \varepsilon^{-2}\sigma(\mathcal{I}y - u_b)) \\ & + \min(0, \varepsilon^{-1}\mathcal{B}^\star p + \varepsilon^{-2}\sigma(\mathcal{I}y - u_a)), \end{aligned}$$

coincides with the Lagrange multiplier λ.

2.3 The primal-dual active set method

To solve ($\mathbf{P}$) numerically, a primal-dual active set strategy is utilized. This method is equivalent to a locally superlinearly convergent semi-smooth Newton algorithm applied to the first-order necessary optimality conditions; see [7,8]. For given iterates $(y^k, p^k) \in W(0,T) \times W(0,T)$, $k \ge 0$, and for $i = 1, \ldots, m$ we introduce the active and inactive sets

$$\begin{aligned} \mathcal{A}_{ai}^k &= \Big\{ t \in [0,T] \,\Big|\, \big(\varepsilon\sigma^{-1}\mathcal{B}^\star p^k + \mathcal{I}y^k\big)_i(t) < u_{ai}(t) \Big\}, \ \mathcal{A}_i^k = \mathcal{A}_{ai}^k \cup \mathcal{A}_{bi}^k, \\ \mathcal{A}_{bi}^k &= \Big\{ t \in [0,T] \,\Big|\, \big(\varepsilon\sigma^{-1}\mathcal{B}^\star p^k + \mathcal{I}y^k\big)_i(t) > u_{bi}(t) \Big\}, \ \ \mathcal{I}_i^k = [0,T] \backslash \mathcal{A}_i^k. \end{aligned} \tag{2.10}$$

In Algorithm 1 we formulate the semismooth Newton method for our problem.

Algorithm 1 (Primal-dual active set strategy)

Require: Starting value (y^0, p^0) and maximal iteration number k_{max}.
Set $k = 0$, determine the active and inactive sets according to (2.10).
repeat
  Compute the solution (y^{k+1}, p^{k+1}) to (2.9) and set $k = k + 1$.
  Compute the active and inactive sets according to (2.10).
until $(\mathcal{A}_{ai}^k = \mathcal{A}_{ai}^{k-1}$ **and** $\mathcal{A}_{bi}^k = \mathcal{A}_{bi}^{k-1})$ **or** $k = k_{\mathsf{max}}$.
Set $u = (\mathcal{B}^\star p^k - \varepsilon\lambda^k)/\sigma$.

2.4 Restatement of (P) as a control constrained problem

It is known [16] that (**P**) can be cast into a purely control constrained optimal control problem. Recall that $\mathcal{I}$ and $\mathcal{S}$ are linear and continuous operators, i.e. the operator $\mathcal{F} = \varepsilon + \mathcal{I}\mathcal{S} : U \to U$ is linear and bounded, too. The proof of Lemma 2.1 can be found in the Appendix.

Lemma 2.1 *Suppose that for every $\epsilon > 0$ there is a constant $C_\epsilon > 0$ satisfying*

$$\|(\mathcal{B}\mathcal{I}\varphi)(t)\|_{V'} \leq C_\epsilon \|\varphi(t)\|_H + \epsilon \|\varphi(t)\|_V \tag{2.11}$$

for all $\varphi \in W(0, T)$ a.e. in $[0, T]$. Then, the linear operator $\mathcal{F} = \varepsilon + \mathcal{I}\mathcal{S}$ has a bounded inverse.

Next we introduce the transformed control variable $v = \mathcal{F}u \in U$ for $u \in U$. From Remark 2.1 we conclude that the solution y to (2.3) satisfies $\mathcal{S}u = y - \hat{y}$. Consequently, v has the representation $v = \varepsilon u + \mathcal{I}(y - \hat{y})$. Using $(y, u) \in X_{\mathsf{ad}}$, it follows that v satisfies the bilateral box constraints

$$v_a \leq v \leq v_b \quad \text{in } [0, T] \text{ a.e.,} \tag{2.12}$$

where we set $v_a = u_a - \mathcal{I}\hat{y}$ and $v_b = u_b - \mathcal{I}\hat{y}$. Let

$$V_{\mathsf{ad}} = \big\{ v \in U \,\big|\, v_a \leq v \leq v_b \text{ a.e. in } [0, T] \big\}.$$

Let us assume that $\mathcal{F}$ has a bounded inverse. We replace the state and control variable in J as follows:

$$\tilde{J}(v) = \hat{J}(\mathcal{F}^{-1}v) = J(\hat{y} + \mathcal{S}\mathcal{F}^{-1}v, \mathcal{F}^{-1}v) \quad \text{for } v \in U.$$

Then, we consider the optimal control problem

$$\min \tilde{J}(v) \quad \text{s.t.} \quad v \in V_{\mathsf{ad}}. \tag{$\tilde{\mathbf{P}}$}$$

If $\bar{v}$ is the solution to ($\tilde{\mathbf{P}}$), then the pair $\bar{x} = (\bar{y}, \bar{u})$ with $\bar{y} = \hat{y} + \mathcal{S}\mathcal{F}^{-1}\bar{v}$ and $\bar{u} = \mathcal{F}^{-1}\bar{v}$ is the solution to (**P**).

Remark 2.3 Let $u \in U$, then $v = \mathcal{F}u$ is given by computing $y = \mathcal{S}u$ and choosing $v = \varepsilon u + \mathcal{I}y$. On the other hand, suppose that we know $v \in U$, then $u = \mathcal{F}^{-1}v$ can be calculated by solving

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\left\langle y(t),\varphi\right\rangle_H + a(t;y(t),\varphi)\\
&\qquad\qquad +\frac{1}{\varepsilon}\left\langle (\mathcal{B}\mathcal{I}y)(t),\varphi\right\rangle_{V',V} = \frac{1}{\varepsilon}\langle \mathcal{B}v(t),\varphi\rangle_{V',V} \quad \forall\varphi\in V \text{ a.e.},\\
&\qquad\qquad\qquad\qquad y(0) = 0 \qquad\qquad\qquad \text{in } H
\end{aligned}$$

and choosing $u = (v - \mathcal{I}y)/\varepsilon$. Hence, the transformations $v \mapsto u$ and $u \mapsto v$ both require to solve a state equation. ◊

2.5 First-order sufficient optimality conditions for ($\tilde{\mathbf{P}}$)

In this subsection we present the first-order optimality conditions for ($\tilde{\mathbf{P}}$). We suppose that $\mathcal{F}$ has a bounded inverse. Assume that $\bar{v}$ denotes the unique optimal control for ($\tilde{\mathbf{P}}$). Using a Lagrangian framework [20] the optimal control $\bar{v}$ satisfies together with the corresponding state variable $\bar{y}$ and Lagrange multiplier $\bar{q}$ the following first-order optimality conditions for ($\tilde{\mathbf{P}}$) :

$$\bar{y} = \hat{y} + \mathcal{S}\mathcal{F}^{-1}\bar{v}, \tag{2.13a}$$

$$\bar{q} = \hat{p} + \mathcal{A}_1\mathcal{F}^{-1}\bar{v} \tag{2.13b}$$

$$\langle \sigma\mathcal{F}^{-\star}\mathcal{F}^{-1}\bar{v} - \mathcal{F}^{-\star}\mathcal{B}^\star\bar{q}, v-\bar{v}\rangle_U \geq 0 \quad \forall v\in V_{\mathsf{ad}}, \tag{2.13c}$$

where $\mathcal{F}^{-\star} : U \to U$ stands for the dual operator of $\mathcal{F}^{-1}$.

The operator $\mathcal{A}_1$ can be expressed in terms of the adjoint $\mathcal{S}^\star : W_0(0,T)' \to L^2(0,T;V)$. This follows from an adaption of Lemma 4.1 in [10] and Lemma 2.4 in [21]:

Lemma 2.2 *Define the linear and bounded operator $\Theta : W(0,T) \to W_0(0,T)'$*

$$\langle \Theta y,\varphi\rangle_{W(0,T)',W(0,T)} = \sigma_Q\int_0^T \langle y(t),\varphi(t)\rangle_H\,\mathrm{d}t + \sigma_\Omega\,\langle y(T),\varphi(T)\rangle_H$$

for $y \in W(0,T)$ and $\varphi \in W_0(0,T)$. Then, we have $\mathcal{B}^\star\mathcal{A}_1 = -\mathcal{S}^\star\Theta\mathcal{S}$.

2.6 A distributed optimal control problem

In this subsection we introduce an example for ($\mathbf{P}$) and discuss the presented theory for this application. Let $\Omega \subset \mathbb{R}^d$, $d \in \{1,2,3\}$, be an open and bounded domain with Lipschitz-continuous boundary $\Gamma = \partial\Omega$. For $T > 0$ we set $Q = (0,T)\times\Omega$ and $\Sigma = (0,T)\times\Gamma$. We choose $H = L^2(\Omega)$ and $V = H_0^1(\Omega)$ endowed with the usual inner products

$$\langle\varphi,\psi\rangle_H = \int_\Omega \varphi\psi\,\mathrm{d}\boldsymbol{x}, \quad \langle\varphi,\psi\rangle_V = \int_\Omega \varphi\psi + \nabla\varphi\cdot\nabla\psi\,\mathrm{d}\boldsymbol{x}$$

and their induced norms, respectively. In (2.1) we have $C_V = 1$. Let $\chi_i \in H$, $1 \le i \le m$, denote given control shape functions. Then, for given control $u \in U$, initial condition $y_\circ \in H$ and inhomogeneity $f \in L^2(0,T;H)$ we consider the linear heat equation

$$\begin{aligned}
y_t(t,\boldsymbol{x}) - \Delta y(t,\boldsymbol{x}) &= f(t,\boldsymbol{x}) + \sum_{i=1}^m u_i(t)\chi_i(\boldsymbol{x}), && \text{a.e. in } Q,\\
y(t,\boldsymbol{x}) &= 0, && \text{a.e. in } \Sigma,\\
y(0,\boldsymbol{x}) &= y_\circ(\boldsymbol{x}), && \text{a.e. in } \Omega.
\end{aligned} \tag{2.14}$$

We introduce the time-independent, symmetric bilinear form

$$a(\varphi,\psi) = \int_\Omega \nabla\varphi\cdot\nabla\psi \,\mathrm{d}\boldsymbol{x} \quad \text{for } \varphi,\psi \in V$$

and the bounded, linear operator $\mathcal{B} : U \to L^2(0,T;H) \hookrightarrow L^2(0,T;V')$ as

$$(\mathcal{B}u)(t,\boldsymbol{x}) = \sum_{i=1}^m u_i(t)\chi_i(\boldsymbol{x}) \quad \text{for } (t,\boldsymbol{x}) \in Q \text{ a.e. and } u \in U.$$

Hence, we have $\alpha = \alpha_1 = \alpha_2 = 1$ in (2.2). It follows that the weak formulation of (2.14) can be expressed in the form (2.3). Moreover, the unique weak solution to (2.14) belongs to the space $L^\infty(0,T;V)$ provided $y_\circ \in V$ holds.

