



**HAL**  
open science

# Estimating the Control Error in discretized PDE-constrained Optimization

Roland Becker

► **To cite this version:**

Roland Becker. Estimating the Control Error in discretized PDE-constrained Optimization. Journal of Numerical Mathematics, 2006, 14 (3), pp.163-185. hal-00158885

**HAL Id: hal-00158885**

**<https://hal.science/hal-00158885>**

Submitted on 1 Jul 2007

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Estimating the Control Error in discretized PDE-constrained Optimization

Roland Becker \*

10th April 2006

**Abstract** — In this article we develop an a posteriori error estimator for discretized optimal control problems. We are interested in estimating the error in the control variable, measured in a natural norm. We prove an error representation formula involving only quantities at hand in a second-order optimization iteration, supposing a strong form of second-order sufficient condition. Possible generalization to the control-constrained case is indicated.

**Keywords:** Finite elements, Optimal control, Adaptive methods

## 1. INTRODUCTION

We consider an abstract optimal control problem with a partial differential equation written in weak form as the state equation. Let  $Q$  and  $V$  be Hilbert spaces for the control variable  $q$  and the state variable  $u$ . They are coupled by the nonlinear state equation:

$$\text{Find } u \in V \text{ such that } a(q, u)(v) = l(v) \quad \forall v \in V, \quad (1.1)$$

where  $a : Q \times V \times V \rightarrow \mathbb{R}$  is a smooth function which is linear with respect to the third argument  $v$  representing the test function. We adopt the convention that the dependency with respect to the variables in the second pair of parenthesis is linear. Further, in (1.1)  $l \in V^*$  represents fixed data.

From an optimization point of view (1.1) defines a constraint which has to be fulfilled, when minimizing the smooth functional

$$J : Q \times V \rightarrow \mathbb{R}. \quad (1.2)$$

In addition, in order to formulate constraints on the control variable, we introduce a closed convex set

$$Q_{ad} \subset Q, \quad (1.3)$$

and the abstract optimization problem reads:

$$\inf_{q \in Q_{ad}} J(q, u) \quad \text{subject to (1.1)}. \quad (1.4)$$

---

\*Laboratoire de Mathématiques Appliquées, Université de Pau et des Pays de l'Adour BP 1155, 64013 PAU Cedex

We will throughout suppose that (1.4) admits a unique solution which is stable in the sense that a continuous second-order condition, which will be formulated later on (see (H1) and (H2) below), holds.

When a finite element method is used to discretize (1.1),  $V$  is replaced by a finite-dimensional subspace  $V_h$ ;  $h$  denotes a generic discretization parameter such as mesh size. The state equation, for given  $q \in Q$ , is accordingly changed to:

$$\text{Find } u_h \in V_h \text{ such that } a(q, u_h)(v_h) = l(v_h) \quad \forall v_h \in V_h. \quad (1.5)$$

In general, any solution of (1.5) leads to an *inadmissible* state in the sense of the continuous optimization problem (1.4). Therefore, when discussing approximate solutions of the optimization problem it is necessary to consider approximations of the whole optimization problem. Further the task of computing such an approximate solution requires the definition of an error measure which might be or might not be intrinsic to the optimization problem. In this article, we are interested in estimating the error in the control variable, measured in a natural norm. Estimation of the control error can be of independent interest, for example when one is only interested in determining the best control. It can also be a step to further error estimation, since an estimate of the error in state variable is implied under appropriate assumptions on the state operator.

In addition to discretization of the state space we consider discretization of the control space. We therefore introduce another finite-dimensional space  $Q_h \subset Q$  using the same generic parameter  $h$  (which could for example represent two different meshes: one for the state and another for the control space) and define  $Q_{ad,h} := Q_h \cap Q_{ad}$ . The discrete optimization problem reads:

$$\inf_{q_h \in Q_{ad,h}} J(q_h, u_h) \quad \text{subject to (1.5)}. \quad (1.6)$$

Error in the control variable is generated by both approximations: First the control space is discretized by  $Q_h \subset Q$ , and second, the discretization of the state equation introduces another source of error. We try to separate the influence of these approximations in order to give an a posteriori error estimate which could yield a basis for adapting the parameters underlying the two discrete spaces individually.

In order to separate difficulties, we first consider the case of unconstrained optimization  $Q_{ad} = Q$  in Section 3. We derive an a posteriori error estimator which can be implemented in a two-mesh iteration and then only uses quantities already present in a second-order optimization iteration.

Section 4 deals with the case that the controls are searched in a closed convex subset of the control space. The considered framework covers a number of challenging problems as the obstacle problem and bound constrained distributed or boundary control. These problems require different functional analytic settings and sophisticated finite element tools as special interpolation operators, see the recent work in these fields. It is not possible to give a precise treatment here. Instead, we are interested in a formal generalization of the error estimate for the unconstrained case which might serve as a guideline in concrete applications.

The crucial questions for an adaptive algorithm are: First, when do we accept the discrete solution as a sufficiently accurate approximation? Second, how can we adapt the discretization parameters efficiently in an iterative way? These questions are at the heart of a posteriori error estimation for finite element discretizations of elliptic partial differential equations, see [1,9,20]. There, the error is usually measured in the energy norm. In the context of optimal control problems, the question of what error measure should be chosen largely depends on the application. In previous work, we have developed an a posteriori error estimator for the value of the cost functional itself as the error measure [2]. An extension to the case of error control with respect to a functional of the control has been developed in [3].

