# POD A-POSTERIORI ERROR BASED INEXACT SQP METHOD FOR BILINEAR ELLIPTIC OPTIMAL CONTROL PROBLEMS\*

# Martin Kahlbacher<sup>1</sup> and Stefan Volkwein<sup>2</sup>

**Abstract.** An optimal control problem governed by a bilinear elliptic equation is considered. This problem is solved by the sequential quadratic programming (SQP) method in an infinite-dimensional framework. In each level of this iterative method the solution of linear-quadratic subproblem is computed by a Galerkin projection using proper orthogonal decomposition (POD). Thus, an approximate (inexact) solution of the subproblem is determined. Based on a POD *a-posteriori* error estimator developed by Tröltzsch and Volkwein [*Comput. Opt. Appl.* **44** (2009) 83–115] the difference of the suboptimal to the (unknown) optimal solution of the linear-quadratic subproblem is estimated. Hence, the inexactness of the discrete solution is controlled in such a way that locally superlinear or even quadratic rate of convergence of the SQP is ensured. Numerical examples illustrate the efficiency for the proposed approach.

Mathematics Subject Classification. 35J47, 49K20, 49M15, 90C20.

Received March 23, 2010. Revised May 24, 2011. Published online December 19, 2011.

# 1. INTRODUCTION

Optimal control problems governed by partial differential equations (PDEs) can often be formulated as an infinite-dimensional optimization problem in the following form (see, e.g., in [23]):

$$\min_{x \in X} J(x) \quad \text{subject to (s.t.)} \quad e(x) = 0. \tag{1.1}$$

The mapping  $J: X \to \mathbb{R}$  denotes the cost functional with a Banach space X. The operator  $e: X \to Y'$  describes the partial differential equations with a Banach space Y and its dual Y'. The Lagrangian for (1.1) is given by

$$L(x,p) = J(x) + \langle e(x), p \rangle_{Y'|Y} \quad \text{for } (x,p) \in X \times Y,$$

where  $\langle \cdot, \cdot \rangle_{Y',Y}$  denotes the dual pairing between Y' and Y. If J and e are twice continuously Fréchetdifferentiable, second-order methods can be applied to solve (1.1) numerically. One favorite method is the

Keywords and phrases. Optimal control, inexact SQP method, proper orthogonal decomposition, *a-posteriori* error estimates, bilinear elliptic equation.

<sup>\*</sup> The authors gratefully acknowledge support by the Austrian Science Fund FWF under grant No. P19588-N18 and by the SFB Research Center "Mathematical Optimization in Biomedical Sciences" (SFB F32).

<sup>&</sup>lt;sup>1</sup> Universität Graz, Institut für Mathematik und Wissenschaftliches Rechnen, Heinrichstraße 36, 8010 Graz, Austria. martin.kahlbacher@hotmail.com

 $<sup>^2</sup>$ Universität Konstanz, Fachbereich Mathematik und Statistik, Universitätsstraße 10, 78457 Konstanz, Germany. Stefan.Volkwein@uni-konstanz.de

sequential quadratic programming (SQP) method, where in each level of the iteration the linear-quadratic programming problem

$$\begin{cases} \min_{x \in X} L_x(x^k, p^k)x + \frac{1}{2} L_{xx}(x^k, p^k)(x, x) \\ \text{s.t. } e(x^k) + e'(x^k)x = 0 \end{cases}$$
(1.2)

is solved. The solution  $\bar{x}$  to (1.2) is given by the solution to the Karush-Kuhn-Tucker (KKT) system

$$A_k \bar{z} = b_k \quad \text{in } X' \times Y' \tag{1.3}$$

with

$$A_k = \begin{pmatrix} L_{xx}(x^k, p^k) \ e'(x^k)^{\star} \\ e'(x^k) \ 0 \end{pmatrix}, \ \bar{z} = \begin{pmatrix} \bar{x} \\ \bar{p} \end{pmatrix}, \ b_k = -\begin{pmatrix} L_x(x^k, p^k) \\ e(x^k) \end{pmatrix}$$

Here,  $X' \times Y'$  is identified with the dual of  $X \times Y$ ,  $e'(x^k)^* : Y \to X'$  is the dual operator of the Fréchet derivative  $e'(x^k) : X \to Y'$  and  $L_x(L_{xx})$  stands for the first (second) Fréchet derivative of the Lagrangian with respect to x.

In the context of PDE constrained optimization (1.3) has to be discretized. Often that leads to very large scale linear systems. Therefore, different techniques of model order reduction methods have been developed to approximate (1.3) by smaller ones that are tractable with less effort. We apply the method of proper orthogonal decomposition (POD), which is based on projecting the system onto subspaces consisting of  $\ell \geq 1$  POD basis elements that contain characteristics of the expected solution; see, *e.g.*, [4, 5, 15, 18, 21]. This is in contrast to, *e.g.*, finite element techniques, where the elements of the subspaces are uncorrelated to the physical properties of the system that they approximate. The discretization of (1.3) leads to a discrete solution which solves (1.3) inexactly. Thus, we obtain an inexact version of the SQP method. Utilizing the convergence theory for inexact Newton methods (see, *e.g.*, [7]) the inexactness can be controlled in such a way that a local superlinear or even local quadratic rate of convergence can be ensured.

Utilizing  $\ell$  POD basis functions for the Galerkin projection of (1.3) we arrive at a finite- and low-dimensional linear system

$$A_k^{\ell} \bar{z}^{\ell} = b_k^{\ell} \quad \text{in } \mathbb{R}^n \tag{1.4}$$

with an integer  $n = n(\ell)$  depending on the number  $\ell$  of POD basis functions. We prolongate the solution  $\bar{z}^{\ell}$  to (1.4) into the space  $X \times Y$  by applying a linear operator  $\mathcal{I} : \mathbb{R}^n \to X \times Y$ . Convergence of the SQP method can be ensured provided the starting value  $(x^0, p^0)$  is appropriately chosen and

$$\|A_k(\mathcal{I}\bar{z}^\ell) - b_k\|_{X' \times Y'} = O\left(\|L'(x^k, p^k)\|_{X' \times Y'}^q\right)$$
(1.5)

with  $q \in [1, 2]$ . Here, L' denotes the Fréchet derivative of the Lagrangian with respect to (x, p). If q = 1 holds, then the iterates converge linearly, if  $q \in (1, 2)$  is satisfied, the rate of convergence is superlinear, and for q = 2 we obtain quadratic rate of convergence. To achieve (1.5) we apply a POD *a-posteriori* error estimator (see [24]) which is derived for linear-quadratic programming problems. Utilizing the quadratic convergence of the SQP method in function spaces we ensure convergence of the iterates – computed by the POD suboptimal control approach – to the solution of the nonlinear optimization problem (1.1).

For the POD method (and also for other model reduction methods like the reduced-basis method [17] and balanced truncation [2, 6]) no reliable *a-priori* error analysis for nonlinear optimal control problems black are available. A priori error estimates for POD Galerkin approximations of linear-quadratic optimal control problems were derived in [12], where the POD basis was computed with the *knowledge* of the optimal solution. In [24] the main focus was on a POD *a-posteriori* analysis for linear-quadratic optimal control problems. It was deduced how far the suboptimal control, computed on the basis of the POD model, is from the (*unknown*) exact one. We use this idea for nonlinear optimal control problems so that we are able to compensate for the lack of *a priori* analysis for POD methods.

In our work we apply the technique developed in [24] to control the discretization error of the POD Galerkin approximation in each level of the SQP method. The approach is illustrated for an optimal control problem governed by a bilinear elliptic partial differential equation. Within the inexact SQP method we tune the number  $\ell$  of basis functions for the POD Galerkin approximation to ensure the locally fast convergence of the algorithm. Thus, in contrast to [14] the POD basis will be fixed during the numerical algorithm. Only the number of the utilized POD ansatz functions is increased, if necessary. We refer to the papers [13, 25], where also bilinear optimal control problems are considered. Let us mention that the presented approach can also be used for nonlinear parabolic equations as well as for reduced-basis approximations; see [22].

The paper is organized in the following manner: in Section 2 the optimal control problem is introduced and optimality conditions are discussed. The SQP method is formulated in Section 3. In Section 4 we turn to the POD discretization of the linear-quadratic subproblem. The inexact SQP method is studied in Section 5. Two numerical examples are presented in Section 6. Finally, two proofs are given in the appendix.

# 2. Optimal control of the bilinear equation

In this section we introduce the optimal control problem. In Section 2.1 we discuss the underlying state equation. The optimal control problem is investigated in Section 2.2, and optimality conditions are presented in Section 2.3.

## 2.1. The state equation

Throughout we suppose that  $\Omega \subset \mathbb{R}^d$ ,  $d \in \{1, 2, 3\}$ , is an open and bounded domain with a smooth boundary  $\partial \Omega = \Gamma$  ensuring the needed Sobolev embeddings. Let  $L^2(\Omega)$  denote the Lebesgue space of all measurable and square integrable functions on  $\Omega$ . For brevity, we set  $V = H^1(\Omega)$  and refer to [8], for instance, for more details on Lebesgue and Sobolev spaces. Recall that V is continuously embedded into  $L^4(\Omega)$  for  $d \leq 3$ . The bilinear elliptic equation is given by

$$-\Delta y(\mathbf{x}) + u(\mathbf{x})y(\mathbf{x}) = f(\mathbf{x}) \qquad \text{for all } \mathbf{x} \in \Omega,$$
(2.1a)

$$\frac{\partial y}{\partial n}(\mathbf{s}) + y(\mathbf{s}) = 0$$
 for all  $\mathbf{s} \in \Gamma$ . (2.1b)

We assume that f belongs to  $L^2(\Omega)$  and the control variable u is of the form

$$u(\mathbf{x}) = \sum_{i=1}^{N} u_i b_i(\mathbf{x}) \text{ for all } \mathbf{x} \in \Omega,$$

where  $b_1, \ldots, b_N$  are linearly independent in  $L^2(\Omega)$ . For instance, the  $b_i$ 's can be step functions satisfying  $b_i \equiv 1$ on  $\Omega_i$  and  $b_i \equiv 0$  on  $\Omega \setminus \Omega_i$  for  $i = 1, \ldots, N$  and  $N = n_{\Omega}$ .

**Remark 2.1.** Let us mention that (2.1) is a simplify model for an identification problem arising in hyperthermia; see [10].

We define the finite-dimensional control space

$$U = \operatorname{span} \left\{ b_1, \dots, b_N \right\} \subset L^2(\Omega)$$

supplied with the topology in  $L^2(\Omega)$ . Note that dim U = N. Let us introduce the Hilbert space

$$X = V \times U$$

endowed with the common product topology. To write the elliptic differential equation (2.1) in a compact form we define the bilinear operator  $e: X \to V'$  by

$$\langle e(x), \varphi \rangle_{V', V} = \int_{\Omega} \nabla y \cdot \nabla \varphi + (uy - f) \varphi \, \mathrm{d}\mathbf{x} + \int_{\Gamma} y \varphi \, \mathrm{d}\mathbf{s}$$

for  $x = (y, u) \in X$  and  $\varphi \in V$ . Moreover,  $\langle \cdot, \cdot \rangle_{V',V}$  denotes the dual pairing associated with V and its dual V'. Moreover,  $u \in U$  holds. Thus, the operator e and its Fréchet-derivatives are well-defined. In particular, at  $x = (u, w) \in X$  we have

$$\langle e'(x)x_{\delta},\varphi\rangle_{V',V} = \int_{\Omega} \nabla y_{\delta} \cdot \nabla \varphi + (u_{\delta}y + uy_{\delta})\varphi \,\mathrm{d}\mathbf{x} + \int_{\Gamma} y_{\delta}\varphi \,\mathrm{d}\mathbf{s},$$
$$\langle e''(x)(x_{\delta},\tilde{x}_{\delta}),\varphi\rangle_{V',V} = \int_{\Omega} (u_{\delta}\tilde{y}_{\delta} + \tilde{u}_{\delta}y_{\delta})\varphi \,\mathrm{d}\mathbf{x}$$

in directions  $x_{\delta} = (y_{\delta}, u_{\delta}), \tilde{x}_{\delta} = (\tilde{y}_{\delta}, \tilde{u}_{\delta}) \in X$  and for  $\varphi \in V$ . Due to the bilinear structure of the mapping e the mapping  $x \mapsto e''(x)$  does not depend on  $x \in X$  so that it is Lipschitz-continuous on X.