We choose certain shape functions $\pi_1,\dots,\pi_m \in H$ and introduce the operator $\mathcal{I} : L^2(0,T;V) \to U$ by

$$(\mathcal{I}\varphi)(t) = \begin{pmatrix} (\mathcal{I}_1\varphi)(t) \\ \vdots \\ (\mathcal{I}_m\varphi)(t) \end{pmatrix} \quad \text{with} \quad (\mathcal{I}_i\varphi)(t) = \int_\Omega \pi_i(\boldsymbol{x})\varphi(t,\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$$

for $\varphi \in L^2(0,T;V)$ a.e. in $[0,T]$. Then, the mixed control-state constraints have the form

$$u_{ai}(t) \le \varepsilon u_i(t) + \int_\Omega \pi_i(\boldsymbol{x})y(t,\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \le u_{bi}(t) \qquad \text{a.e. in } [0,T]$$

for $(y,u) \in X$ and $i \in 1,\dots,m$ with $u_{ai}, u_{bi} \in L^2(0,T)$. Notice that $\mathcal{I}$ is even a bounded operator from $L^2(0,T;H)$ to U.

The adjoint operators $\mathcal{B}^\star : L^2(0,T;V) \to U$ and $\mathcal{I}^\star : U \to L^2(0,T;V')$ have the explicit representations

$$(\mathcal{B}_i^\star p)(t) = \int_\Omega \chi_i(\boldsymbol{x})p(t,\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \;(1 \le i \le m), \quad (\mathcal{I}^\star\lambda)(t,\boldsymbol{x}) = \sum_{i=1}^m \lambda_i(t)\pi_i(\boldsymbol{x})$$

for $(t,\boldsymbol{x}) \in Q$ a.e., $p \in L^2(0,T;V)$ and $\lambda \in U$. In particular, if $\chi_i = \pi_i$ holds for $1 \le i \le m$, then we have $\mathcal{B}^\star = \mathcal{I}$ as well as $\mathcal{I} = \mathcal{B}^\star$.

Let us discuss condition (2.11). For $\varphi \in W(0,T)$ we have $(\mathcal{B}\mathcal{I}\varphi)(t) \in H$ for $t \in [0,T]$ a.e. Using (2.1) with $C_V = 1$ we obtain

$$\|(\mathcal{B}\mathcal{I}\varphi)(t)\|_{V'} \leq \sum_{i=1}^{m} \big|\langle \pi_i, \varphi(t)\rangle_H\big| \, \|\chi_i\|_H \leq c_1 \, \|\varphi(t)\|_H^{1/2} \|\varphi(t)\|_V^{1/2},$$

where we set $c_1 = \sum_{i=1}^{m} \|\pi_i\|_H \|\chi_i\|_H$. Now (2.11) follows directly from Young's inequality (A.1) with $\mathsf{a} = 2\|\varphi(t)\|_V^{1/2}$ and $\mathsf{b} = c_1 \|\varphi(t)\|_N^{1/2}/2$. In particular, we find $C_\epsilon = c_1^2/(8\epsilon)$ for every $\epsilon > 0$.

3 A-posteriori error analysis

The goal of this section is to derive an a-posteriori error estimate for $(\hat{\mathbf{P}})$. For that purpose we utilize the technique in [21, Section 3] for the control constrained problem $(\tilde{\mathbf{P}})$.

3.1 Derivation of the a-posteriori error estimate

Suppose that $\mathcal{F}$ has a bounded inverse and that u_p is an arbitrary control in U_{ad}. Our goal is to estimate the norm $\|\bar{u} - u_p\|_U$ without the knowledge of the optimal solution $\bar{u} = \mathcal{F}^{-1}\bar{v}$. We set $v_p = \mathcal{F}u_p$, i.e., $v_p = \varepsilon u_p + \mathcal{I}(y_p - \hat{y})$ with $y_p = \hat{y} + \mathcal{S}u_p$. If $u_p \neq \bar{u}$ holds, then $v_p \neq \bar{v}$. Thus, v_p does not satisfy the necessary (and by convexity sufficient) optimality condition (2.13c). However, there exists a function $\zeta \in U$ such that

$$\langle \sigma \mathcal{F}^{-\star}\mathcal{F}^{-1} v_p - \mathcal{F}^{-\star}\mathcal{B}^\star q_p + \zeta, v - v_p \rangle_U \geq 0 \qquad \forall v \in V_{\mathsf{ad}}, \tag{3.1}$$

with $q_p = \hat{q} + \mathcal{A}u_p$. Therefore, v_p satisfies the optimality condition of a perturbed parabolic optimal control problem with "perturbation" ζ. The smaller ζ is, the closer v_p is to $\bar{v}$.

Next we estimate $\|\bar{u} - u_p\|_U$ in terms of $\|\zeta\|_U$. By Lemma 2.2 we have

$$\mathcal{B}^\star\big(\bar{q} - \bar{q}_p\big) = \mathcal{B}^\star\mathcal{A}_1^\star\big(\bar{u} - \bar{u}_p\big) = \mathcal{B}^\star\mathcal{S}^\star\Theta\mathcal{S}\big(u_p - \bar{u}\big) = \mathcal{B}^\star\mathcal{S}^\star\Theta\big(y_p - \bar{y}\big), \tag{3.2}$$

with $\bar{y} = \hat{y} + \mathcal{S}\bar{u}$. Choosing $v = v_p$ in (2.13c), $v = \bar{v}$ in (3.1) and using (3.2) we obtain

$$\begin{aligned}
0 &\leq \langle -\sigma\mathcal{F}^{-\star}(\bar{u} - u_p) + \mathcal{F}^{-\star}\mathcal{B}^\star(\bar{q} - q_p) + \zeta, \mathcal{F}(\bar{u} - u_p)\rangle_U \\
&= -\sigma \,\|\bar{u} - u_p\|_U^2 - \langle \mathcal{B}^\star\mathcal{S}^\star\Theta(\bar{y} - y_p), \bar{u} - u_p\rangle_U + \langle \zeta, \mathcal{F}(\bar{u} - u_p)\rangle_U \\
&= -\sigma \,\|\bar{u} - u_p\|_U^2 - \langle \Theta(\bar{y} - y_p), \bar{y} - y_p\rangle_{W(0,T)',W(0,T)} + \langle \mathcal{F}^\star\zeta, \bar{u} - u_p\rangle_U \\
&= -\sigma \,\|\bar{u} - u_p\|_U^2 + \langle \mathcal{F}^\star\zeta, \bar{u} - u_p\rangle_U \leq -\sigma \,\|\bar{u} - u_p\|_U^2 + \|\mathcal{F}^\star\zeta\|_U \|\bar{u} - u_p\|_U.
\end{aligned}$$

Hence, we get the a-posteriori error estimation

$$\|\bar{u} - u_p\|_U \leq \frac{1}{\sigma} \, \|\mathcal{F}^\star\zeta\|_U \tag{3.3}$$

If the evaluation of $\mathcal{F}^\star$ is much more expensive compared to the calculation of ζ, it may be advisable to deduce a-priori estimates for the operator norm of $\mathcal{F}^\star$ by the application of energy estimates for the state equations. However, these bounds are not rigorous in general, i.e. the true value of $\|\mathcal{F}^\star\zeta\|_U$ might be overestimated significantly.

3.2 Computation of a perturbation ζ

Now we want to determine an appropriate perturbation $\zeta \in U$ satisfying (3.1). Suppose that the suboptimal control $u_p = \mathcal{F}^{-1}v_p \in U$ is known, we derive from (3.1) the variational inequality

$$\langle \xi + \zeta, v - v_p\rangle_U \geq 0 \quad \forall v \in V_{\mathsf{ad}}, \tag{3.4}$$

where we set $\xi = \mathcal{F}^{-\star}(\sigma u_p - \mathcal{B}^\star q_p) \in U$. Hence, the element ξ solves

$$\big(\varepsilon + \mathcal{S}^\star\mathcal{I}^\star\big)\xi = \sigma u_p - \mathcal{B}^\star q_p \quad \text{in } U. \tag{3.5}$$

To determine an appropriate perturbation ζ we have to compute ξ. This is formulated in the next theorem. Its proof follows from Proposition 3.2 in [21].