The next proposition ensures existence and uniqueness of a weak solution to the state equation for arbitrary non-negative  $u \in U$ . For a proof we refer to [10], Theorem 2.1.

**Proposition 2.2.** For every  $u \in U$  with  $u \ge 0$  in  $\Omega$  there exists a unique solution  $y = y(u) \in V$  of the equation e(y, u) = 0. Moreover, y is uniformly bounded in V with respect to u.

The following result ensures a standard constraint qualification that is needed to ensure the existence of Lagrange multipliers. For the proof we refer to [10], Theorem 2.2.

**Proposition 2.3.** For every  $x = (y, u) \in X$  with  $u \ge 0$  in  $\Omega$ , the Fréchet derivative  $e_y(x) : V \to V'$  of the operator e with respect to y is bijective. In particular, e'(x) is surjective, and there exists a constant  $C_{ker} > 0$  such that

$$\|y_{\delta}\|_{V} \leq C_{ker} \|u_{\delta}\|_{L^{2}(\Omega)}$$
 for all  $(y_{\delta}, u_{\delta}) \in \ker e'(x) \subset X$ .

### 2.2. The optimal control problem

Motivated by Propositions 2.2 and 2.3 we define the set of admissible nonnegative control functions by

$$U_{ad} = \left\{ u \in U \, \big| \, u(\mathbf{x}) \ge u_a \text{ for all } \mathbf{x} \in \Omega \right\} \subset L^2(\Omega),$$

where  $u_a$  is a nonnegative real number. We set  $X_{ad} = V \times U_{ad}$  and introduce a cost functional  $J : X \to \mathbb{R}$  of tracking type

$$J(x) = \frac{1}{2} \|y - y_d\|_{L^2(\Omega)}^2 + \frac{\sigma}{2} \|u\|_{L^2(\Omega)}^2 \quad \text{for } x = (y, u) \in X,$$

where  $y_d \in L^2(\Omega)$  is a desired state, and  $\sigma > 0$  denotes a regularization parameter.

**Remark 2.4.** Let  $\Omega_m$  be a subset of  $\Omega$  and  $u^{\circ} \in U$  arbitrarily chosen. In our numerical examples we consider the more general cost functional

$$J(x) = \frac{1}{2} \|y - y_d\|_{L^2(\Omega_m)}^2 + \frac{\sigma}{2} \|u - u^\circ\|_{L^2(\Omega)}^2 \quad \text{for } x = (y, u) \in X,$$

which does not effect significantly the analysis of the optimal control problem.

It follows by standard arguments that J is twice continuously Fréchet-differentiable and the mapping  $x \mapsto J''(x)$  is Lipschitz-continuous on X. In particular, the first and second derivatives at  $x = (y, u) \in X$  are

$$J'(x)x_{\delta} = \int_{\Omega} (y - y_d)y_{\delta} + \sigma u u_{\delta} \,\mathrm{d}\mathbf{x}, \quad J''(x)(x_{\delta}, \tilde{x}_{\delta}) = \int_{\Omega} y_{\delta} \tilde{y}_{\delta} + \sigma u_{\delta} \tilde{u}_{\delta} \,\mathrm{d}\mathbf{x}$$

for directions  $x_{\delta} = (y_{\delta}, u_{\delta})$  and  $\tilde{x}_{\delta} = (\tilde{y}_{\delta}, \tilde{u}_{\delta})$ .

Then, the optimal control problem is given by

$$\min J(x) \quad \text{s.t.} \quad x \in \mathcal{F}(\mathbf{P}),\tag{P}$$

where the feasible set is  $\mathcal{F}(\mathbf{P}) = \{x \in X_{ad} | e(x) = 0 \text{ in } V'\}$ . Since  $U_{ad} \neq \emptyset$  holds, it follows by standard arguments that there exists at least one optimal solution  $x^* = (y^*, u^*)$  to  $(\mathbf{P})$ .

# 2.3. Optimality conditions

Let us introduce the Lagrange functional  $L: X \times V \to \mathbb{R}$  associated with (**P**):

$$L(x,p) = J(x) + \langle e(x), p \rangle_{V',V} \quad \text{for } (x,p) \in X \times V.$$

It follows from the properties of J and e that the Lagrange functional is twice continuously Fréchet-differentiable and the mapping  $(x, p) \mapsto L''(x, p)$  is Lipschitz-continuous on X.

In the following theorem we state first-order necessary optimality conditions for  $(\mathbf{P})$ . The existence of a unique Lagrange multiplier is shown in [10].

**Theorem 2.5** (first-order necessary optimality conditions). Suppose that  $x^* = (y^*, u^*)$  is a local solution to (**P**). Then there exists a unique Lagrange multiplier  $p^* \in V$  satisfying together with  $x^*$  the dual equation

$$-\Delta p^* + u^* p^* = y_d - y^* \text{ on } \Omega, \quad \frac{\partial p^*}{\partial n} + p^* = 0 \text{ on } \Gamma.$$

$$(2.2)$$

Furthermore, the variational inequality

$$\int_{\Omega} \left( \sigma u^* + y^* p^* \right) \left( u - u^* \right) \mathrm{d}\mathbf{x} \ge 0 \quad \text{for all } u \in U_{ad}$$

holds.

For the convergence of the SQP method second-order sufficient optimality conditions are required, at least in a neighborhood of the solution  $x^* = (y^*, u^*)$ . The second Fréchet-derivative of the Lagrangian at  $(x^*, p^*) \in X_{ad} \times V$  with respect to x in the direction  $x = (y, u) \in X$  is

$$L_{xx}(x^*, p^*)(x, x) = \int_{\Omega} y^2 + \sigma u^2 + 2uyp^* \, \mathrm{d}\mathbf{x} \ge \sigma \, \|u\|_{L^2(\Omega)}^2 + 2 \int_{\Omega} uyp^* \, \mathrm{d}\mathbf{x}.$$

Since V is continuously embedded into  $L^4(\Omega)$  there exists a constant  $C_{emb} > 0$  such that

$$\|\varphi\|_{L^4(\Omega)} \le C_{emb} \, \|\varphi\|_V \quad \text{for all } \varphi \in V.$$

$$(2.3)$$

Due to Proposition 2.3 we also have  $||y||_V \leq C_{ker} ||u||_{L^2(\Omega)}$  for all  $(y, u) \in \ker e'(x^*)$  with a constant  $C_{ker} > 0$ . We set  $C = C_{emb}C_{ker}$  and derive

$$L_{xx}(x^*, p^*)(x, x) \ge \frac{\sigma}{2} \|u\|_{L^2(\Omega)}^2 + \frac{\sigma}{2} \|u\|_{L^2(\Omega)}^2 - 2 \|u\|_{L^2(\Omega)} \|y\|_{L^4(\Omega)} \|p^*\|_{L^4(\Omega)}$$
$$\ge \frac{\sigma}{2} \|u\|_{L^2(\Omega)}^2 + \frac{\sigma}{2C_{ker}} \|y\|_V^2 - 2C \|u\|_{L^2(\Omega)}^2 \|p^*\|_{L^4(\Omega)}$$
$$\ge \min\left(\frac{\sigma}{4}, \frac{\sigma}{2C_{ker}}\right) \|x\|_X^2 + \frac{1}{4} \left(\sigma - 8C \|p^*\|_{L^4(\Omega)}\right) \|u\|_{L^2(\Omega)}^2$$

for all  $x = (y, u) \in \ker e'(x^*)$ . Thus, we have proved the following result.

**Theorem 2.6** (second-order sufficient optimality conditions). Suppose that  $x^* = (y^*, u^*)$  is a local solution to (**P**) and  $p^* \in V$  is the associated unique Lagrange multiplier. Let the constants  $C_{emb}$  and  $C_{ker}$  be given by (2.3) and Proposition 2.3, respectively. If

$$\sigma - 8C_{emb}C_{ker} \|p^*\|_{L^4(\Omega)} \ge 0 \tag{2.4}$$

holds, the second-order sufficient optimality condition is satisfied at  $(x^*, p^*)$ , i.e., there exists a  $\gamma > 0$  so that

$$L_{xx}(x^*, p^*)(x_{\delta}, x_{\delta}) \ge \gamma \|x_{\delta}\|_X^2 \quad \text{for all } x_{\delta} = (y_{\delta}, u_{\delta}) \in \ker e'(x^*).$$

Note that (2.4) can be ensured provided the Lagrange multiplier satisfies

$$\|p^*\|_{L^4(\Omega)} \le \frac{\sigma}{8C_{emb}C_{ker}}.$$
(2.5)

**Remark 2.7.** It follows from standard arguments that (2.5) holds if the residuum  $\|y^* - y_d\|_{L^2(\Omega)}$  is sufficiently small.

## 3. The inexact sop method

In this section we formulate the SQP method for  $(\mathbf{P})$ . Moreover, the *a-posteriori* error estimator for the linear-quadratic subproblems are introduced.

## 3.1. The SQP method

To solve  $(\mathbf{P})$  numerically, we apply the SQP method. The principal idea is to replace J and e by a quadratic approximation of the Lagrangian and a linearization of the constraint. For the readers convenience we recall the SQP method in Algorithm 1.

## Algorithm 1 (Lagrange-SQP method)

1: Choose  $x^0 = (y^0, u^0) \in X_{ad}, p^0 \in V, \mu > 0$ , and set k = 1.

2: repeat

Compute  $J'(x^k)$ ,  $L_{xx}(x^k, p^k)$ ,  $e(x^k)$ , and  $e'(x^k)$ . 3:

4: Solve the linear-quadratic minimization problem

$$\min_{x \in X} J^{k}(x) = J'(x^{k})x + \frac{1}{2}L_{xx}(x^{k}, p^{k})(x, x)$$
  
s.t.  $e'(x^{k})x + e(x^{k}) = 0$  and  $x^{k} + x \in X_{ad}$ . (**P**<sup>k</sup>)

- Determine a step length parameter  $t_k \in (0,1]$  by an Armijo backtracking line search for the  $\ell_1$  merit function 5:
  $$\begin{split} \varPhi(x;\mu) &= J(x) + \mu \, \|e(x)\|_{V'} \text{ (see, e.g., [11]).} \\ \text{Set } x^{k+1} &= x^k + t_k x \in X_{ad} \text{ and } k = k+1. \\ \text{Choose a new estimate } p^k \text{ for the Lagrange multiplier.} \end{split}$$
- 6:
- 7:

8: until a given stopping criterium is satisfied.