Theorem 3.1 (A-posteriori error estimate for $(\hat{\mathbf{P}})$) *Suppose that u_p is an arbitrary control in U_{ad} and $q_p = \hat q + \mathcal{A}_1 u_p$. Moreover, the function ξ solves* (3.5)*. Define $\zeta \in U$ as follows:*

$$\zeta_i(t) = \begin{cases} -\min(0, \xi_i(t)) & \text{a.e. in } \mathcal{A}^p_{ai} = \big\{t \in [0,T]\,|\, v_{pi}(t) = v_{ai}(t)\big\}, \\ -\max(0, \xi_i(t)) & \text{a.e. in } \mathcal{A}^p_{bi} = \big\{t \in [0,T]\,|\, v_{pi}(t) = v_{bi}(t)\big\}, \\ -\xi_i(t) & \text{a.e. in } [0,T] \setminus \big(\mathcal{A}^p_{ai} \cup \mathcal{A}^p_{bi}\big) \end{cases} \tag{3.6}$$

for $1 \leq i \leq m$. Then, we have

$$\|\bar u - u_p\|_U \leq \epsilon_{\mathsf{ape}} \tag{3.7}$$

with the a-posteriori error estimator $\epsilon_{\mathsf{ape}} = \|\mathcal{F}^\star\zeta\|_U/\sigma$.

We call (3.7) an a-posteriori error estimate, since, in the next section, we will apply it to suboptimal controls u_p that have already been computed from a POD model.

4 Galerkin approximation for $(\mathbf{P})$ and $(\tilde{\mathbf{P}})$

The goal of this section is to apply the a-posteriori error analysis for a suboptimal control $u_p = \bar u^\ell$, which is computed by a POD Galerkin scheme for $(\hat{\mathbf{P}})$. In particular, we prove that the a-posteriori error estimator tends to zero if the number of POD ansatz functions tends to infinity. Let us mention that we avoid a discretization of the control space U; see also [9].

4.1 Proper orthogonal decomposition (POD)

Suppose that $w_1, \ldots, w_{\mathsf{k}} \in L^2(0,T;\mathcal{V})$ are $\mathsf{k} \geq 1$ given trajectories. Moreover, we introduce the linear subspace

$$\mathcal{V} = \operatorname{span}\Big\{ w_k(t) \,|\, t \in [0,T] \text{ and } k \in \{1,\ldots,\mathsf{k}\} \Big\}$$

with dimension $d \leq \infty$. The method of POD consists in choosing an orthonormal basis in V such that for every $\ell \in \{1,\ldots,d\}$ the mean square error between the elements $w_k(t)$, $t \in [0,T]$ and $1 \leq k \leq \mathsf{k}$, and the corresponding ℓ-th partial sum is minimized on average:

$$\begin{cases} \displaystyle\min_{\tilde\psi_1,\ldots,\tilde\psi_\ell \in V} \sum_{k=1}^{\mathsf{k}} \Big\| w_k - \sum_{i=1}^{\ell} \langle w_k, \tilde\psi_i \rangle_V \, \tilde\psi_i \Big\|^2_{L^2(0,T;V)} \\ \text{s.t. } \langle \tilde\psi_i, \tilde\psi_j \rangle_V = \delta_{ij}, \ 1 \leq i,j \leq \ell \end{cases} \tag{4.1}$$

with

$$\Big\| w_k - \sum_{i=1}^{\ell} \langle w_k, \tilde\psi_i \rangle_V \, \tilde\psi_i \Big\|^2_{L^2(0,T;V)} = \int_0^T \Big\| w_k(t) - \sum_{i=1}^{\ell} \langle w_k(t), \tilde\psi_i \rangle_V \, \tilde\psi_i \Big\|^2_V \, \mathrm{d}t.$$

A solution to (4.1) is called a *POD basis of rank ℓ*. We introduce the linear, bounded, nonnegative operator $\mathcal{R} : V \to V$ by

$$\mathcal{R}\psi = \sum_{k=1}^{\mathsf{k}} \int_0^T \langle w_k(t), \psi \rangle_V \, w_k(t) \, \mathrm{d}t \quad \text{for } \psi \in V.$$

The solution of (4.1) can be found in [11,22], for instance.

Proposition 4.1 *For $w_1, \ldots, w_{\mathsf{k}} \in L^2(0,T;V)$ the linear operator $\mathcal{R}$ is nonnegative, self-adjoint and compact. Let $\{\lambda_i\}_{i\in\mathbb{N}}$ and $\{\psi_i\}_{i\in\mathbb{N}}$ denote the nonnegative eigenvalues and associated orthonormal eigenfunctions of $\mathcal{R}$ satisfying*

$$\mathcal{R}\psi_i = \lambda_i \psi_i, \quad \lambda_1 \geq \lambda_2 \geq \ldots, \quad \textit{and } \lambda_i \to 0 \textit{ as } i \to \infty.$$

Then a POD basis of rank $\ell \leq d$ is given by $\{\psi_i\}_{i=1}^{\ell}$, and we have

$$\sum_{k=1}^{\mathsf{k}} \Big\| w_k - \sum_{i=1}^{\ell} \langle w_k, \psi_i \rangle_V \, \psi_i \Big\|^2_{L^2(0,T;V)} = \sum_{i=\ell+1}^{\infty} \lambda_i. \tag{4.2}$$

Remark 4.1 From the Hilbert-Schmidt theorem [18, p. 203] it follows that $\{\psi_i\}_{i\in\mathbb{N}}$ form a complete orthonormal basis for the separable space V. $\Diamond$

4.2 The POD Galerkin approximation

In this subsection we introduce the POD schemes for the first-order optimality system using a POD Galerkin approximation for the primal and dual variables. Moreover, we study the convergence of the POD discretizations, where we make use of the analysis in [10, 14, 21]. We make use of the following hypothesis.

Assumption 1 *Let $u, \lambda \in U$ be chosen such that the functions $\mathcal{S}u$, $\mathcal{A}_1 u$ and $\mathcal{A}_2\lambda$ are nonzero and belong to $H^1(0,T;V)$. In the context of Section 4.1 we choose $\mathsf{k} = 6$, $w_1 = \mathcal{S}u$, $w_2 = (\mathcal{S}u)_t$, $w_3 = \mathcal{A}_1 u$, $w_4 = (\mathcal{A}_1 u)_t$, $w_5 = \mathcal{A}_2\lambda$ and $w_6 = (\mathcal{A}_2\lambda)_t$.*

For $\ell \geq 1$ we denote by $\{\psi_i\}_{i=1}^{\ell}$ a POD basis of rank ℓ and set $V^\ell = \operatorname{span}\{\psi_1, \ldots, \psi_\ell\}$. Let us define the linear and bounded projection operator

$$\mathcal{P}^\ell \psi = \sum_{i=1}^{\ell} \langle \psi, \psi_i \rangle_V \, \psi_i \in V^\ell \quad \text{for } \psi \in V.$$

Suppose that $(\bar y, \bar u) \in X$ is the solution to $(\tilde{\mathbf{P}})$ and $(\bar p, \bar\lambda)$ the associated unique Lagrange mulitplier pair. Analogous to the operator $\mathcal{S}$ introduced in Remark 2.1 we define the operator $\mathcal{S}^\ell : U \to W_0(0,T)$ as follows: for $u \in U$ the function $y^\ell = \mathcal{S}^\ell u$ is the unique solution to

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle y^\ell(t), \psi \rangle_H + a(t; y^\ell(t), \psi) = \langle (\mathcal{B}u)(t), \psi \rangle_{V',V} \quad \forall \psi \in V^\ell \text{ a.e.}$$

Notice that for any $u \in U$ the element $\mathcal{S}^\ell u$ belongs even to $H^1(0,T;V^\ell)$ which is continuously embedded into $W(0,T)$.

Similar to Remark 2.2 we introduce the linear and bounded operators $\mathcal{A}_1^\ell$, $\mathcal{A}_2^\ell : U \to W(0,T)$: for given $u \in U$ the function $p^\ell = \mathcal{A}_1^\ell u$ uniquely solves

$$\begin{aligned} -\frac{\mathrm{d}}{\mathrm{d}t} \langle p^\ell(t), \psi \rangle_H + a(t; p^\ell(t), \psi) &= -\sigma_Q \, \langle (\mathcal{S}^\ell u)(t), \psi \rangle_{V',V} && \forall \psi \in V^\ell \text{ a.e.},\\ p^\ell(T) &= -\sigma_\Omega (\mathcal{S}^\ell u)(T) && \text{in } H \end{aligned}$$

and for given $\lambda \in U$ the function $p^\ell = \mathcal{A}_2\lambda$ is the unique solution to

$$\begin{aligned} -\frac{\mathrm{d}}{\mathrm{d}t} \langle p^\ell(t), \psi \rangle_H + a(t; p^\ell(t), \psi) &= -\langle (\mathcal{I}^\star \lambda)(t), \psi \rangle_{V',V} && \forall \psi \in V^\ell \text{ a.e.},\\ p^\ell(T) &= 0 && \text{in } H. \end{aligned}$$

Now, the POD Galerkin scheme for (2.9) is given by

$$\begin{aligned} y^\ell &= \hat y + \frac{1}{\sigma} \, \mathcal{S}^\ell \big(\mathcal{B}^\star p^\ell - \varepsilon \mathcal{N}(y^\ell, p^\ell) \big),\\ p^\ell &= \hat p + \frac{1}{\sigma} \, \mathcal{A}_1^\ell \big(\mathcal{B}^\star p^\ell - \varepsilon \mathcal{N}(y^\ell, p^\ell) \big) + \mathcal{A}_2^\ell \mathcal{N}(y^\ell, p^\ell). \end{aligned} \tag{4.3a}$$

Then we set

$$\lambda^\ell = \mathcal{N}(y^\ell, p^\ell), \quad u^\ell = \big(\mathcal{B}^\star p^\ell - \varepsilon\lambda^\ell)\big)/\sigma \quad \text{in } U. \tag{4.3b}$$

For a proof of the following proposition we refer the reader to [6, Theorems 3.1 and 3.2].

Proposition 4.2 *Let Assumption* 1 *hold. If* $\mathcal{S}\tilde{u}$ *even belongs to* $H^1(0,T;V)$ *for every* $\tilde{u} \in U$*, then we have*

$$\lim_{\ell\to\infty} \|\mathcal{S} - \mathcal{S}^\ell\|_{\mathcal{L}(U,W(0,T))} = \lim_{\ell\to\infty} \|\mathcal{A}_1 - \mathcal{A}_1^\ell\|_{\mathcal{L}(U,W(0,T))} = 0,$$

where $\mathcal{L}(U,W(0,T))$ *stands for the Banach space of all linear and bounded operators from* U *to* $W(0,T)$ *equipped with its natural operator norm.*

To introduce a POD Galerkin scheme for (2.13) let us introduce the POD aproximation $\mathcal{F}^\ell : U \to U$ of the operator $\mathcal{F}$ by $\mathcal{F}^\ell = \varepsilon + \mathcal{I}\mathcal{S}^\ell : U \to U$. The proof of the following proposition is given in the Appendix.

Proposition 4.3 *Suppose that Assumption* 1 *is valid. Then, we have:*

1) $\mathcal{F}^\ell$ *is linear, bounded (uniformly with respect to* ℓ*),* $\lim\limits_{\ell\to 0} \|\mathcal{F} - \mathcal{F}^\ell\|_{\mathcal{L}(U)} = 0$.
2) *Suppose that* $\mathcal{F}$ *has a bounded inverse. Then,* $\mathcal{F}^\ell$ *has a bounded inverse as well. We write* $\mathcal{F}^{\ell,-1} = (\mathcal{F}^\ell)^{-1}$*. If* $L \in \mathbb{N}$ *be chosen arbitrarily large so that* $\|\mathcal{F}^{-1}\|_{\mathcal{L}(U)}\|\mathcal{F} - \mathcal{F}^\ell\|_{\mathcal{L}(U)} < 1$ *holds for all* $\ell > L$*, we obtain*

$$\|\mathcal{F}^{\ell,-1}\|_{\mathcal{L}(U)} \le \frac{\|\mathcal{F}^{-1}\|_{\mathcal{L}(U)}}{1 - \|\mathcal{F}^{-1}\|_{\mathcal{L}(U)}\|\mathcal{F} - \mathcal{F}^\ell\|_{\mathcal{L}(U)}} \quad \text{for all } \ell \ge L.$$

3) *We have* $\lim_{\ell\to\infty} \|1 - \mathcal{F}\mathcal{F}^{\ell,-1}\|_{\mathcal{L}(U)} = 0$, $\lim_{\ell\to\infty} \|1 - \mathcal{F}^{\ell,-1}\mathcal{F}\|_{\mathcal{L}(U)} = 0$ *and* $\lim_{\ell\to\infty} \|1 - \mathcal{F}^\star\mathcal{F}^{\ell,-\star}\|_{\mathcal{L}(U)} = 0$*, where* $\mathcal{F}^{\ell,-\star}$ *denotes the adjoint operator of* $\mathcal{F}^{\ell,-1}$.

Let us formulate a discrete version of the primal-dual active set method which is utilized in our numerical tests to solve (4.3a) and (4.3b). For given iterates $(y^{\ell k}, p^{\ell k}) \in W(0,T)\times W(0,T)$, $k \ge 0$, and for $i = 1,\dots,m$ we introduce the active and inactive sets

$$\begin{aligned}
\mathcal{A}_{ai}^{\ell k} &= \Big\{ t \in [0,T] \,\Big|\, (\varepsilon\sigma^{-1}\mathcal{B}^\star p^{\ell k} + \mathcal{I}y^{\ell k})_i(t) < u_{ai}(t) \Big\}, \ \mathcal{A}_i^{\ell k} = \mathcal{A}_{ai}^{\ell k} \cup \mathcal{A}_{bi}^{\ell k},\\
\mathcal{A}_{bi}^{\ell k} &= \Big\{ t \in [0,T] \,\Big|\, (\varepsilon\sigma^{-1}\mathcal{B}^\star p^{\ell k} + \mathcal{I}y^{\ell k})_i(t) > u_{bi}(t) \Big\}, \ \ \mathcal{I}_i^{\ell k} = [0,T]\backslash\mathcal{A}_i^{\ell k};
\end{aligned} \tag{4.4}$$

compare (2.10). In Algorithm 2 we state the semismooth Newton method for the POD discretized problem.

Algorithm 2 (POD discretized primal-dual active set strategy)

Require: POD basis ψ^ℓ, starting value $(y^{\ell 0}, p^{\ell 0})$ and maximal iteration number k_{max}.
1: Set $k = 0$, determine the active and inactive sets according to (4.4).
2: **repeat**
3: Determine the solution $(y^{\ell,k+1}, p^{\ell,k+1})$ to (4.3a) and set $k = k + 1$.
4: Compute the active and inactive sets according to (4.4).
5: **until** $(\mathcal{A}^{\ell k}_{ai} = \mathcal{A}^{\ell,k-1}_{ai}$ **and** $\mathcal{A}^{\ell k}_{bi} = \mathcal{A}^{\ell,k-1}_{bi})$ **or** $k = k_{\mathsf{max}}$.
6: Set $u^\ell = (\mathcal{B}^\star p^{\ell k} - \varepsilon\lambda^{\ell k})/\sigma$.

Utilizing the operator $\mathcal{F}^\ell$ we introduce the POD Galerkin scheme for (2.13) as follows

$$\bar y^\ell = \hat y + \mathcal{S}^\ell \mathcal{F}^{\ell,-1}\bar v^\ell, \tag{4.5a}$$

$$\bar q^\ell = \hat p + \mathcal{A}_1^\ell \mathcal{F}^{\ell,-1}\bar v^\ell \tag{4.5b}$$

$$\langle \sigma\mathcal{F}^{\ell,-\star}\mathcal{F}^{\ell,-1}\bar v - \mathcal{F}^{\ell,-\star}\mathcal{B}^\star\bar q^\ell, v - \bar v^\ell\rangle_U \geq 0 \quad \forall v \in V_{\mathsf{ad}}, \tag{4.5c}$$

In contrast to the error analysis in [10,21] the discretized variational inequality involves the operators $\mathcal{F}^\ell$. This reflects the fact that the POD discretization

$$U^\ell_{\mathsf{ad}} = \big\{u \in U \,\big|\, u_a \leq \varepsilon u + \mathcal{I}\big(\hat y + \mathcal{S}^\ell u\big) \leq u_b \text{ in } \mathbb{R}^m \text{ a.e. in } [0,T]\big\}$$

of the admissible set U_{ad} depends on the POD Galerkin scheme, whereas in [10, 21] the admissible set for the controls is independent of the POD discretization. On the other hand, V_{ad} is independent of ℓ. In particular, we have $\bar v^\ell = \mathcal{F}^\ell \bar u^\ell$. The proof of the next theorem is given in the Appendix.

Theorem 4.1 *Suppose that Assumption* 1 *holds.*

1) *For $\bar q$ and $\bar q^\ell$ we have* $\lim_{\ell\to\infty}\|\bar q - \bar q^\ell\|_{W(0,T)} = 0$
2) *Let $\bar v$ and $\bar v^\ell$ be the solutions to* (2.13c) *and* (4.5c), *respectively. Then,* $\lim\limits_{\ell\to\infty}\|\bar v - \bar v^\ell\|_U = 0$.
3) *If $\bar u$ and $\bar u^\ell$ are the solutions to* (2.7) *and* (4.3), *respectively, we obtain* $\lim\limits_{\ell\to\infty}\|\bar u - \bar u^\ell\|_U = 0$.
4) *Define, according to* (3.6), *the function $\zeta^\ell \in U$ by*

$$\zeta^\ell_i(t) = \begin{cases} -\min(0, \xi^\ell_i(t)) & \text{a.e. in } \mathcal{A}^\ell_{ai} = \big\{t \in [0,T] \,|\, \bar v^\ell_i(t) = v_{ai}(t)\big\}, \\ -\max(0, \xi^\ell_i(t)) & \text{a.e. in } \mathcal{A}^\ell_{bi} = \big\{t \in [0,T] \,|\, v^\ell_i(t) = v_{bi}(t)\big\}, \\ -\xi^\ell_i(t) & \text{a.e. in } [0,T] \setminus \big(\mathcal{A}^\ell_{ai} \cup \mathcal{A}^\ell_{bi}\big), \end{cases}$$

where $\xi^\ell \in U$ solves $\mathcal{F}^\star\xi^\ell = \sigma\bar u^\ell - \mathcal{B}^\star\bar q^\ell$ in U; compare (3.5). *Then,*

$$\|\bar u - \bar u^\ell\|_U \leq \epsilon_{\mathsf{ape}} \quad \text{with } \epsilon_{\mathsf{ape}} = \frac{\|\mathcal{F}^\star\zeta^\ell\|_U}{\sigma} \tag{4.6}$$

and $\lim\limits_{\ell\to\infty}\big\|\zeta^\ell\big\|_U = 0$.

Remark 4.2 1) In addition to the a-posteriori error analysis in [21], we have to solve the linear system $\mathcal{F}^\star \xi^\ell = \sigma \bar{u}^\ell - \mathcal{B}^\star \bar{q}^\ell$.

2) Part 3) of Theorem 4.1 shows that $\|\zeta^\ell\|_U$ can be expected smaller than any $\epsilon > 0$ provided that ℓ is taken sufficiently large. Motivated by this result we set up Algorithm 3. $\Diamond$

Algorithm 3 (POD reduced-order method with a-posteriori estimator)

Require: Initial control $u^{0\ell} \in U$, initial number ℓ for the POD ansatz functions, a maximal number $\ell_{\mathsf{max}} > \ell$ of POD ansatz functions, and a stopping tolerance $\epsilon > 0$.