## Remark 3.1.

- (a) Two choices for the update of the Lagrange multiplier in step 7 are the Lipschitz-continuous or a Newton update; see, e.g., in [20,26]. If  $e'(x^*)$  is surjective and  $L_{xx}(x^*, p^*)$  is coercive on ker  $e'(x^*)$  one can prove that Algorithm 1 has a locally quadratic rate of convergence if J and e have Lipschitz-continuous second Fréchet derivatives:
- (b) the linear-quadratic minimization problem ( $\mathbf{P}^k$ ) is well-defined provided the operator  $L_{rr}(x^k, p^k)$  is coercive on ker  $e'(x^k)$  and  $e'(x^k)$  is surjective. Thus, Algorithm 1 is not globally convergent;
- (c) by Proposition 2.3 the operator e'(x) is surjective for all  $x \in X_{ad}$ . However,  $L_{xx}(x^k, p^k)$  need not to be coercive on ker  $e'(x^k)$ ; compare Remark 2.7. To ensure that  $(\mathbf{P}^k)$  has a unique solution we modify  $L_{xx}(x^k, p^k)$ in the case if coercivity does not hold. For  $\beta \in [0,1]$  let the bilinear operator  $B^{k,\beta}: X \times X \to \mathbb{R}$  be given by

$$B^{k,\beta}(x,\tilde{x}) = J''(x^k)(x,\tilde{x}) + \beta \left\langle e''(x^k)(x,\tilde{x}), p^k \right\rangle_{V',V} \quad \text{for } x, \tilde{x} \in X$$

Then,  $B^{k,1} = L_{xx}(x^k, p^k)$  and  $B^{k,0} = J''(x^k)$ . Due to Proposition 2.3 we have

$$B^{k,0}(x,x) \ge \frac{\sigma}{2} \min\left(\frac{1}{C_{ker}}, 1\right) \|x\|_X^2 \quad \text{for all } x \in \ker e'(x^k),$$

*i.e.*,  $B^{k,0}$  is positive definite. Thus, in the case if coercivity does not hold, we replace  $(\mathbf{P}^k)$  by

$$\min_{x \in X} J^{k,\beta}(x) = J'(x^k)x + \frac{1}{2} B^{k,\beta}(x,x)$$
  
s.t.  $e'(x^k)x + e(x^k) = 0$  and  $x^k + x \in X_{ad}$  ( $\mathbf{P}^{k,\beta}$ )

with a coercive operator  $B^{k,\beta}$  (e.g., with  $\beta = 0$ ).

Next we derive the optimality conditions for the linear-quadratic subproblem  $(\mathbf{P}^{k,\beta})$ . Throughout we suppose that the parameter  $\beta \in [0,1]$  is chosen in such a way that  $B^{k,\beta}$  is coercive on  $X \times X$ , *i.e.*,  $(\mathbf{P}^{k,\beta})$  has a unique solution. For our problem the cost  $J^{k,\beta}$  in  $(\mathbf{P}^{k,\beta})$  has the form

$$J^{k,\beta}(x) = \int_{\Omega} (y^k - y_d)y + \sigma u^k u + \frac{1}{2} \left(y^2 + 2\beta uyp^k + \sigma u^2\right) \,\mathrm{d}\mathbf{x}$$

for  $x = (y, u) \in X$ . The equation  $e'(x^k)x + e(x^k) = 0$  is equivalent with the fact that x = (y, u) satisfies the linearized state equation

$$\int_{\Omega} \nabla y \cdot \nabla \varphi + \left( u^{k} y + u y^{k} \right) \varphi \, \mathrm{d}\mathbf{x} + \int_{\Gamma} y \varphi \, \mathrm{d}\mathbf{s} = -\langle e(x^{k}), \varphi \rangle_{V', V}$$

for all  $\varphi \in V$ . To obtain  $x^k + x \in X_{ad}$  we have to ensure that  $u^k + u \in U_{ad}$ . Setting  $u_a^k = u_a - u^k$  we require

$$u \in U_{ad}^k = \big\{ \tilde{u} \in U \, \big| \, \tilde{u} \ge u_a^k \text{ in } \Omega \big\}.$$

**Remark 3.2.** We introduce the linear operator  $S : U \to V$  as follows: for  $u \in U$  the function y = Su is the unique solution to

$$\int_{\Omega} \nabla y \cdot \nabla \varphi + u^{k} y \varphi \, \mathrm{d}\mathbf{x} + \int_{\Gamma} y \varphi \, \mathrm{d}\mathbf{s} = -\int_{\Omega} u y^{k} \varphi \, \mathrm{d}\mathbf{x} \quad \text{for all } \varphi \in V.$$
(3.1)

Since  $u^k \in U_{ad}$  holds, it follows from the Lax-Milgram lemma that S is well-defined and bounded. Moreover,  $\hat{y}_k \in V$  is the unique solution to

$$\int_{\Omega} \nabla \hat{y}_k \cdot \nabla \varphi + u^k \hat{y}_k \varphi \, \mathrm{d}\mathbf{x} + \int_{\Gamma} \hat{y}_k \varphi \, \mathrm{d}\mathbf{s} = -\langle e(x^k), \varphi \rangle_{V', V}$$

Then, x = (y, u) with  $y = \hat{y}_k + Su$  solves  $e'(x^k)x + e(x^k) = 0$ . The adjoint operator  $S^* : V' \to U'$  of S is given as follows [23]: for arbitrary  $r \in V'$  compute the solution  $v \in V$  to the variational problem

$$\int_{\Omega} \nabla v \cdot \nabla \varphi + u^k v \varphi \, \mathrm{d}\mathbf{x} + \int_{\Gamma} v \varphi \, \mathrm{d}\mathbf{s} = \langle r, \varphi \rangle_{V', V} \quad \text{for all } \varphi \in V$$
(3.2)

and set  $\mathcal{S}^* r = -y^k v$ . In particular,  $\mathcal{S}^* r \in L^2(\Omega)$ .

Suppose that there is a unique solution  $\bar{x} = (\bar{y}, \bar{u})$  to  $(\mathbf{P}^k)$ . To derive the optimality conditions, we define the Lagrangian functional  $L^{k,\beta} : X \times V \to \mathbb{R}$  associated with  $(\mathbf{P}^k)$  by

$$L^{k,\beta}(x,p) = J^{k,\beta}(x) + \langle e'(x^k)x + e(x^k), p \rangle_{V',V} \quad \text{for } (x,p) \in X \times V.$$

From  $L_p^{k,\beta}(\bar{x},\bar{p})p = 0$  for all  $p \in V$  we infer that the pair  $(\bar{y},\bar{u})$  solves (3.3a). The equation  $L_y^{k,\beta}(\bar{x},\bar{p})y = 0$  for all  $y \in V$  implies

$$\int_{\Omega} \nabla y \cdot \nabla \bar{p} + u^k y \bar{p} \, \mathrm{d}\mathbf{x} + \int_{\Gamma} y \bar{p} \, \mathrm{d}\mathbf{s} = -\int_{\Omega} \left( y^k + \bar{y} - y_d + \beta \bar{u} p^k \right) y \, \mathrm{d}\mathbf{x}.$$

Thus,  $\bar{p}$  satisfies the dual problem

$$-\Delta \bar{p} + u^k \bar{p} = y_d - y^k - \bar{y} - \beta \bar{u} p^k \text{ in } \Omega, \quad \frac{\partial \bar{p}}{\partial n} + \bar{p} = 0 \text{ on } \Gamma.$$

Finally, the optimality condition  $L_u^{k,\beta}(\bar{x},\bar{p})(u-\bar{u}) \ge 0$  for all  $u \in U_{ad}^k$  implies:

$$\int_{\Omega} \left( \beta \bar{y} p^k + \sigma (u^k + \bar{u}) + y^k \bar{p} \right) (u - \bar{u}) \, \mathrm{d}\mathbf{x} \ge 0$$

Summarizing, the solution  $\bar{x} = (\bar{y}, \bar{u})$  to  $(\mathbf{P}^{k,\beta})$  satisfies together with the Lagrange multiplier  $\bar{p} \in V$  the following optimality system:

(1) The (linearized) state equation

$$\int_{\Omega} \nabla \bar{y} \cdot \nabla \varphi + \left( u^{k} \bar{y} + \bar{u} y^{k} \right) \varphi \, \mathrm{d}\mathbf{x} + \int_{\Gamma} \bar{y} \varphi \, \mathrm{d}\mathbf{s} = -\langle e(x^{k}), \varphi \rangle_{V', V}$$
(3.3a)

for all  $\varphi \in V$ ;

(2) the (linearized) dual equation

$$\int_{\Omega} \nabla \bar{p} \cdot \nabla \varphi + u^k \bar{p} \varphi \, \mathrm{d}\mathbf{x} + \int_{\Gamma} \bar{p} \varphi \, \mathrm{d}\mathbf{s} = \int_{\Omega} \left( y_d - y^k - \bar{y} - \beta \bar{u} p^k \right) \varphi \, \mathrm{d}\mathbf{x}$$
(3.3b)

for all  $\varphi \in V$ ; and

(3) the (linearized) variational inequality

$$\int_{\Omega} \left( \beta \bar{y} p^k + \sigma (u^k + \bar{u}) + y^k \bar{p} \right) (u - \bar{u}) \, \mathrm{d}\mathbf{x} \ge 0 \tag{3.3c}$$

for all  $u \in U_{ad}^k$ .

Recall that  $\beta \in [0,1]$  is chosen in such a way that (3.3) has a unique solution  $(\bar{y}, \bar{u}, \bar{p}) \in V \times U_{ad}^k \times V$ .

# **3.2.** A-posteriori error analysis for $(\mathbf{P}^{k,\beta})$

Utilizing an *a-posteriori* error analysis we can ensure that  $(\mathbf{P}^{k,\beta})$  is solved with a given tolerance. Therefore, we consider an inexact version of Algorithm 1, where the inexactness arises due to the inexact solution of the optimality system (3.3). Within the SQP method we control the error tolerance for the POD discretization to guarantee the overall convergence of the optimization method. The presented approach is not limited to POD model reduction, but can easily be applied to other reduced-order techniques, *e.g.*, to the reduced-basis method. We refer to [22] as a first step in this direction.

The idea of *a-posteriori* error estimates was used by Malanowski *et al.* [16] in the context of error estimates for the optimal control of ODEs. It was extended later to elliptic optimal control problems in [3]. Let us explain this basic idea for our application.

Let  $u^p = \sum_{i=1}^N u_i^p b_i \in U_{ad}^k$  be chosen arbitrarily. Our goal is to estimate the difference

$$\|\bar{u}-u^p\|_{L^2(\Omega)}$$

without the knowledge of the optimal solution  $(\bar{y}, \bar{u}, \bar{p})$  to (3.3). If  $u^p \neq \bar{u}$  then  $u^p$  does not satisfy the necessary (and by convexity sufficient) optimality conditions (3.3c). However, there exists a (unique) function  $\zeta \in L^2(\Omega)$  such that

$$\int_{\Omega} \left( \beta y^p p^k + \sigma(u^k + u^p) + y^k p^p + \zeta \right) (u - u^p) \, \mathrm{d}\mathbf{x} \ge 0 \quad \text{for all } u \in U^k_{ad}, \tag{3.4}$$

where  $y^p$  and  $p^p$  solve (3.3a) and (3.3b), respectively, with  $(\bar{y}, \bar{u}, \bar{p})$  replaced by  $(y^p, u^p, p^p)$ . Therefore,  $u^p$  satisfies the optimality condition of a perturbed elliptic optimal control problem with 'perturbation'  $\zeta$ :

$$\min_{x=(y,u)} J^k(x) = J'(x^k)x + \frac{1}{2} B^{k,\beta}(x,x) + \int_{\Omega} \zeta u \, \mathrm{d}\mathbf{x}$$
  
s.t.  $e'(x^k)x + e(x^k) = 0$  and  $x^k + x \in X_{ad}$ .