Determine $\hat{y}$, $\hat{q}$, $w_1 = \mathcal{S}u^{0\ell}$, $w_2 = \mathcal{A}u^{0\ell}$.
Compute a POD basis $\{\psi_i\}_{i=1}^{\ell_{\mathsf{max}}}$ choosing w_1 and w_2. Set $\ell = 1$.
repeat
    Establish the POD Galerkin discretization using $\{\psi_i\}_{i=1}^{\ell}$.
    Call Algorithm 2 to compute suboptimal control $\bar{u}^\ell$.
    Determine ϵ_{ape} from (4.6).
    if $\epsilon_{\mathsf{ape}} < \epsilon$ **or** $\ell = \ell_{\mathsf{max}}$ **then**
        Return ℓ and suboptimal control $\bar{u}^\ell$ and STOP.
    end if
    Set $\ell = \ell + 1$.
until $\ell > \ell_{\mathsf{max}}$

5 Numerical experiments

In this section we carry out numerical test examples for the presented theoretical findings.

Run 5.1 In the context of Section 2.6 we choose $d = 1$, $\Omega = (0, 2)$, $\Omega_i = \frac{2}{m}[i-1, i]$ for $1 \le i \le m$, $\chi_i(\boldsymbol{x}) = \pi_i(\boldsymbol{x}) = \chi_{\Omega_i}(\boldsymbol{x})$ the characteristic functions on the subdomains Ω_i and $T = 3$. Let $\sigma =$ 1e-3, $\sigma_Q = 1$, $\sigma_\Omega = 0$, $\varepsilon =$ 1e-5 and $f(t, \boldsymbol{x}) = t - \boldsymbol{x}^3$, $y_Q(t, \boldsymbol{x}) = t(1 - (\boldsymbol{x} - 1)^2)$, $y_\Omega(\boldsymbol{x}) = y_Q(T, \boldsymbol{x})$, $y_\circ(\boldsymbol{x}) = (\chi_{[0.4,1.0]} - \chi_{[1.0,1.6]})(\boldsymbol{x})$.

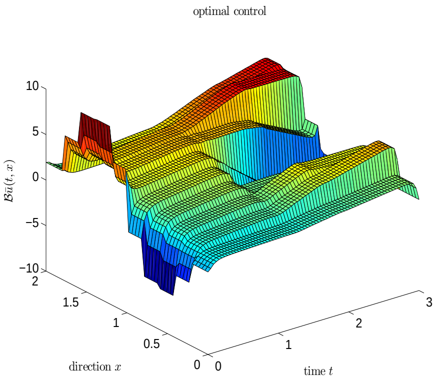

Fig. 5.1 Run 5.1: The optimal control $\mathcal{B}\bar{u}$ and the optimal state $\bar{y}$ for $u_a = -0.25$, $u_b = 0.25$ and $m = 10$ calculated by solving the full order model.

Notice that for large m and small regularization ε, the mixed control-state constraints can be interpreted as pointwise state constraints: As one sees in Figure 5.1, $\frac{m}{2}u_a \leq y \leq \frac{m}{2}u_b$, i.e. $-1.25 \leq y(t,x) \leq 1.25$, holds approximatively.

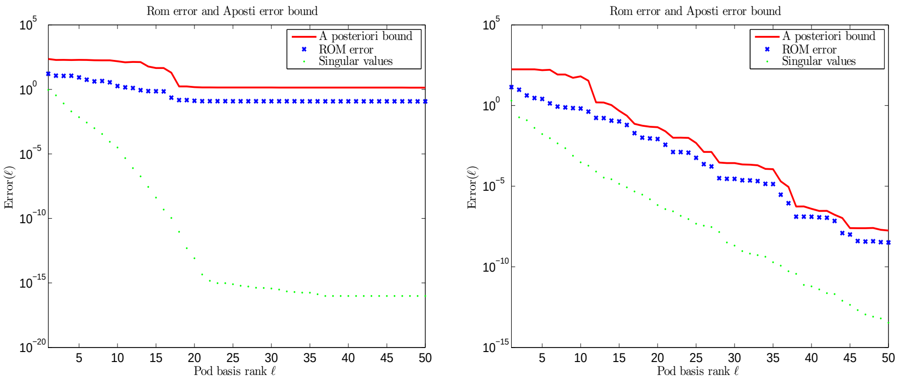

Fig. 5.2 Run 5.1: The ROM errors of the control $\bar{u}^\ell$ for different POD basis ranks ℓ with initial control guesses $u^{0\ell} \equiv 1$ (left) and $u^{0\ell} = \bar{u}$ (right).

In Figure 5.2 we investigate the decay of the ROM errors for increasing POD basis rank ℓ. For this purpose, we apply Algorithm 2 for the (fixed) POD bases corresponding to the snapshots $\mathcal{S}u^{0\ell}$ with $u^{0\ell} \equiv 1$ and $u^{0\ell} = \bar{u}$. One sees that increasing the basis rank of the arbitrarily chosen POD basis does not lead to a satisfying accuracy of the reduced order model solution since the state solution $\mathcal{S}u^{0\ell}$ does not cover enough of the dynamics of the optimal state $\bar{y}$, so POD basis updates which exploite the information gained from the actual suboptimal control are required to get snapshots which fit to $\bar{u}$.

Algorithm 4 (POD Primal-dual active set strategy with basis adaptivity)

Require: Initial control guess $u^{0\ell}$, tolerance ϵ and maximal iterations k_{max}.
Set $k = 0$.
repeat
  Compute the solution $y^k = \hat{y} + \mathcal{S}u^{k\ell} \in W(0,T)$ to the system (2.3).
  Compute a rank-ℓ POD basis $\psi^\ell \subseteq V$ by solving (4.1) with $w = y^k$.
  Execute Algorithm 2 and set $u^{k\ell} = \bar{u}^\ell$, $k = k + 1$.
until $\epsilon_{\mathsf{ape}}(u^{k\ell}) < \epsilon$ **or** $k = k_{\mathsf{max}}$.
Return $\bar{u}^\ell = u^{k\ell}$.

Since in practice we do not have enough knowledge about $\bar{u}$ to choose a-priori an admissible POD basis, we present an adaptive strategy in Algorithm 4 which actualizes the snapshots allocations iteratively with the information gained from the ROM solutions. Now, one single basis update is sufficient to achieve the same decay order of the error caused by the model reduction as we have for the (in general unknown) optimal POD basis.

In contrast to our experiences with pure control constrained optimal control problems, the a posteriori error bounds are not rigorous, i.e. the true ROM error has a smaller order than the error estimation. This is due to the fact that we have chosen the regularization parameter ε of the state constraints very small which leads to a badly scaled discretization of the differential equations required to determine the perturbation vector ζ. For larger values of ε, the error bounds are rigorous if the POD basis rank ℓ is not chosen too small, see Figure 5.3. In our example, the primal-dual active set strategy does not converge within the claimed maximal number of active set actualizations if $\ell < 12$.

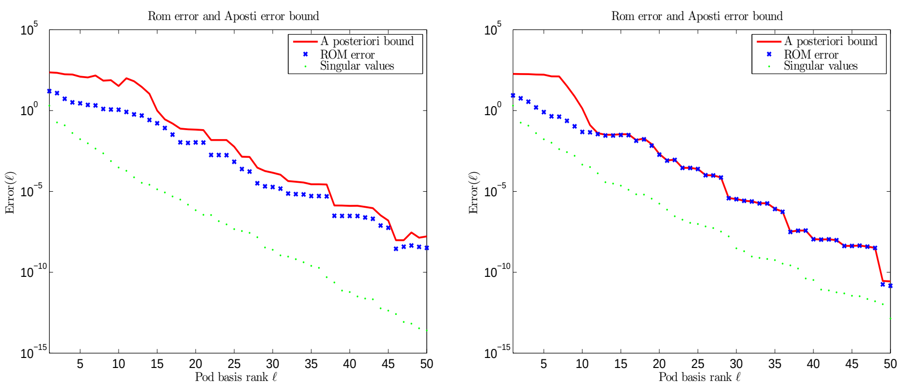

Fig. 5.3 Run 5.1: The ROM errors of the control $\bar{u}^\ell$ for different POD basis ranks ℓ with adaptive basis selections and regularization parameters $\varepsilon = $ 1e-5 (left) and $\varepsilon = $ 5e-3 (right).

Process	Time	#	Total
Assemble full system (2.9)	0.66 sec	9×	5.97 sec
Solve full system (2.9)	22.27 sec	9×	200.43 sec
Total			206.40 sec
Solve full snapshots equations (2.13a)	0.11 sec	2×	0.21 sec
Solve eigenvalue problem (4.1)	0.24 sec	2×	0.84 sec
Assemble ROM system (4.3a)	0.53 sec	17×	9.01 sec
Solve ROM system (4.3a)	0.45 sec	17×	7.72 sec
Evaluate error estimator (4.6)	0.11 sec	2×	0.23 sec
Total			18.01 sec

Table 1 Run 5.1: The calculation times for solving the optimization problem with the primal-dual active set strategy with and without model reduction. With 25 POD elements, the reduced-order problem has to be solved two times; solvings of two eigenvalue problems are required in addition for updating the POD basis. Nevertheless, 91.27% of the calculation time is spared in total where the true error is 9.07e-05 and the a-posteriori error bound totals up to 5.68e-4.

We finish the first example with a look at Table 1 where the effort of the model reduction on the calculation times is illustrated. Since the most expensive part of the optimization process is the simultaneous solving of the primal and dual equations, ROM is very effective here even for POD basis ranks which are chosen so large that the accuracy of the full order model is reached.