#### Remark 3.3.

- (a) The variable  $\zeta$  measures the violation of the optimality conditions. The computation of  $\zeta$  is possible on the basis of the known data  $u^p$ ,  $y^p$ , and  $p^p$ ;
- (b) the smaller  $\zeta$  is, the closer  $u^p$  is to  $\bar{u}$ . Up to now it is not clear that  $\|\zeta\|_{L^2(\Omega)}$  can be made small. We will address this issue in Theorem 4.3 for POD approximations;
- (c) if the sequence  $\{(x^k, p^k)\}_{k \in \mathbb{N}}$  is uniformly bounded in  $X \times V$ , then there exists a constant  $C_p > 0$  which is independent on  $(x^k, p^k)$  so that

$$\|(\bar{y},\bar{p}) - (y^p,p^p)\|_{V \times V} \le C_p \|\bar{u} - u^p\|_U$$

holds true.

We proceed by deriving an estimate for  $\|\bar{u}-u^p\|_{L^2(\Omega)}$  in terms of  $\|\zeta\|_{L^2(\Omega)}$ . The proof is given in the appendix. The proof is based on the same methodology as the one for the Falk lemma for variational inequalities; see, *e.g.*, [9].

**Theorem 3.4.** Let  $(\bar{y}, \bar{u}, \bar{p})$  be the solution to (3.3) and  $u^p \in U^k_{ad}$  be chosen arbitrarily. Then, it follows that

$$\|\bar{u} - u^p\|_{L^2(\Omega)} \le \frac{1}{\sigma} \|\zeta\|_{L^2(\Omega)}$$

where  $\zeta$  is chosen such that (3.4) holds.

**Remark 3.5** (see [24]). We introduce  $\mathcal{G}^{k,\beta} : X \times V \to L^2(\Omega)$  by

$$\mathcal{G}^{k,\beta}(x,p) = \beta p^k y + y^k p + \sigma \left( u^k + u \right) \quad \text{for } x = (y,u) \in X \text{ and } p \in V.$$

Then, (3.4) can be expressed as

$$\int_{\Omega} \left( \mathcal{G}^{k,\beta}(x^p, p^p) + \zeta \right) (u - u^p) \, \mathrm{d}\mathbf{x} \ge 0 \quad \text{for all } u \in U^k_{ad}.$$

Define  $\zeta \in L^2(\Omega)$  as follows

$$\zeta(\mathbf{x}) = \begin{cases} \left[ \mathcal{G}^{k,\beta}(y^p, u^p, p^p)(\mathbf{x}) \right]_{-} & \text{for all } \mathbf{x} \in \mathcal{A}^k = \left\{ \mathbf{x} \in \Omega \,|\, u^p(\mathbf{x}) = u_a^k(\mathbf{x}) \right\}, \\ - \,\mathcal{G}^{k,\beta}(y^p, u^p, p^p)(\mathbf{x}) & \text{for all } \mathbf{x} \in \Omega \setminus \mathcal{A}^k, \end{cases}$$

where  $[s]_{-} = -\min(0, s)$  for  $s \in \mathbb{R}$ . Then the estimate

$$\|\bar{u} - u^p\|_{L^2(\Omega)} \le \frac{1}{\sigma} \|\zeta\|_{L^2(\Omega)}$$
 (3.5)

holds true.

We call (3.5) an *a-posteriori* error estimate, since, in the next section, we shall apply it to suboptimal solutions  $u^p$  to the optimality system (3.3) that have already be computed by a POD Galerkin method. After having computed  $u^p$ , we determine the associated state  $y^p$  and adjoint state  $p^p$ . Then we can determine  $\zeta$  and its  $L^2$ -norm and (3.5) gives an upper bound for the distance of  $u^p$  to  $\bar{u}$ . In this way, the error caused by the POD method can be estimated *a-posteriorly*. If the error is too large, then we have to include more POD basis functions in our Galerkin approximation for (3.3).

M. KAHLBACHER AND S. VOLKWEIN

# 4. The pod Galerkin discretization of $(\mathbf{P}^{k,\beta})$

In this section we briefly introduce the POD method and derive the reduced-order model for the optimality system (3.3) of  $(\mathbf{P}^{k,\beta})$ . Moreover, *a-priori* error estimates for POD Galerkin schemes for the state as well as for the adjoint equation are shown.

# 4.1. The POD method

Let  $u \in U$  be given. Then there exists a vector  $\mathbf{u} = (u_1, \ldots, u_N)^T \in \mathbb{R}^N$  such that

$$u(\mathbf{x}) = \sum_{i=1}^{N} u^{i} b_{i}(\mathbf{x}) \quad \text{for all } \mathbf{x} \in \Omega.$$
(4.1)

Furthermore, we suppose that

$$\mathbf{u} \in \mathcal{D} = [\underline{u}_1, \overline{u}_1] \times \dots [\underline{u}_N, \overline{u}_N] \subset \mathbb{R}^N \text{ with } 0 < \underline{u}_i \leq \overline{u}_i \text{ for } i = 1, \dots, N.$$

By  $y = y(\mathbf{u})$  we denote the unique solution to (3.3a), where u is given as in (4.1). The snapshot ensemble is chosen to be

$$\mathcal{V} = \operatorname{span}\left\{y(\mathbf{u}) \,|\, \mathbf{u} \in \mathcal{D}\right\} \subset V. \tag{4.2}$$

Then,  $d = \dim \mathcal{V} \leq \infty$ . Let  $\ell < \infty$  satisfy  $1 \leq \ell \leq d$ . The POD basis  $\{\psi_i\}_{i=1}^{\ell}$  of rank  $\ell$  is given by the solution to the following minimization problem:

$$\min_{\{\psi_i\}_{i=1}^{\ell} \subset V} \int_{\mathcal{D}} \left\| y(\mathbf{u}) - \sum_{i=1}^{\ell} \left\langle y(\mathbf{u}), \psi_i \right\rangle_V \psi_i \right\|_V^2 \mathrm{d}\mathbf{u} \quad \text{s.t.} \quad \left\langle \psi_i, \psi_j \right\rangle_V = \delta_{ij}, \tag{P}^{\ell}$$

where  $\delta_{ij} = 1$  if i = j and  $\delta_{ij} = 0$  if  $i \neq j$ . It is well-known that the solution to  $(\mathbf{P}^{\ell})$  can be derived by the methods of snapshots [21]: solve the symmetric eigenvalue problem

$$\mathcal{K}v_i = \lambda_i v_i \quad \text{for } i = 1, \dots, \ell$$

in  $L^2(\mathcal{D})$ , where  $\mathcal{K}: L^2(\mathcal{D}) \to L^2(\mathcal{D})$  is given by

$$(\mathcal{K}v)\left(\tilde{\mathbf{u}}\right) = \int_{\mathcal{D}} \langle y(\mathbf{u}), y(\tilde{\mathbf{u}}) \rangle_{V} v(\mathbf{u}) \, \mathrm{d}\mathbf{u} \quad \text{for } \tilde{\mathbf{u}} \in \mathcal{D} \text{ and } v \in L^{2}(\mathcal{D}),$$

and set

$$\psi_i = \frac{1}{\sqrt{\lambda_i}} \int_{\mathcal{D}} y(\mathbf{u}) v_i(\mathbf{u}) \,\mathrm{d}\mathbf{u} \quad \text{for } i = 1, \dots, \ell.$$

From the Hilbert-Schmidt theorem [19], page 29, it follows that there exists a complete orthogonal basis  $\{\psi_i\}_{i=1}^d$  for  $\mathcal{V} = \operatorname{range}(\mathcal{R})$  and a sequence  $\{\lambda_i\}_{i=1}^d$  of real numbers such that

$$\mathcal{R}\psi_i = \lambda_i\psi_i \text{ for } i = 1, \dots, d \text{ and } \lambda_1 \ge \lambda_2 \ge \dots \ge \lambda_d \ge 0$$

To obtain a complete orthogonal basis in the separable Hilbert space V we need an orthogonal basis for  $(\operatorname{range}(\mathcal{R}))^{\perp}$ . This can be done by the Gram-Schmidt procedure. Hence, we suppose in the following that  $\{\psi_i\}_{i=1}^{\infty}$  is a complete orthogonal basis for V. In particular, we have

$$\int_{\mathcal{D}} \left\| y(\mathbf{u}) - \sum_{i=1}^{\ell} \left\langle y(\mathbf{u}), \psi_i \right\rangle_V \psi_i \right\|_V^2 \mathrm{d}\mathbf{u} = \sum_{i=\ell+1}^{\infty} \lambda_i.$$
(4.3)

If  $1 \leq d = \dim \mathcal{V} \leq \infty$  holds, it follows that  $\lambda_i > 0$  for  $1 \leq i \leq d$  and  $\mathcal{R}\psi_i = 0$  for i > d.

## Remark 4.1.

(a) In real computations, we do not have the  $y(\mathbf{u})$  for all  $\mathbf{u} \in \mathcal{D}$  at hand. For that purpose let  $\{\mathbf{u}_j\}_{j=1}^M$  define grid points in  $\mathcal{D}$  and  $y_j = y(\mathbf{u}_j), j = 1, \ldots, M$ , be approximations for u at the grid points  $\mathbf{u}_j$ . We set

$$\mathcal{V}_M = \operatorname{span}\left\{y_1, \ldots, y_M\right\}$$

with  $d_M = \dim \mathcal{V}_M \leq M$ . Then, for given  $\ell \leq d_M$  we consider the minimization problem

$$\min_{\{\psi_i\}_{i=1}^{\ell} \subset V} \sum_{j=1}^{M} \alpha_j \left\| y_j - \sum_{i=1}^{\ell} \langle y_j, \psi_i \rangle_V \psi_i \right\|_V^2 \quad \text{s.t.} \quad \langle \psi_i, \psi_j \rangle_V = \delta_{ij}, \tag{\mathbf{P}_M^{\ell}}$$

instead of  $(\mathbf{P}^{\ell})$ . In  $(\mathbf{P}_{M}^{\ell})$  the  $\alpha_{j}$ 's stand for weights in the used quadrature rule;

(b) in our numerical experiments in Section 6 we determine a POD basis before the optimization utilizing snapshots from the state and the adjoint equation for (**P**). More precisely, we choose a grid  $\{\mathbf{u}_j\}_{j=1}^M$  in the parameter set  $\mathcal{D}$  and compute the states  $y_j = y(\mathbf{u}_j)$  by solving (2.1). Then, using  $\mathbf{u}_j$  and  $y_j$  we compute the solution  $p_j = p(\mathbf{u}_j)$  to (2.2) for  $j = 1, \ldots, M$ . Then, the snapshot ensemble is given by the linear space  $\mathcal{V}_M = \text{span}\{y_1, \ldots, y_M, p_1, \ldots, p_M\}$ . In [12] it is shown that the error of the POD Galerkin approximation can be improved significantly by incorporating adjoint information into the snapshot ensemble.

## 4.2. POD Galerkin scheme for the optimality system

The error analysis presented in this section shows that there is a real chance to decrease the error by increasing the number of snapshots used by the POD method.

Let  $y = \hat{y}^k + Su$  be the state associated with some control  $u \in U_{ad}^k$ , and let  $\mathcal{V}$  by given as in (4.2). We fix  $\ell$  with  $\ell \leq \dim \mathcal{V}$  and compute the first  $\ell$  POD basis functions  $\psi_1, \ldots, \psi_\ell \in V$  by solving  $\mathcal{K}v_i = \lambda_i v_i$  for  $i = 1, \ldots, \ell$ . Then we define the finite-dimensional linear space

$$V^{\ell} = \operatorname{span}\left\{\psi_1, \dots, \psi_{\ell}\right\} \subset V_{\ell}$$

Endowed with the topology in V it follows that  $V^{\ell}$  is a Hilbert space. Let  $\mathcal{P}^{\ell}$  denote the orthogonal projection of V onto  $V^{\ell}$  defined by

$$\mathcal{P}^{\ell}\psi = \sum_{i=1}^{\ell} \langle \psi, \psi_i \rangle_V \psi_i \quad \text{for } \psi \in V.$$
(4.4)

Using (4.3) we have

$$\int_{\mathcal{D}} \left\| y(\mathbf{u}) - \mathcal{P}^{\ell} y(\mathbf{u}) \right\|_{V}^{2} \mathrm{d}\mathbf{u} = \left\| y - \mathcal{P}^{\ell} y \right\|_{L^{2}(\mathcal{D};V)}^{2} = \sum_{i=\ell+1}^{\infty} \lambda_{i}.$$

Using standard arguments the POD Galerkin appoximation of (3.3) yields the following linear system: determine  $(y^{\ell}, u^{\ell}, p^{\ell}) \in V^{\ell} \times U^k_{ad} \times V^{\ell}$  satisfying

(1) The (linearized) state equation

$$\int_{\Omega} \nabla y^{\ell} \cdot \nabla \psi + \left( u^{k} y^{\ell} + u^{\ell} y^{k} \right) \psi \, \mathrm{d}\mathbf{x} + \int_{\Gamma} y^{\ell} \psi \, \mathrm{d}\mathbf{s} = -\langle e(x^{k}), \psi \rangle_{V', V}$$
(4.5a)

for all  $\psi \in V^{\ell}$ ;

(2) the (linearized) dual equation

$$\int_{\Omega} \nabla p^{\ell} \cdot \nabla \psi + u^{k} p^{\ell} \psi \, \mathrm{d}\mathbf{x} + \int_{\Gamma} p^{\ell} \psi \, \mathrm{d}\mathbf{s} = \int_{\Omega} \left( y_{d} - y^{k} - y^{\ell} - \beta u^{\ell} p^{k} \right) \psi \, \mathrm{d}\mathbf{x}$$
(4.5b)

for all  $\psi \in V^{\ell}$ ; and

(3) the (linearized) variational inequality

$$\int_{\Omega} \left(\beta y^{\ell} p^{k} + \sigma(u^{k} + u^{\ell}) + y^{k} p^{\ell}\right) (u - u^{\ell}) \,\mathrm{d}\mathbf{x} \ge 0 \tag{4.5c}$$

for all  $u \in U_{ad}^k$ .

**Remark 4.2.** Using similar arguments as in [24] it follows that

$$\|\bar{u} - \bar{u}^{\ell}\|_{L^{2}(\Omega)} \leq C\left(\|\bar{y}(\bar{\mathbf{u}}) - \mathcal{P}^{\ell}\bar{y}(\bar{\mathbf{u}})\|_{V} + \|\bar{p}(\bar{\mathbf{u}}) - \mathcal{P}^{\ell}\bar{p}(\bar{\mathbf{u}})\|_{V}\right)$$
(4.6)

for a constant C > 0; see in the appendix. In particular, we have  $\lim_{\ell \to \infty} \|\bar{u} - \bar{u}^{\ell}\|_{L^2(\Omega)} = 0$ .

## 4.3. A-posteriori error estimate for the POD approximation

In this subsection we complete the discussion of the *a*-posteriori estimate by combining Remarks 3.5 and 4.2. The proposition permits to estimate  $\|\bar{u} - \bar{u}^{\ell}\|_{L^2(\Omega)}$  by the norm of an appropriate  $\zeta$ , while Remark 4.2 will be used to show that  $\zeta$  tends to zero as  $\ell \to \infty$ , since it ensures the convergence of  $\bar{u}^{\ell}$  to the optimal control  $\bar{u}$  for  $(\mathbf{P}^{k,\beta})$ .

For any  $\ell$  let  $\bar{u}^{\ell} \in U_{ad}^{k}$  be the optimal control solving (4.5) together with  $\bar{y}$  and  $\bar{p}$ . Then,  $\bar{u}^{\ell}$  is taken as a suboptimal  $u^{p}$  for  $(\mathbf{P}^{k,\beta})$ , *i.e.*, in Remark 4.2 we choose  $u^{p} := \bar{u}^{\ell}$ .

**Theorem 4.3.** Suppose that  $(\bar{y}, \bar{u}, \bar{p}) \in V \times U_{ad}^k \times V$  is the solution to (3.3).

(1) Let  $\ell \leq d$  be arbitrarily given and  $(\bar{y}^{\ell}, \bar{u}^{\ell}, \bar{p}^{\ell}) \in V \times U_{ad}^k \times V$  be the solution to (4.5). Using  $u^p = \bar{u}^{\ell}$  compute the residuum  $\zeta^{\ell} = \zeta$  as in Remark 3.5. Then,

$$\|\bar{u} - \bar{u}^{\ell}\|_{L^{2}(\Omega)} \le \frac{1}{\sigma} \|\zeta^{\ell}\|_{L^{2}(\Omega)}$$

(2) If  $\{\psi_i\}_{i=1}^{\infty}$  is a complete orthonormal basis for V, then  $\lim_{\ell \to \infty} \|\zeta^{\ell}\|_{L^2(\Omega)} = 0.$ 

The proof is a variant of the proof of Theorem 4.11 in [24].

**Remark 4.4.** Part (2) of Theorem 4.3 shows that  $\|\zeta^{\ell}\|_{L^2(\Omega)}$  can be expected smaller than any  $\varepsilon > 0$  provided that  $\ell$  is taken sufficiently large. Motivated by this result we set up Algorithm 2.

Algorithm 2 (POD method for  $(\mathbf{P}^{k,\beta})$  with *a-posteriori* estimator)

1: Choose a maximal number  $\ell_{max} > 0$  of POD basis function, an  $\ell < \ell_{max}$ , and a stopping criterium  $\varepsilon > 0$ .

- 2: Compute a POD basis of rank  $\ell$  by solving  $(\mathbf{P}^{\ell})$ .
- 3: repeat

- 5: Calculate the suboptimal control  $\bar{u}^{\ell}$  to  $(\mathbf{P}^{k,\beta})$ .
- 6: Using  $u^p = \bar{u}^\ell$  compute the residuum  $\zeta^\ell$  as in Remark 3.5.
- 7: **if**  $\|\zeta^{\ell}\|_{L^{2}(\Omega)} \geq \varepsilon$  **then**
- 8: Set  $\ell = \ell + 1$ .
- 9: end if

10: **until**  $\|\zeta^{\ell}\|_{L^{2}(\Omega)} < \varepsilon$  or  $\ell > \ell_{max}$ 

11: Return  $\ell$  and suboptimal control  $\bar{u}^{\ell}$ .

**Remark 4.5.** Of, course, step 8 can be replaced by 8: Set  $\ell = \ell + L$ . with any natural number L.

<sup>4:</sup> Derive a reduced-order model of rank  $\ell$  for  $(\mathbf{P}^{k,\beta})$ .

# 5. Convergence of the inexact SQP method

Let us assume that the solution  $\bar{z} = (\bar{y}, \bar{u}, \bar{p})$  to (3.3) satisfies  $\bar{u} > u_a^k$  (inactive control). Then, the variational inequality (3.3c) can be replaced by the equation

$$\int_{\Omega} \left( \beta \bar{y} p^k + \sigma (u^k + \bar{u}) + y^k \bar{p} \right) b_i \, \mathrm{d}\mathbf{x} = 0 \quad \text{for } 1 \le i \le N.$$
(3.3c')

Thus, (3.3a), (3.3b), and (3.3c') leads to a linear operator equation in  $X' \times V'$  of the form (1.3) for the variable  $\bar{z}$ . Since the mapping  $(x, p) \mapsto L^{k,\beta}(x, p)$  is twice continuously Fréchet-differentiable, it can be shown that there exists a constant C > 0 independent of the iteration level k so that  $||A_k||_{L(X \times V, X' \times V')} \leq C$ , where  $L(X \times V, X' \times V')$  denotes the Banach space of all bounded linear operators from  $X \times V$  to  $X' \times V'$  endowed with the common operator norm.

Let the solution  $\bar{z}^{\ell} = (\bar{y}^{\ell}, \bar{u}^{\ell}, \bar{p}^{\ell})$  to (4.5) satisfy  $\bar{u}^{\ell} > u_a$ . Then, the variational inequality (4.5c) yields the equation

$$\int_{\Omega} \left( \beta \bar{y}^{\ell} p^{k} + \sigma (u^{k} + \bar{u}^{\ell}) + y^{k} \bar{p}^{\ell} \right) b_{i} \, \mathrm{d}\mathbf{x} = 0 \quad \text{for } 1 \le i \le N,$$

$$(4.5c')$$

so that (4.5a), (4.5b), and (4.5c') can be formulated as a finite-dimensional linear system of the form (1.4) for the variable  $\bar{z}^{\ell} \in \mathbb{R}^n$  with  $n = 2\ell + N$  ( $\ell$  coefficients  $\bar{y}_i$  for  $\bar{y}^{\ell}$ , N coefficients for  $\bar{u}^{\ell}$ , and  $\ell$  coefficients  $\bar{p}_i$  for  $\bar{p}^{\ell}$ ). From (1.4) we obtain the coefficients  $\bar{u}_i$ ,  $1 \leq i \leq N$ , for the suboptimal control  $\bar{u}^{\ell}$ . Then, we define the bounded operator  $\mathcal{I} : \mathbb{R}^n \to X \times Y$  as follows:

$$\mathbb{R}^{n} \ni \bar{z}^{\ell} = \begin{pmatrix} \bar{y}_{1} \\ \vdots \\ \bar{y}_{\ell} \\ \bar{u}_{1} \\ \vdots \\ \bar{u}_{N} \\ \bar{p}_{1} \\ \vdots \\ \bar{p}_{\ell} \end{pmatrix} \mapsto \mathcal{I} \bar{z}^{\ell} = \begin{pmatrix} \tilde{y}^{\ell} \\ \sum_{i=1}^{N} \bar{u}_{i} b_{i} \\ \sum_{i=1}^{N} \bar{p}^{\ell} \end{pmatrix} \in V \times U \times V$$

where  $\tilde{y}^{\ell}$  and  $\tilde{p}^{\ell}$  solve (3.3a) and (3.3b), respectively, with  $\bar{u}^{\ell}$  instead of  $\bar{u}$ . From (1.3) and Remark 3.3 it follows

$$\begin{split} \|A_{k}(\mathcal{I}\bar{z}^{\ell} - b_{k}\|_{X' \times V'} &= \|A_{k}(\mathcal{I}\bar{z}^{\ell} - \bar{z})\|_{X' \times V'} \leq \|A_{k}\|_{L(X \times V, X' \times V')} \|\mathcal{I}\bar{z}^{\ell} - \bar{z}\|_{X \times V} \\ &\leq C \sqrt{\|\tilde{y}^{\ell} - \bar{y}\|_{V}^{2} + \|\bar{u}^{\ell} - \bar{u}\|_{L^{2}(\Omega)}^{2} + \|\tilde{p}^{\ell} - \bar{p}\|_{V}^{2}} \\ &\leq \tilde{C} \|\bar{u}^{\ell} - \bar{u}\|_{L^{2}(\Omega)} \end{split}$$

with  $\tilde{C} = C \sqrt{1 + C_p^2}$ . Consequently, Theorem 4.3 implies that

$$\|A_k(\mathcal{I}\bar{z}^\ell) - b_k\|_{X' \times V'} \le \frac{\tilde{C}}{\sigma} \|\zeta^\ell\|_{L^2(\Omega)} \stackrel{\ell \to \infty}{\longrightarrow} 0.$$

Therefore, we combine Algorithms 1 and 2 to arrive an POD *a-posteriori* error based inexact SQP method for the bilinear optimal control problem; see Algorithm 3. By L' we denote the Fréchet derivative of the Lagrangian with respect to (x, p).