Run 5.2 Choosing $\varepsilon = 1$ and $\mathcal{I} = 0$, the techniques presented also cover the case of pure control constraints. The optimal control and state solutions induced by the same data functions as we used before for the state constrained problem are shown in Figure 5.4:

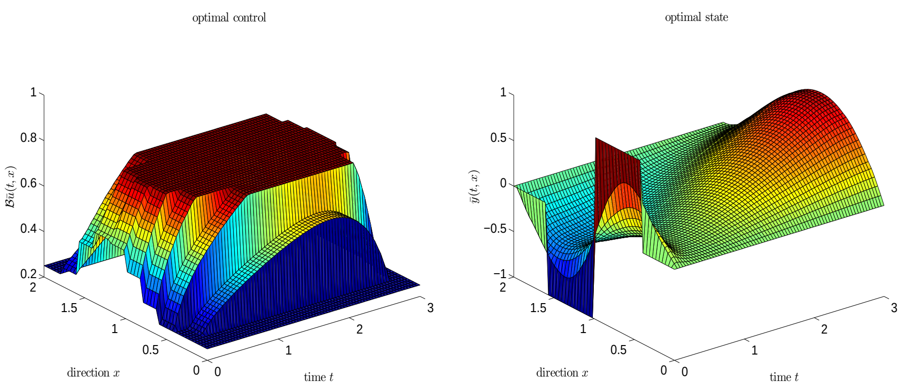

Fig. 5.4 Run 5.2: The optimal control $\mathcal{B}\bar{u}$ and the optimal state $\bar{y}$ for $u_a = 0.25$, $u_b = 0.75$ and $m = 10$ calculated by solving the full order model.

In this case, the a-posteriori error bounds are rigorous. Figure 5.5 illustrates that even the arbitrary POD basis corresponding to the initial control guess $u^{0\ell} \equiv 1$ coveres enough dynamics of $\bar{u}$ to decrease the ROM errors below the FEM accuracy which is of the order 1.0e-04 without basis updates.

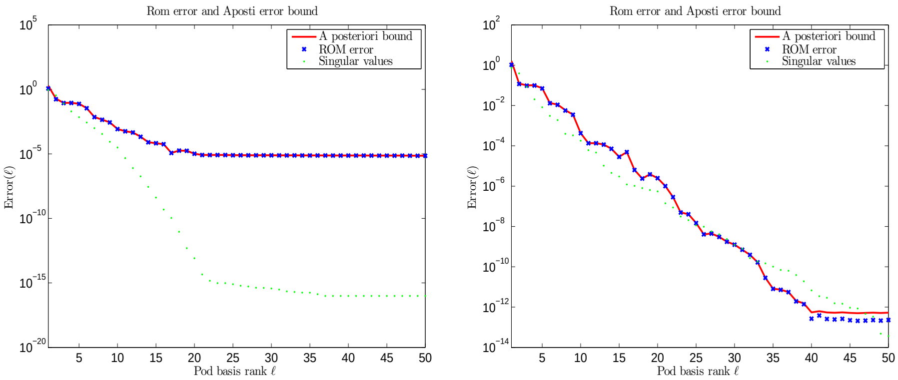

Fig. 5.5 Run 5.2: The ROM errors of the control $\bar{u}^\ell$ for different POD basis ranks ℓ with a fixed POD basis (left) and with basis adaptivity (right).

For practical application it does not make sense to enlarge the reduced order model, i.e. the POD basis rank, as soon as the a-posteriori error estimator indicates that the exactness of the full order model is reached since the FEM error is part of the snapshots and hence of the reduced order model anyway.

Process	**Time**	**#**	**Total**
Assemble full system (2.9)	0.79 sec	4×	3.14 sec
Solve full system (2.9)	18.53 sec	4×	74.11 sec
Total			77.25 sec
Solve full snapshots equations (2.13a)	0.10 sec	1×	0.10 sec
Solve eigenvalue problem (4.1)	0.17 sec	1×	0.17 sec
Assemble ROM system (4.3a)	0.40 sec	4×	1.61 sec
Solve ROM system (4.3a)	0.12 sec	4×	0.47 sec
Evaluate error estimator (4.6)	0.13 sec	1×	0.13 sec
Total			2.48 sec

Table 2 Run 5.2: The calculation times for solving the optimization problem with the primal-dual active set strategy with and without model reduction. With 25 POD elements, the reduced-order problem has to be solved two times; solvings of two eigenvalue problems are required in addition to update the POD basis. Nevertheless, 96.79% of the calculation time is spared in total and both the true error as well as the a-posteriori error bound amount to 5.26e-6.

Run 5.3 We consider now a two-dimensional setting: Let $\Omega = (0, \pi) \times (0, \pi)$, $\Omega_{ij} = \frac{\pi}{m}[i-1, i] \times [j-1, j]$ for $1 \le i, j \le m$, $\chi_{ij}(\boldsymbol{x}) = \pi_{ij}(\boldsymbol{x}) = \chi_{\Omega_{ij}}(\boldsymbol{x})$ be the characteristic functions on the subdomains Ω_{ij} and $T = \frac{\pi}{2}$. Let $\sigma =$ 1e-3, $\sigma_Q = 1$, $\sigma_\Omega = 1$, $\varepsilon =$ 1e-5. We choose $f \equiv 0$, $y_\circ \equiv 0$ and consistent desired states $y_Q(t, \boldsymbol{x}) = \sin(t) \sin(\boldsymbol{x}_1) \sin(\boldsymbol{x}_2)$, $y_\Omega = y_Q(T, \cdot)$. Obviously, $y = y_Q$ is an optimal state solution to the optimization problem without state or control constraints if $\sigma = 0$. Figure 5.6 shows that the state solution which respects the weakly regularized pointwise state constraints resembles the projection of y_Q on the admissible range.

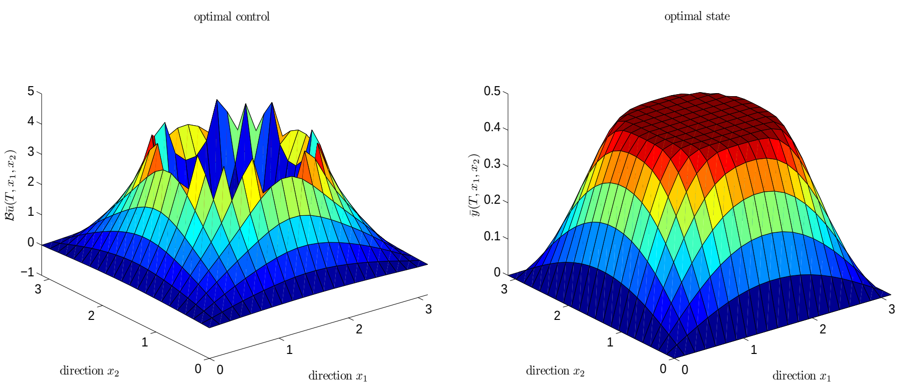

Fig. 5.6 Run 5.3: The optimal control $(\mathcal{B}\bar{u})(T)$ and the optimal state $\bar{y}(T)$ for $u_{a,ij} = -0.5|\Omega_{ij}|$, $u_{b,ij} = +0.5|\Omega_{ij}|$ and $m = 400$ calculated by solving the full order model.

An arbitrary POD basis is not able to establish the area of the graph where the upper constraints are active in this case, see Figure 5.7: Basis updates are required instead to build up an accurate reduced-order model.

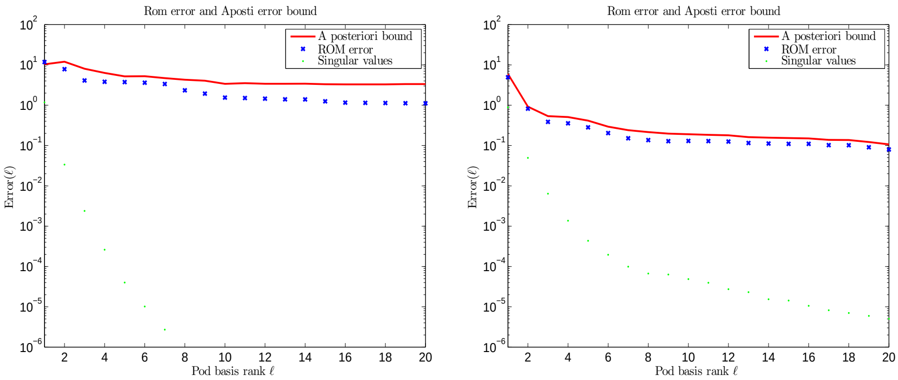

Fig. 5.7 Run 5.3: The ROM errors of the control $\bar{u}^\ell$ for different POD basis ranks ℓ with a fixed POD basis (left) and with basis adaptivity (right).

Process	**Time**	**#**	**Total**
Assemble full system (4.3a)	0.25 sec	39×	9.90 sec
Solve full system (4.3a)	32.16 sec	39×	1254.11 sec
Total			1264.01 sec
Solve full snapshots equations (2.13a)	0.21 sec	2×	0.42 sec
Solve eigenvalue problem (4.1)	0.10 sec	2×	0.19 sec
Assemble ROM system (4.3a)	0.12 sec	50×	6.18 sec
Solve ROM system (4.3a)	0.03 sec	50×	1.69 sec
Evaluate error estimator (4.6)	0.54 sec	2×	1.08 sec
Total			9.56 sec

Table 3 Run 5.3: The calculation times for solving the optimization problem with the primal-dual active set strategy with and without model reduction. With 25 POD elements, the reduced-order problem has to be solved two times; solvings of two eigenvalue problems are required in addition to update the POD basis. Nevertheless, 99.24% of the calculation time is spared in total. The a posteriori error bound is of the same order as the ROM error which amount to 8.74e-2.

Appendix

Proof of Lemma 2.1

We utilize arguments from the proof of Lemma 4.1 in [16]. First we show that $\mathcal{F}$ is injective. From $\mathcal{F}u = 0$ we infer $(\varepsilon + \mathcal{I}\mathcal{S})u = 0$ in U. Since $\varepsilon > 0$ holds, we derive $u = -\frac{1}{\varepsilon}\mathcal{I}\mathcal{S}u$ in U. By Remark 2.1, we conclude that $y = \mathcal{S}u \in W(0,T)$ satisfies

$$\begin{aligned} \langle y_t(t), \varphi\rangle_{V',V} + a(t; y(t), \varphi) &= -\frac{1}{\varepsilon}\langle (\mathcal{B}\mathcal{I}y)(t), \varphi\rangle_{V',V}, \qquad && \forall \varphi \in V \text{ a.e. in } [0,T],\\ \langle y(0), \varphi\rangle_H &= 0 && \forall \varphi \in H. \end{aligned}$$

Choosing $\varphi = y(t)$ and using (2.2), (2.11) and Young's inequality [5, p. 622],

$$|\mathsf{ab}| \leq \frac{\tilde{\epsilon}\mathsf{a}^2}{2} + \frac{\mathsf{b}^2}{2\tilde{\epsilon}} \quad \text{for all } \mathsf{a}, \mathsf{b} \in \mathbb{R}, \tilde{\epsilon} > 0 \tag{A.1}$$

with $\mathsf{a} = \|y(t)\|_V$, $\mathsf{b} = \|y(t)\|_H/(C_\epsilon \varepsilon)$, we deduce that

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|y(t)\|_H^2 + \alpha_1\|y(t)\|_V^2 - \alpha_2\|y(t)\|_H^2 \leq \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|y(t)\|_H^2 + a(t;y(t),y(t)) \\
&\leq \frac{1}{\varepsilon}\|(\mathcal{B}\mathcal{I}y)(t)\|_{V'}\|y(t)\|_V \leq \frac{1}{\varepsilon}(C_\epsilon\|y(t)\|_H + \epsilon\|y(t)\|_V)\|y(t)\|_V \\
&\leq \left(\frac{\tilde{\epsilon}}{2} + \frac{\epsilon}{\varepsilon}\right)\|y(t)\|_V^2 + \frac{1}{2\tilde{\epsilon}}\frac{C_\epsilon^2}{\varepsilon^2}\|y(t)\|_H^2 \qquad \text{a.e. in } [0,T].
\end{aligned}$$

Taking $\epsilon = \varepsilon\alpha_1/4$ and $\tilde{\epsilon} = \alpha_1/2$, we get the a-priori energy estimate

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|y(t)\|_H^2 + \frac{\alpha_1}{2}\|y(t)\|_V^2 - \left(\alpha_2 + \frac{C_\epsilon^2}{\varepsilon^2\alpha_1}\right)\|y(t)\|_H^2 \leq 0 \quad \text{a.e. in } [0,T]. \tag{A.2}$$

Applying Gronwall's lemma [5, pp. 624-625] and $y(0) = 0$ in H, we obtain by standard arguments for linear evolution problems that $y \equiv 0$ holds in $W(0,T)$. Thus, the operator $\mathcal{F}$ is injective. Next we prove that $\mathcal{F}$ is surjective. Let $v \in U$ be chosen arbitrarily. Then, $\mathcal{F}$ is surjective if there exists an $u \in U$ satisfying $\varepsilon u + \mathcal{I}\mathcal{S}u = v$ in U. Deriving again an a-priori energy estimate analogously to (A.2), we conclude that

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\langle y(t),\varphi\rangle_H + a(t;y(t),\varphi) + \frac{1}{\varepsilon}\langle\mathcal{B}\mathcal{I}y(t),\phi\rangle_{V',V} &= \frac{1}{\varepsilon}\langle\mathcal{B}v(t),\varphi\rangle_{V',V} \qquad \forall\varphi \in V \text{ a.e.},\\
\langle y(0),\varphi\rangle_H &= 0 \qquad \forall\varphi \in H
\end{aligned}$$

admits a unique solution $y \in W(0,T)$. Defining $u = (v - \mathcal{I}y)/\varepsilon \in U$, $y = \mathcal{S}u$ holds, i.e. $\mathcal{F}u = (\varepsilon + \mathcal{I}\mathcal{S})u = v$ in U which implies the surjectivity of $\mathcal{F}$. Now the claim follows from the bounded inverse theorem.

Proof of Proposition 4.3

1) Clearly, $\mathcal{F}^\ell$ is linearand bounded by a constant which is independent of ℓ. Since

$$\|\mathcal{F} - \mathcal{F}^\ell\|_{\mathcal{L}(U)} = \|\mathcal{I}(\mathcal{S} - \mathcal{S}^\ell)\|_{\mathcal{L}(U)} \leq \|\mathcal{I}\|_{\mathcal{L}(L^2(0,T;V),U)}\|\mathcal{S} - \mathcal{S}^\ell\|_{\mathcal{L}(U,W(0,T))}$$

is satisfied, the convergence follows directly from Proposition 4.2.

2) By part 1) there exists a constant $L \in \mathbb{N}$ so that

$$\|\mathcal{F} - \mathcal{F}^\ell\|_{\mathcal{L}(U)} < \frac{1}{\|\mathcal{F}^{-1}\|_{\mathcal{L}(U)}} \quad \text{for all } \ell \geq L.$$

Then, the claim follows from the perturbation lemma [17, p. 45].

3) Using

$$\begin{aligned}
\left\|1 - \mathcal{F}\mathcal{F}^{\ell,-1}\right\|_{\mathcal{L}(U)} &= \left\|(\mathcal{F}^\ell - \mathcal{F})\mathcal{F}^{\ell,-1}\right\|_{\mathcal{L}(U)} \leq \left\|\mathcal{F}^\ell - \mathcal{F}\right\|_{\mathcal{L}(U)}\left\|\mathcal{F}^{\ell,-1}\right\|_{\mathcal{L}(U)},\\
\left\|1 - \mathcal{F}^{\ell,-1}\mathcal{F}\right\|_{\mathcal{L}(U)} &= \left\|\mathcal{F}^{\ell,-1}(\mathcal{F}^\ell - \mathcal{F})\right\|_{\mathcal{L}(U)} \leq \left\|\mathcal{F}^{\ell,-1}\right\|_{\mathcal{L}(U)}\left\|\mathcal{F}^\ell - \mathcal{F}\right\|_{\mathcal{L}(U)}
\end{aligned}$$

and parts 2), 3) the limits hold.

Proof of Theorem 4.1

1) The claim follows from [6, Theorem 3.2].
2) Choosing $v = \bar v^\ell$ in (2.13c) and $v = \bar v$ in (4.5c) we get the variational inequality

$$\Big\langle \sigma\big(\mathcal{F}^{-\star}\mathcal{F}^{-1}\bar v - \mathcal{F}^{\ell,-\star}\mathcal{F}^{\ell,-1}\bar v^\ell\big) - \mathcal{F}^{\ell,-\star}\mathcal{B}^\star\bar q^\ell + \mathcal{F}^{-\star}\mathcal{B}^\star\bar q, \bar v^\ell - \bar v\Big\rangle_U \ge 0. \tag{A.3}$$

From $\bar v^\ell \in V_{\mathsf{ad}}$ we infer that $\|\bar v^\ell\|$ is bounded. Thus, there exists a constant $C_1 > 0$ with $\|\bar v^\ell\|_U \le C_1$. Moreover, $\|\mathcal{F}^{\ell,-1}\|_{\mathcal{L}(U)} \le 2C_2$ for ℓ sufficiently large with $C_2 = \|\mathcal{F}^{-1}\|_{\mathcal{L}(U)}$. Thus,

$$\begin{aligned}
&\sigma\,\langle \mathcal{F}^{-\star}\mathcal{F}^{-1}\bar v - \mathcal{F}^{\ell,-\star}\mathcal{F}^{\ell,-1}\bar v^\ell, \bar v^\ell - \bar v\rangle_U\\
&\quad= \sigma\,\langle \mathcal{F}^{-\star}\mathcal{F}^{-1}(\bar v - \bar v^\ell) + \mathcal{F}^{-\star}\mathcal{F}^{-1}\bar v^\ell - \mathcal{F}^{\ell,-\star}\mathcal{F}^{\ell,-1}\bar v^\ell, \bar v^\ell - \bar v\rangle_U\\
&\quad= -\sigma\,\|\mathcal{F}^{-1}(\bar v - \bar v^\ell)\|_U^2 + \sigma\,\langle(\mathcal{F}^{-1} - \mathcal{F}^{\ell,-1})\bar v^\ell, \mathcal{F}^{-1}(\bar v^\ell - \bar v)\rangle_U\\
&\qquad+ \langle(1 - \mathcal{F}^\star\mathcal{F}^{\ell,-\star})\mathcal{F}^{\ell,-1}\bar v^\ell, \mathcal{F}^{-1}(\bar v^\ell - \bar v)\rangle_U\\
&\quad\le \sigma\,\Big(-\|\mathcal{F}^{-1}(\bar v - \bar v^\ell)\|_U + C_3\big(\|1 - \mathcal{F}\mathcal{F}^{\ell,-1}\|_{\mathcal{L}(U)} + \|1 - \mathcal{F}^\star\mathcal{F}^{\ell,-\star}\|_{\mathcal{L}(U)}\big)\Big)\\
&\qquad\quad\cdot \|\mathcal{F}^{-1}(\bar v^\ell - \bar v)\|_U
\end{aligned}$$

with $C_3 = 2C_1C_2$. We set $C_4 = C_1\max(2C_2, \|\mathcal{B}^\star\mathcal{A}_1\mathcal{F}^{-1}\|_{\mathcal{L}(U)}$ and $C_5 = \|\mathcal{B}\|_{\mathcal{L}(U)}$. Recall that the operator $\mathcal{S}^\ell$ is bounded independently of ℓ. Hence, the constant $C_6 = 2C_1C_2\,\|\mathcal{S}^{\ell,\star}\Theta\mathcal{S}^\ell\|_{\mathcal{L}(U)}$ does not depend on ℓ. Moreover, $\mathcal{B}^\star\mathcal{A}_1^\ell = -\mathcal{S}^{\ell,\star}\Theta\mathcal{S}^\ell$ is uniformly bounded. Hence, there exists a constant $C_7 > 0$ which does not depend on ℓ so that $\|\mathcal{B}^\star\mathcal{A}_1^\ell\|_{\mathcal{L}(U)} \le C_7$. Consequently, Lemma 2.2 implies that