**Remark 5.1.** In our numerical experiments it is more efficient to compute the POD basis of rank  $\ell_{\text{max}}$  only once at the beginning of the SQP method. As snapshots  $y = y(\mathbf{u})$ ,  $\mathbf{u} = (u_1, \ldots, u_N)$ , we take solutions from the bilinear problem (2.1) for different controls  $\mathbf{u} \in \mathcal{D}$ . Then we apply Algorithm 3 without the step 4.

Algorithm 3 (POD *a-posteriori* error based inexact SQP method)

1: Choose  $x^0 = (y^0, u^0) \in X_{ad}, p^0 \in V, 0 < \varepsilon_a \leq \varepsilon_r \ll 1, q \in [1, 2], \text{ and set } k = 1.$ 2: while  $||L'(x^k, p^k)|| \ge \varepsilon_a + \varepsilon_r ||L'(x^0, p^0)||$  do Choose  $\ell_{max} > 0$ ,  $\ell < \ell_{max}$ , and  $\varepsilon_k = \min(0.5, \|L'(x^k, p^k)\|^q)$ . 3: Call Algorithm 2 with  $\ell_{max}$  and stopping criterium  $\varepsilon_k$ . 4: 5: if  $\ell > \ell_{max}$  then STOP and restart the algorithm (*e.g.*, with a larger  $\ell_{\max}$ ). 6: end if 7: Determine  $t_k \in (0, 1]$  by a line search; see Algorithm 1. 8: Set  $(x^{k+1}, p^{k+1}) = (x^k, p^k) + t_k(\tilde{y}^\ell, \bar{u}^\ell, \tilde{p}^\ell)$  and k = k+1. 9: 10: end while

For Algorithm 3 we have proved the next convergence theorem.

**Theorem 5.2.** Let  $x^* \in X_{ad}$  be a local solution to (**P**),  $p^*$  the associated Lagrange multiplier, and  $z^* = (x^*, p^*)$ . Suppose that

- (A1) the starting value  $(x^0, p^0)$  of Algorithm 3 is sufficiently close to  $(x^*, p^*)$ ;
- (A2) the optimality system (3.3) admits a (unique) solution  $(\bar{y}, \bar{u}, \bar{p})$  so that  $\bar{u} > u_a^k$  in  $\Omega$  a.e. (inactive  $\bar{u}$ );
- (A3) for sufficiently large  $\ell \leq \ell_{\max}$  the optimality system (4.5) admits a (unique) solution  $(\bar{y}^{\ell}, \bar{u}^{\ell}, \bar{p}^{\ell})$  so that  $\bar{u}^{\ell} > u_a^k$  in  $\Omega$  a.e. (inactive  $\bar{u}^{\ell}$ ).

Let the iterates  $\{z^k\}_{k\in\mathbb{N}}$ ,  $z^k = (x^k, p^k)$ , be generated by Algorithm 3. Then,  $\lim_{k\to\infty} z^k = z^*$  in  $X \times V$ . In particular, we obtain superlinear and quadratic rate of convergence:

$$\begin{aligned} \|z^{k+1} - z^*\|_{X \times V} &\leq c_k \|z^k - z^*\|_{X \times V} & \text{for all } k \text{ if } q \in (1, 2), \\ \|z^{k+1} - z^*\|_{X \times V} &\leq c \|z^k - z^*\|_{X \times V}^2 & \text{for all } k \text{ if } q = 2, \end{aligned}$$

where  $c_k$  satisfies  $\lim_{k\to\infty} c_k = 0$  and c is a positive constant independent of k.

## Remark 5.3.

- (a) Assumption (A1) ensures that the iterates  $(x^k, p^k)$  belong to a neighborhood of  $(x^*, p^*)$ , where the convergence of the SQP method is ensured without any globalization strategy. In particular, at each level k of the SQP method the linear-quadratic optimal control problem  $(\mathbf{P}^{k,\beta})$  admits a unique solution for  $\beta = 1$  and we can choose  $t_k = 1$ ;
- (b) if Assumption (A2) and (A3) do not hold, we have to deal with the variational inequalities (3.3c) and (4.5c), respectively. Thus, (3.3) and (4.5) are generalized equations. We have to apply the theory of Newton methods for generalized equations; see [1];
- (c) in Step 9 of Algorithm 3 the dual variable can also be chosen as the least-squares multiplier, *i.e.*,

$$p^{k+1} = \arg\min_{p \in V} \|J'(x^{k+1} + e'(x^{k+1})^*p\|_{X'},$$

see [20], for instance.

### 6. Numerical experiments

We present two examples concerning *a-posteriori* error estimates for POD. The numerical tests are executed on a standard 3.0 GHz desktop PC. We are using the MATLAB 7.1 package including its integrated PDE Toolbox for the FE discretization.



FIGURE 1. Run 1: domain  $\Omega$  with subdomains (left plot) and decay of the 22 largest eigenvalues (right plot).

**Run 1.** Let the domain  $\Omega$  be given by

$$\Omega = \left\{ \mathbf{x} = (x_1, x_2) \left| \frac{x_1^2}{0.8^2} + \frac{x_2^2}{0.7^2} < 1 \right\} \subset \mathbb{R}^2.$$

Moreover, we assume that  $\Omega$  consists of two disjunct subdomains ( $\Omega_1$  and  $\Omega_2$ ), where  $\Omega_2$  is given as the quadrilateral with corners  $(0.22, -0.28), (0.35, 0.34), (-0.30, 0.41), \text{ and } (-0.18, -0.32), \text{ and } \Omega_1 = \Omega \setminus \Omega_2$ ; see left plot of Figure 1. We choose two characteristic functions as shape functions for the control, *i.e.*, N = 2 and  $b_i = \chi_{\Omega_i}$  for i = 1, 2. In  $\Omega_1$  let f = 1, whereas f = -10 in  $\Omega_2$ . The domain is discretized by a FE grid that consists of 4862 degrees of freedom. In the context of Remark 4.1-b we utilize 441 snapshots computed on an equidistant grid for  $\mathbf{u} = (u_1, u_2) \in [0, 10] \times [0, 10] =: \mathcal{D}$  and choose  $\ell_{max} = 25$ . The computation of the snapshots requires 51 s. The POD basis computation needs 2 s. The 22 largest eigenvalues for the POD computation are shown in the right plot of Figure 1. Notice that the relative error in the  $H^1$ -norm between the finite element solution to (2.1) and the POD solution for  $\ell = \ell_{max}$  is about  $1.06 \times 10^{-8}$ . For the cost functional (see Rem. 2.4) let  $\Omega_m = \Omega_2$ ,  $y_d$  be the solution to (2.1) for  $u = (b_1 + 3b_2)/2$ ,  $u^{\circ} = 0$ , and  $\sigma = 10^{-4}$ . In Algorithm 3 (step 1) we choose  $\varepsilon_r = 10^{-4}$ ,  $\varepsilon_a = 10^{-6}$  and q = 1.05. Thus, we have a locally superlinear rate of convergence. We initialize the SQP algorithm as follows:  $u^0 = 0$ ,  $y^0 = y(u^0)$  is the solution to (2.1) for  $u = u^0$  and  $p^0 = p(u^0)$  solves the adjoint equation (2.2) with  $y^* = y^0$  and  $u^* = u^0$ . In each SQP iteration, we solve the reduced system (4.5) to obtain  $\bar{z}^{\ell} = (\bar{y}^{\ell}, \bar{u}^{\ell}, \bar{p}^{\ell})$ . Then, we evaluate  $\mathcal{I}\bar{z}^{\ell} = (\tilde{y}^{\ell}, \bar{u}^{\ell}, \tilde{p}^{\ell})$ . If the *a-posteriori* error estimator ensures a small error for  $\|\bar{u} - \bar{u}^{\ell}\|_{L^{2}(\Omega)}$ , we set  $x^{k+1} = x^{k} + (\tilde{y}^{\ell}, \bar{u}^{\ell})$  and compute the associated least-squares update  $p^{k+1}$ ; see Remark 5.3-c. The SQP method stops after nine iterations and requires 6 s. The optimal control is  $0.77b_1 + 1.07b_2$ . The optimal state and the residuum are presented in Figure 2. In Table 1 the convergence behaviour is shown. We observe that the globalized SQP method converges to a local optimal solution. The inexactness can be controlled by the *a-posteriori* error estimator. In the first SQP iterations the step size  $t_k$  is less than one, so that the region of superlinear rate of convergence is reached in the 7th SQP iteration. Then, we have fast convergence in the last three iterations.

**Run 2.** Let  $\Omega$  consist of 3 subdomains; see Figure 3. We choose the associated three characteristic functions as shape functions for the control, *i.e.*, N = 3 and  $b_i = \chi_{\Omega_i}$  for i = 1, 2, 3. In  $\Omega_1$  let f = 2, in  $\Omega_2$  we have f = 0.1, whereas f = 1 in  $\Omega_3$ . The domain is discretized by a FE grid that consists of 4819 degrees of freedom. We apply a POD computation (using 2197 snapshots computed on an equidistant grid for  $\bar{u} = (u_1, u_2, u_3) \in [0,3] \times [0,3] \times [0,3] =: \mathcal{D}$  and choose  $\ell_{max} = 16$ . The 16 largest eigenvalues of the eigenvalue problem in the POD computation are shown in the right plot of Figure 3. The relative error in the  $H^1$ -norm between the POD

| k | $  L'(x^{k-1}, p^{k-1})  $ | $  L'(x^{k-1}, p^{k-1})  ^q$ | $\ \zeta^\ell\ /\sigma$ | $\ell$ | $t_k$ |
|---|----------------------------|------------------------------|-------------------------|--------|-------|
| 1 | 9.34e-02                   | 8.29e-02                     | 3.36e-04                | 6      | 0.5   |
| 2 | 6.01e-02                   | 5.23e-02                     | 1.44e-04                | 6      | 0.5   |
| 3 | 4.51e-02                   | 3.86e-02                     | 4.62e-05                | 6      | 0.5   |
| 4 | 3.33e-02                   | 2.80e-02                     | 4.64e-05                | 6      | 0.5   |
| 5 | 2.27e-02                   | 1.88e-02                     | 8.89e-05                | 6      | 0.5   |
| 6 | 1.27e-02                   | 1.02e-02                     | 1.05e-04                | 6      | 1.0   |
| 7 | 1.95e-03                   | 1.43e-03                     | 5.77e-05                | 6      | 1.0   |
| 8 | 1.71e-05                   | 9.90e-06                     | 2.20e-07                | 10     | 1.0   |
| 9 | 4.04e-11                   | -                            | _                       | -      | _     |

TABLE 1. Run 1: SQP iterations, stopping criterium, tolerance for inexactness, *a-posteriori* error estimate, number  $\ell$  of POD ansatz functions and step size parameter  $t_k$ .



FIGURE 2. Run 1: optimal state  $y^*$  (left plot) and residuum  $|y^* - y^d|/|y^d|$  in percent (right plot).



FIGURE 3. Run 2: domain  $\Omega$  with subdomains (left plot) and decay of the 16 largest eigenvalues (right plot).

TABLE 2. Run 2: SQP iterations, stopping criterium, tolerance for inexactness, *a-posteriori* error estimate, number  $\ell$  of POD ansatz functions and step size parameter  $t_k$ .

| k | $  L'(x^{k-1}, p^{k-1})  $ | $  L'(x^{k-1}, p^{k-1})  ^q$ | $\ \zeta^\ell\ /\sigma$ | $\ell$ | Hessian mod. | $t_k$ |
|---|----------------------------|------------------------------|-------------------------|--------|--------------|-------|
| 1 | 3.01e-02                   | 2.53e-02                     | 6.71e-03                | 10     | yes          | 0.5   |
| 2 | 4.49e-01                   | 4.31e-01                     | 2.88e-04                | 10     | no           | 1.0   |
| 3 | 3.63e-01                   | 3.45e-01                     | 2.92e-04                | 10     | no           | 1.0   |
| 4 | 5.16e-02                   | 4.45e-02                     | 1.17e-04                | 10     | no           | 1.0   |
| 5 | 1.49e-02                   | 1.21e-02                     | 7.33e-05                | 10     | no           | 1.0   |
| 6 | 1.63e-03                   | 1.19e-03                     | 7.50e-06                | 10     | no           | 1.0   |
| 7 | 6.66e-07                   | -                            | _                       | _      | -            | _     |



FIGURE 4. Run 2: optimal state  $y^*$  (left plot) and relative residuum  $|y^* - y^d|/|y^d|$  in percent (right plot).

solution with  $\ell = \ell_{max}$  and the FE solution is  $2.58 \times 10^{-6}$ . The computing time for the snapshot computation is 249 s, the computation of the 16 POD basis functions costs 34 s.

For the cost functional let  $\Omega_m = \Omega_1$ ,  $u^{\circ} = (3b_1 + 2b_2 + b_3)/2$ ,  $y_d$  be the solution to (2.1) for  $u = u^{\circ}$ , and  $\sigma = 10^{-4}$ . In Algorithm 3 we choose the same values for  $\varepsilon_r$ ,  $\varepsilon_a$  and q as in Run 1. As starting values we choose  $u^0 = 3b_1 + 4b_2 + 3b_3$ ,  $y^0$  as the solution to (2.1) for  $u = u^0$  and  $p^0$  as the solution to (2.2) with  $y^* = y^0$  and  $u^* = u^0$ . The SQP methods stops after seven SQP iterations and requires 4 s. The optimal control is  $1.49b_1 + 1.01b_2 + 0.52b_3$ . The optimal state and the residuum are presented in Figure 4. We observe a fast rate of convergence in the last two iterations. The results are stated in Table 2. Again, the convergence of the SQP method can be ensured by the *a*-posteriori error estimator. Fast rate of convergence can be observed in the last iteration, whereas at the beginning a globalization of the SQP method is required.

### APPENDIX

Proof of Theorem 3.4. Choosing  $u = \overline{u}$  in (3.3c) and  $u = u^p$  in (3.4) we obtain

$$0 \leq \left\langle p^{k}\bar{y} + y^{k}\bar{p} + \sigma(u^{k} + \bar{u}) - p^{k}y^{p} - y^{k}p^{p} - \sigma(y^{k} + y^{p}) - \zeta, u^{p} - \bar{u}\right\rangle_{L^{2}(\Omega)}$$
  
=  $-\sigma \left\|\bar{u} - u^{p}\right\|_{L^{2}(\Omega)}^{2} + \left\langle p^{k}(\bar{y} - y^{p}) + y^{k}(\bar{p} - p^{p})\right) + \zeta, u^{p} - \bar{u}\right\rangle_{L^{2}(\Omega)}.$ 

Let  $y = S(y^p - \bar{y})$ . Using Remark 3.2 and (3.3b) we find

$$\begin{split} \left\langle y^k \bar{p}, u^p - \bar{u} \right\rangle_{L^2(\Omega)} &= \int_{\Omega} (u^p - \bar{u}) y^k \bar{p} \, \mathrm{d} \mathbf{x} \\ &= -\int_{\Omega} \nabla y \cdot \nabla \bar{p} + u^k y \bar{p} \, \mathrm{d} \mathbf{x} - \int_{\Gamma} y \bar{p} \, \mathrm{d} \mathbf{x} = -\int_{\Omega} \nabla \bar{p} \cdot \nabla y + u^k \bar{p} y \, \mathrm{d} \mathbf{x} - \int_{\Gamma} \bar{p} y \, \mathrm{d} \mathbf{x} \\ &= \int_{\Omega} \left( \bar{y} p^k + y^k + \bar{y} - y_d \right) y \, \mathrm{d} \mathbf{x} = \left\langle \bar{u} p^k + y^k + \bar{y} - y_d, \mathcal{S}(u^p - \bar{u}) \right\rangle_{V', V} \\ &= \left\langle \mathcal{S}^\star \left( \bar{u} p^k + y^k + \bar{y} - y_d \right), u^p - \bar{u} \right\rangle_{L^2(\Omega)}. \end{split}$$

Analogously, we obtain

$$\left\langle y^{k}p^{p}, u^{p} - \bar{u} \right\rangle_{L^{2}(\Omega)} = \left\langle \mathcal{S}^{\star} \left( u^{p}p^{k} + y^{k} + y^{p} - y_{d} \right), u^{p} - \bar{u} \right\rangle_{L^{2}(\Omega)}.$$

Therefore,

$$\begin{split} \left\langle y^{k}(\bar{p}-p^{p}), u^{p}-\bar{u} \right\rangle_{L^{2}(\Omega)} &= \left\langle y^{k}\bar{p}, u^{p}-\bar{u} \right\rangle_{L^{2}(\Omega)} - \left\langle y^{k}p^{p}, u^{p}-\bar{u} \right\rangle_{L^{2}(\Omega)} \\ &= \left\langle \mathcal{S}^{\star}\left(\bar{y}-y^{p}+(\bar{u}-u^{p})p^{k}\right), u^{p}-\bar{u} \right\rangle_{L^{2}(\Omega)}. \end{split}$$

Recall that  $\mathcal{S}(u^p - \bar{u}) = y^p - \bar{y}$ . Hence,

$$\begin{split} \left\langle y^{k}(\bar{p}-p^{p}), u^{p}-\bar{u} \right\rangle_{L^{2}(\Omega)} &= -\left\| \bar{y}-y^{p} \right\|_{L^{2}(\Omega)}^{2} + \left\langle (\bar{u}-u^{p})p^{k}, y^{p}-\bar{y} \right\rangle_{L^{2}(\Omega)} \\ &\leq \left\langle p^{k}(y^{p}-\bar{y}), \bar{u}-u^{p} \right\rangle_{L^{2}(\Omega)}. \end{split}$$

Summarizing, we obtain

$$0 \leq -\sigma \|\bar{u} - u^p\|_{L^2(\Omega)}^2 + \langle y^k(\bar{p} - p^p) - p^k(\bar{y} - y^p) - \zeta, u^p - \bar{u} \rangle_{L^2(\Omega)}$$
  
$$\leq -\sigma \|\bar{u} - u^p\|_{L^2(\Omega)}^2 + \langle p^k(y^p - \bar{y}) - p^k(\bar{y} - y^p) - \zeta, \bar{u} - u^p \rangle_{L^2(\Omega)}$$
  
$$= -\sigma \|\bar{u} - u^p\|_{L^2(\Omega)}^2 - \langle \zeta, u^p - \bar{u} \rangle_{L^2(\Omega)}.$$

Consequently, we have

$$\sigma \|\bar{u} - u^p\|_{L^2(\Omega)}^2 \le -\langle \zeta, u^p - \bar{u} \rangle_{L^2(\Omega)} \le \|\zeta\|_{L^2(\Omega)} \|u^p - \bar{u}\|_{L^2(\Omega)},$$

which gives the claim.

Proof of (4.6). From (3.3c) and (4.5c) we find

$$\int_{\Omega} \mathcal{G}^{k,\beta}(\bar{y},\bar{u},\bar{p})(\bar{u}^{\ell}-\bar{u}) \,\mathrm{d}\mathbf{x} \ge 0 \quad \text{and} \quad \int_{\Omega} \mathcal{G}^{k,\beta}(\bar{y}^{\ell},\bar{u}^{\ell},\bar{p}^{\ell})(\bar{u}-\bar{u}^{\ell}) \,\mathrm{d}\mathbf{x} \ge 0.$$

Adding both inequalities we deduce

$$\sigma \|\bar{u} - \bar{u}^{\ell}\|_{L^{2}(\Omega)}^{2} \leq \int_{\Omega} \left(\beta(p^{k}(\bar{y} - \bar{y}^{\ell})) + y^{k}(\bar{p} - \bar{p}^{\ell})\right) (\bar{u}^{\ell} - \bar{u}) \,\mathrm{d}\mathbf{x}.$$
(A.1)

508

We estimate

$$\begin{split} \int_{\Omega} y^{k} (\bar{p} - \bar{p}^{\ell}) (\bar{u}^{\ell} - \bar{u}) \, \mathrm{d}\mathbf{x} &= \int_{\Omega} y^{k} \bar{p} (\bar{u}^{\ell} - \bar{u}) \, \mathrm{d}\mathbf{x} \\ &= \int_{\Omega} \mathcal{S}^{\star} \left( y^{k} + \bar{y} - y_{d} + \bar{u}p^{k} \right) \left( \bar{u}^{\ell} - \bar{u} \right) \, \mathrm{d}\mathbf{x} \\ &+ \int_{\Omega} (\mathcal{S}^{\ell})^{\star} \left( y^{k} + \bar{y}^{\ell} - y_{d} + \bar{u}^{\ell}p^{k} \right) \left( \bar{u}^{\ell} - \bar{u} \right) \, \mathrm{d}\mathbf{x} \\ &= \int_{\Omega} \left( \mathcal{S}^{\star} - (\mathcal{S}^{\ell})^{\star} \right) \left( y^{k} - y_{d} \right) \left( \bar{u}^{\ell} - \bar{u} \right) \, \mathrm{d}\mathbf{x} \\ &+ \int_{\Omega} \left( \mathcal{S}^{\star} \left( \bar{y} + \bar{u}p^{k} \right) - (\mathcal{S}^{\ell})^{\star} \left( \bar{y}^{\ell} + \bar{u}^{\ell}p^{k} \right) \right) \left( \bar{u}^{\ell} - \bar{u} \right) \, \mathrm{d}\mathbf{x} \\ &= \int_{\Omega} \left( \left( \mathcal{S}^{\star} - (\mathcal{S}^{\ell})^{\star} \right) \left( y^{k} - y_{d} \right) + \mathcal{S}^{\star} (p^{k}\bar{u}) - (\mathcal{S}^{\ell})^{\star} (p^{k}\bar{u}^{\ell}) \right) \left( \bar{u}^{\ell} - \bar{u} \right) \, \mathrm{d}\mathbf{x} \\ &+ \int_{\Omega} \left( \mathcal{S}^{\star} \bar{y} - (\mathcal{S}^{\ell})^{\star} \bar{y}^{\ell} \right) \left( \bar{u}^{\ell} - \bar{u} \right) \, \mathrm{d}\mathbf{x}. \end{split}$$

Recall that  $\bar{y} = \hat{y}_k + S\bar{u}$  and  $\bar{y}^{\ell} = \hat{y}_k + S^{\ell}\bar{u}^{\ell}$ . Therefore,

$$\int_{\Omega} \left( \mathcal{S}^{\star} \bar{y} - (\mathcal{S}^{\ell})^{\star} \bar{y}^{\ell} \right) \left( \bar{u}^{\ell} - \bar{u} \right) \mathrm{d}\mathbf{x} = \int_{\Omega} \left( \mathcal{S}^{\star} (\hat{y}_{k} + \mathcal{S}\bar{u}) - (\mathcal{S}^{\ell})^{\star} (\hat{y}_{k}^{\ell} + \mathcal{S}^{\ell} \bar{u}^{\ell}) \right) \left( \bar{u}^{\ell} - \bar{u} \right) \mathrm{d}\mathbf{x}.$$