$$\begin{aligned}
&\Big\langle \mathcal{F}^{-\star}\mathcal{B}^\star\bar q - \mathcal{F}^{\ell,-\star}\mathcal{B}^\star\bar q^\ell, \bar v^\ell - \bar v\Big\rangle_U\\
&= \Big\langle \mathcal{F}^{-\star}\mathcal{B}^\star\mathcal{A}_1\mathcal{F}^{-1}\big(\bar v - \bar v^\ell\big) + \mathcal{F}^{-\star}\mathcal{B}^\star\mathcal{A}_1\big(\mathcal{F}^{-1} - \mathcal{F}^{\ell,-1}\big)\bar v^\ell, \bar v^\ell - \bar v\Big\rangle_U\\
&\quad+ \Big\langle \mathcal{F}^{-\star}\mathcal{B}^\star\big(\mathcal{A}_1 - \mathcal{A}_1^\ell\big)\mathcal{F}^{\ell,-1}\bar v^\ell + \big(\mathcal{F}^{-\star} - \mathcal{F}^{\ell,-\star}\mathcal{B}^\star\mathcal{A}_1^\ell\mathcal{F}^{\ell,-1}\bar v^\ell, \bar v^\ell - \bar v\Big\rangle_U\\
&\le -\Big\langle \Theta\mathcal{S}\mathcal{F}^{-1}\big(\bar v - \bar v^\ell\big), \mathcal{S}\mathcal{F}^{-1}\big(\bar v^\ell - \bar v\big)\Big\rangle_U\\
&\quad+ C_4\left(\big\|1 - \mathcal{F}\mathcal{F}^{\ell,-1}\big\|_{\mathcal{L}(U)} + \big\|\mathcal{B}^\star(\mathcal{A}_1 - \mathcal{A}_1^\ell)\big\|_{\mathcal{L}(U)}\right)\big\|\mathcal{F}^{-1}\big(\bar v - \bar v^\ell\big)\big\|_U\\
&\quad+ C_6\,\big\|1 - \mathcal{F}^\star\mathcal{F}^{\ell,-\star}\big\|_{\mathcal{L}(U)}\big\|\mathcal{F}^{-1}\big(\bar v - \bar v^\ell\big)\big\|_U\\
&\le C_7\left(\big\|1 - \mathcal{F}\mathcal{F}^{\ell,-1}\big\|_{\mathcal{L}(U)} + \big\|1 - \mathcal{F}^\star\mathcal{F}^{\ell,-\star}\big\|_{\mathcal{L}(U)} + \big\|\mathcal{B}^\star(\mathcal{A}_1 - \mathcal{A}_1^\ell)\big\|_{\mathcal{L}(U)}\right)\\
&\qquad\cdot \big\|\mathcal{F}^{-1}\big(\bar v - \bar v^\ell\big)\big\|_U
\end{aligned}$$

with $C_7 = \max(C_4, C_6)$. Hence, we have

$$\begin{aligned}
\|\mathcal{F}^{-1}(\bar v - \bar v^\ell)\|_U &\le \frac{C_8}{\sigma}\left(\|1 - \mathcal{F}\mathcal{F}^{\ell,-1}\|_{\mathcal{L}(U)} + \big\|1 - \mathcal{F}^\star\mathcal{F}^{\ell,-\star}\big\|_{\mathcal{L}(U)}\right)\\
&\quad+ \frac{C_8}{\sigma}\big\|\mathcal{B}^\star(\mathcal{A}_1 - \mathcal{A}_1^\ell)\big\|_{\mathcal{L}(U)}
\end{aligned}$$

with $C_8 = (C_3 + C_7)$. From Proposition 4.2 we infer that

$$\lim_{\ell\to\infty}\big\|\mathcal{B}^\star(\mathcal{A}_1 - \mathcal{A}_1^\ell)\big\|_{\mathcal{L}(U)} = 0$$

so that $\|\mathcal{F}^{-1}(\bar v - \bar v^\ell)\|_U \to \infty$ for $\ell \to \infty$. Since $\mathcal{F}$ is invertible, we conclude that

$$\lim_{\ell\to\infty}\|\bar v - \bar v^\ell\|_U = 0.$$

3) From $\bar{u} = \mathcal{F}^{-1}\bar{v}$, $\bar{u}^\ell = \mathcal{F}^{\ell,-1}\bar{v}^\ell$ and Theorem 4.1 we infer that

$$\begin{aligned}\|\bar{u} - \bar{u}^\ell\|_U &= \|\mathcal{F}^{-1}\bar{v} - \mathcal{F}^{\ell,-1}\bar{u}^\ell\|_U \\ &\le \|\mathcal{F}^{-1}\|_{\mathcal{L}(U)}\left(\|\bar{v} - \bar{v}^\ell\|_U + \|1 - \mathcal{F}\mathcal{F}^{\ell,-1}\|_{\mathcal{L}(U)}\|\bar{v}^\ell\|_U\right) \stackrel{\ell\to\infty}{\longrightarrow} 0.\end{aligned}$$

4) The first part of the claim follows directly from Theorem 3.1. We infer from part 1) and part 3) that $\{\mathcal{B}^\star\bar{q}^\ell\}_{\ell\in\mathbb{N}}$ and $\{\bar{u}^\ell\}_{\ell\in\mathbb{N}}$ converge to $\bar{u}$ respectively $\mathcal{B}^\star\bar{q}$. Hence, $\{\sigma\bar{u}^\ell - \mathcal{B}^\star\bar{q}^\ell\}_{\ell\in\mathbb{N}}$ tends to $\sigma\bar{u} - \mathcal{B}^\star\bar{q}$. Since $\mathcal{F}^{-\star}$ is bounded, we conclude that $\{\xi^\ell\}_{\ell\in\mathbb{N}}$ converge to $\xi = \mathcal{F}^{-\star}(\sigma\bar{u} - \mathcal{B}^\star\bar{q})$. Now, the proof follows by the same arguments as the proof of Theorem 4.11, part (2), in [21].

References

1. K. Afanasiev and M. Hinze. Adaptive control of a wake flow using proper orthogonal decomposition. *Lecture Notes in Pure and Applied Mathematics*, 216: 317-332, 2001.
2. E. Arian, M. Fahl, and E.W. Sachs. Trust-region proper orthogonal decomposition for flow control. Technical Report 2000-25, ICASE, 2000.
3. R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology. Volume 5: Evolution Problems I.* Springer, Berlin, 2000.
4. 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.
5. L.C. Evans. *Partial Differential Equations.* American Math. Society, Providence, Rhode Island, 2008.
6. M. Gubisch and S. Volkwein. Modified POD Galerkin schemes for linear-quadratic optimal control problems. In preparation, 2013
7. M. Hintermüller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semismooth Newton method. *SIAM Journal on Optimization*, 13:865-888, 2003.
8. M. Hintermüller, I. Kopacka, and S. Volkwein. Mesh-independence and preconditioning for solving control problems with mixed control-state constraints. *ESAIM: Control, Optimisation and Calculus of Variations*, 15:626-652, 2009.
9. M. Hinze and F. Tröltzsch. Discrete concepts versus error analysis in pde constrained optimization. *GAMM-Mitteilungen*, 33:148-162, 2010.
10. M. Hinze and S. Volkwein. Error estimates for abstract linear-quadratic optimal control problems using proper orthogonal decomposition. *Computational Optimization and Applications*, 39:319-345, 2008.
11. P. Holmes, J.L. Lumley, G. Berkooz, and C.W. Rowley. *Turbulence, Coherent Structures, Dynamical Systems and Symmetry.* Cambridge Monographs on Mechanics, Cambridge University Press, second edition, 2012.
12. 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.
13. E. Kammann, F. Tröltzsch and S. Volkwein. A method of a-posteriori error estimation with application to proper orthogonal decomposition. *ESAIM: Mathematical Modelling and Numerical Analysis*, to appear, 2013.
14. K. Kunisch and S. Volkwein. Galerkin proper orthogonal decomposition methods for parabolic problems. *Numerische Mathematik*, 90:117-148, 2001.
15. J.L. Lions. *Optimal Control of Systems Governed by Partial Differential Equations.* Springer, Berlin, 1971.
16. C. Meyer, U. Prüfert, and F. Tröltzsch. On two numerical methods for state-constrained elliptic control problems. *Optimization Methods and Software*, 22:871-899, 2007.
17. J.M. Ortega and W.C. Rheinboldt *Iterative Solution of Nonlinear Equations in Several Variables.* Computer Science and Applied Mathematics. Academic Press, 1970.
18. M. Reed and B. Simon. *Methods of Modern Mathematical Physics I: Functional Analysis.* Academic Press, New York, 1980.

19. A. Studinger and S. Volkwein. Numerical analysis of POD a-posteriori error estimation for optimal control To appear in *Control and Optimization with PDE Constraints*, International Series of Numerical Mathematics, Kunisch, Clason (eds), Birkhäuser, Basel, 2013.
20. F. Tröltzsch. *Optimal Control of Partial Differential Equations. Theory, Methods and Applications*. American Math. Society, Providence, volume 112, 2010.
21. F. Tröltzsch and S. Volkwein. POD a-posteriori error estimates for linear-quadratic optimal control problems. *Computational Optimization and Applications*, 44:83-115, 2009.
22. S. Volkwein. *Proper Orthogonal Decomposition: Theory and Reduced-Order Modelling.* Lecture Notes, University of Konstanz, 2012.