From

$$\mathcal{S}^* \mathcal{S} \bar{u} - (\mathcal{S}^\ell)^* \mathcal{S}^\ell \bar{u}^\ell = \mathcal{S}^* \mathcal{S} \bar{u} - (\mathcal{S}^\ell)^* \mathcal{S}^\ell \bar{u} + (\mathcal{S}^\ell)^* \mathcal{S}^\ell (\bar{u} - \bar{u}^\ell)$$

and

$$\int_{\Omega} \left( (\mathcal{S}^{\ell})^{\star} \mathcal{S}^{\ell}(\bar{u} - \bar{u}^{\ell}) \right) (\bar{u}^{\ell} - \bar{u}) \, \mathrm{d}\mathbf{x} = - \left\| \mathcal{S}^{\ell}(\bar{u} - \bar{u}^{\ell}) \right\|_{L^{2}(\Omega)}^{2} \leq 0,$$

we infer that

$$\int_{\Omega} \left( \mathcal{S}^{\star} \bar{y} - (\mathcal{S}^{\ell})^{\star} \bar{y}^{\ell} \right) (\bar{u}^{\ell} - \bar{u}) \, \mathrm{d}\mathbf{x} = \int_{\Omega} \left( \mathcal{S}^{\star} \hat{y}_{k} - (\mathcal{S}^{\ell})^{\star} \hat{y}_{k}^{\ell} + \mathcal{S}^{\star} \mathcal{S} \bar{u} - (\mathcal{S}^{\ell})^{\star} \mathcal{S}^{\ell} \bar{u}^{\ell} \right) (\bar{u}^{\ell} - \bar{u}) \, \mathrm{d}\mathbf{x}$$
$$= \int_{\Omega} \left( \mathcal{S}^{\star} (\hat{y}_{k} + \mathcal{S} \bar{u}) - (\mathcal{S}^{\ell})^{\star} (\hat{y}_{k}^{\ell} + \mathcal{S}^{\ell} \bar{u}) \right) (\bar{u}^{\ell} - \bar{u}) \, \mathrm{d}\mathbf{x}$$
$$= \int_{\Omega} \left( \mathcal{S}^{\star} \bar{y} - (\mathcal{S}^{\ell})^{\star} \tilde{y}^{\ell} \right) (\bar{u}^{\ell} - \bar{u}) \, \mathrm{d}\mathbf{x},$$

where  $\tilde{y}^{\ell} = \hat{y}^{k}_{\ell} + S\bar{u}$ . Consequently,

$$\begin{split} \int_{\Omega} y^{k} (\bar{p} - \bar{p}^{\ell}) (\bar{u}^{\ell} - \bar{u}) \, \mathrm{d}\mathbf{x} &= \int_{\Omega} \left( \left( \mathcal{S}^{\star} - (\mathcal{S}^{\ell})^{\star} \right) (y^{k} - y_{d}) + \mathcal{S}^{\star} (p^{k} \bar{u}) - (\mathcal{S}^{\ell})^{\star} (p^{k} \bar{u}^{\ell}) \right) (\bar{u}^{\ell} - \bar{u}) \, \mathrm{d}\mathbf{x} \\ &+ \int_{\Omega} \left( \mathcal{S}^{\star} \bar{y} - (\mathcal{S}^{\ell})^{\star} \right) (y^{k} - y_{d}) + \mathcal{S}^{\star} (p^{k} \bar{u}) - (\mathcal{S}^{\ell})^{\star} (p^{k} \bar{u}^{\ell}) \right) (\bar{u}^{\ell} - \bar{u}) \, \mathrm{d}\mathbf{x} \\ &= \int_{\Omega} \left( \left( \mathcal{S}^{\star} - (\mathcal{S}^{\ell})^{\star} \right) (y^{k} - y_{d}) + \mathcal{S}^{\star} (p^{k} \bar{u}) - (\mathcal{S}^{\ell})^{\star} (p^{k} \bar{u}^{\ell}) \right) (\bar{u}^{\ell} - \bar{u}) \, \mathrm{d}\mathbf{x} \\ &+ \int_{\Omega} \left( \mathcal{S}^{\star} \bar{y} - (\mathcal{S}^{\ell})^{\star} \tilde{y}^{\ell} \right) (\bar{u}^{\ell} - \bar{u}) \, \mathrm{d}\mathbf{x} \\ &= \int_{\Omega} \mathcal{S}^{\star} \left( y^{k} + \bar{y} - y_{d} + \bar{u}p^{k} \right) (\bar{u}^{\ell} - \bar{u}) \, \mathrm{d}\mathbf{x} \\ &- \int_{\Omega} \left( \mathcal{S}^{\ell} \right)^{\star} \left( y^{k} + \tilde{y}^{\ell} - y_{d} + \bar{u}^{\ell} p^{k} \right) (\bar{u}^{\ell} - \bar{u}) \, \mathrm{d}\mathbf{x} = \int_{\Omega} y^{k} (\bar{p} - \tilde{p}^{\ell}) (\bar{u}^{\ell} - \bar{u}) \, \mathrm{d}\mathbf{x}, \end{split}$$

where  $\tilde{p}^{\ell}$  solves

$$\int_{\Omega} \nabla \tilde{p}^{\ell} \cdot \nabla \psi + u^k \tilde{p}^{\ell} \psi \, \mathrm{d}\mathbf{x} + \int_{\Gamma} p^{\ell} \psi \, \mathrm{d}\mathbf{s} = \int_{\Omega} y_d - y^k - \tilde{y}^{\ell} - \beta \bar{u} p^k) \psi \, \mathrm{d}\mathbf{x} \text{ for all } \psi \in V^{\ell}.$$

Inserting this into (A.1) and using (2.3) we find

$$\sigma \|\bar{u} - \bar{u}^{\ell}\|_{L^{2}(\Omega)}^{2} \leq \int_{\Omega} \left(\beta p^{k}(\bar{y} - \bar{y}^{\ell}) + y^{k}(\bar{p} - \tilde{p}^{\ell})\right) (\bar{u}^{\ell} - \bar{u}) \,\mathrm{d}\mathbf{x}$$
$$\leq C_{1} \left(\|\bar{y} - \bar{y}^{\ell}\|_{V} + \|\bar{p} - \tilde{p}^{\ell}\|_{V}\right) \|\bar{u}^{\ell} - \bar{u}\|_{L^{2}(\Omega)}$$

with  $C_1 = C_{emb} \max(\beta \| p^k \|_{L^4(\Omega)}, \| y^k \|_{L^4(\Omega)})$ , which was the claim.

## References

- W. Alt, The Lagrange-Newton method for infinite-dimensional optimization problems. Numer. Funct. Anal. Optim. 11 (1990) 201–224.
- [2] A.C. Antoulas, Approximation of Large-Scale Dynamical Systems. Advances in Design and Control, SIAM, Philadelphia (2005).
- [3] N. Arada, E. Casas and F. Tröltzsch. Error estimates for the numerical approximation of a semilinear elliptic control problem. Comput. Optim. Appl. 23 (2002) 201–229.
- [4] E. Arian, M. Fahl and E.W. Sachs, Trust-region proper orthogonal decomposition for flow control. Technical Report 2000-25, ICASE (2000).
- [5] J.A. Atwell, J.T. Borggaard and B.B. King, Reduced order controllers for Burgers' equation with a nonlinear observer. Int. J. Appl. Math. Comput. Sci. 11 (2001) 1311–1330.
- [6] P. Benner and E.S. Quintana-Ortí, Model reduction based on spectral projection methods, in *Reduction of Large-Scale Systems, Lect. Notes Comput. Sci. Eng.* 45, edited by P. Benner, V. Mehrmann and D.C. Sorensen (2005) 5–48.
- [7] P. Deuflhard, Newton Methods for Nonlinear Problems: Affine Invariance and Adaptive Algorithms, Springer Series in Comput. Math. 35 (2004).
- [8] L.C. Evans, Partial Differential Equations, Graduate Studies in Mathematics. American Mathematical Society, Providence, Rhode Island 19 (2002).
- [9] R.S. Falk, Error estimates for the approximation of a class of variational inequalities. Math. Comput. 28 (1974) 963–971.
- [10] T. Gänzler, S. Volkwein and M. Weiser, SQP methods for parameter identification problems arising in hyperthermia. Optim. Methods Softw. 21 (2006) 869–887.
- [11] M. Hintermüller, On a globalized augmented Lagrangian SQP-algorithm for nonlinear optimal control problems with box constraints, in *Fast solution methods for discretized optimization problems, International Series of Numerical Mathematics.* edited by K.-H. Hoffmann, R.H.W. Hoppe and V. Schulz, Birkhäuser publishers, Basel **138** (2001) 139–153.
- [12] M. Hinze and S. Volkwein, Error estimates for abstract linear-quadratic optimal control problems using proper orthogonal decomposition. Comput. Optim. Appl. 39 (2008) 319–345.
- [13] A. Kröner and B. Vexler, A priori error estimates for elliptic optimal control problems with bilinear state equation. J. Comput. Appl. Math. 230 (2009) 781–802.
- [14] K. Kunisch and S. Volkwein, Proper orthogonal decomposition for optimality systems. ESAIM: M2AN 42 (2008) 1–23.
- [15] H.V. Ly and H.T. Tran, Modeling and control of physical processes using proper orthogonal decomposition. Math. Comput. Model. 33 (2001) 223–236.
- [16] K. Malanowski, C. Büskens and H. Maurer, Convergence of approximations to nonlinear control problems, in *Mathematical Programming with Data Perturbation*, edited by A.V. Fiacco and M. Dekker. Inc., New York (1997) 253–284.
- [17] A.T. Patera and G. Rozza, Reduced Basis Approximation and A Posteriori Error Estimation for Parametrized Partial Differential Equations. MIT Pappalardo Graduate Monographs in Mechanical Engineering (2006).
- [18] S.S. Ravindran, Adaptive reduced order controllers for a thermal flow system using proper orthogonal decomposition. SIAM J. Sci. Comput. 28 (2002) 1924–1942.
- [19] M. Read and B. Simon, Methods of Modern Mathematical Physics I: Functional Analysis. Academic Press, Boston (1980).

- [20] E.W. Sachs and S. Volkwein, Augmented Lagrange-SQP methods with Lipschitz-continuous Lagrange multiplier updates. SIAM J. Numer. Anal. 40 (2002) 233–253.
- [21] L. Sirovich, Turbulence and the dynamics of coherent structures, parts I-III. Quart. Appl. Math. XLV (1987) 561-590.
- [22] 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. Math. Comput. Modelling of Dynam. Systems 17 (2011) 355-369.
- [23] F. Tröltzsch, Optimal Control of Partial Differential Equations: Theory, Methods and Applications, Graduate Studies in Mathematics. American Mathematical Society 112 (2010).
- [24] F. Tröltzsch and S. Volkwein, POD a-posteriori error estimates for linear-quadratic optimal control problems. Comput. Optim. Appl. 44 (2009) 83–115.
- [25] M. Vallejos and A. Borzì, Multigrid optimization methods for linear and bilinear elliptic optimal control problems. Computing 82 (2008) 31–52.
- [26] S. Volkwein, Mesh-independence of an augmented Lagrangian-SQP method in Hilbert spaces. SIAM J. Control Optimization 38 (2000) 767–785.