In this paper, we present an a posteriori error estimator for the control error measured in a natural norm:

$$\|q - q_h\|_Q \leq \eta(q_h, u_h). \quad (1.7)$$

We call this norm natural since it is related to a second-order sufficient condition for the continuous optimization problem. This condition does not directly rely on a coercivity assumption on the state operator. As a matter of fact, the estimator  $\eta$  in (1.7) depends on two continuous auxiliary equations. These equations turn out to be the state and adjoint equations corresponding to the computed fixed control. We propose to approximate their solutions by a two-level algorithm. Then, the necessary information can be computed in the first step of the optimization iteration on the next finer mesh.

The analysis of approximate optimization problems is a classical subject of optimization theory, see for example [17]. The new contribution of our work is the derivation of an estimator which allows for *local* mesh refinement, which is practically important in the presence of singularities. For such an estimator local information about the influence of the local residuals of the partial differential equation on the error in the objective functional is the crucial information. Singularities may be caused either by the differential operator, the domain, or other components of the optimization problem as the cost functional. They lead to different asymptotic behavior of the residuals due to reduction of regularity of the solution (state, control, or Lagrange multiplier). For example, a usual feature of inverse problems is the occurrence of point measurements, which lead to singularities in the adjoint variable.

A posteriori error estimators for finite element discretizations of optimal control problems have also been proposed in [13,12]. There, specific optimal control problems involving elliptic second-order partial differential equations are considered and the coercivity of the state operator is used; the structure of the a posteriori error estimator is different from ours since it involves the sum of squares of the residuals of the different equations. In contrast to this, the estimator discussed here contains the product of residuals from the primal and dual equations; see below for details.

### 1.1. Notation.

We denote by  $(\cdot, \cdot)_Q$  the scalar product and by  $\|\cdot\|_Q$  its induced norm in the Hilbert space  $Q$ , and similarly for  $V$ . The dual space of  $Q$  is denoted as  $Q^*$ . If no confusion

is caused, we suppress the subscript in  $(\cdot, \cdot)_Q$ . However, in general,  $(u, v)$  denotes the  $L^2(\Omega)$  scalar product. For a differentiable function  $f : Q \rightarrow \mathbb{R}$ , we define as usual its gradient and Hessian via the Ritz identification:  $f'(q)(p) = (\nabla f(q), p)_Q$  and  $f''(q)(p_1, p_2) = (\nabla^2 f(q)p_1, p_2)_Q$ . We denote the partial derivatives of a function of two variables  $f(q, p)$  by  $f'_p(q, p)$  and  $f'_q(q, p)$ ; the partial gradients are denoted by  $\nabla_p f(q, p)$  and  $\nabla_q f(q, p)$ .

## 2. HYPOTHESIS AND STRUCTURE OF THE OPTIMISATION PROBLEM

We suppose that the continuous minimization problem (1.4) admits a unique local solution  $(q, u)$ . Further, we suppose that there exists a neighborhood  $U$  of  $q$  such that the state equation is uniquely solvable for all  $p \in U$ . This leads to the definition of the *solution operator*:

$$S : U \rightarrow V : q \mapsto u, \quad u \text{ is the solution of (1.1) for given } q. \quad (2.1)$$

This operator inherits its smoothness properties from  $a$ . Now we can eliminate the constraint defined through the state equation by introduction of the *reduced functional*:

$$j : U \rightarrow \mathbb{R} \quad j(q) := J(q, S(q)). \quad (2.2)$$

The minimization problem (1.4) is now changed into:

$$j(q) \rightarrow \inf, \quad q \in Q_{ad}. \quad (2.3)$$

The first-order necessary condition for this minimization problem relies on the convexity of  $Q_{ad}$  and reads:

$$j'(q)(p - q) \geq 0 \quad \forall p \in Q_{ad}. \quad (2.4)$$

In addition, we require the following second-order condition to be verified: There exist  $\gamma > 0$  and  $\xi > 0$  such that for all  $p \in B_\xi(q)$

$$j''(p)(\delta q, \delta q) \geq \gamma \|\delta q\|_Q^2 \quad \forall \delta q \in Q. \quad (\mathbf{H1})$$

The condition **(H1)** is stronger than standard sufficient second-order conditions [5,15]. First, we do not exploit the (convex) constraints in order to weaken the requirement. Second, we impose uniformity of the coercivity condition. Note however, that in case of a quadratic functional  $j$ , we precisely assume that the smallest eigenvalue  $\gamma$  of the symmetric positive operator  $\nabla j^2(q)$  is bounded away from zero. In this case we also have  $\xi = +\infty$ .

In the following we present some examples in order to illustrate condition **(H1)**.

**Example 2.1.** We consider a linear-quadratic problem of the form

$$Au = f + Bq, \quad J(q, u) = \frac{1}{2} \|Cu - d\|_Z^2 + \frac{\alpha}{2} \|q\|_Q^2, \quad (2.5)$$

where  $A : V \rightarrow V^*$ ,  $B : Q \rightarrow V^*$  are linear continuous operators and  $A$  is an isomorphism. In addition  $C : V \rightarrow Z$  is a linear continuous observation operator and  $d \in Z$  represents given observations. We find the reduced functional to be the quadratic functional

$$j(q) = \frac{1}{2} \|Gq - \tilde{d}\|_Z^2 + \frac{\alpha}{2} \|q\|_Q^2, \quad (2.6)$$

where  $G := CA^{-1}B$  is the control-to-observation map and  $\tilde{d} := d - CA^{-1}f$ . The reduced functional is quadratic with gradient and Hessian given by:

$$\nabla j(q) = (G^*G + \alpha I)q - G^*\tilde{d}, \quad \nabla^2 j(q) = G^*G + \alpha I, \quad (2.7)$$

where  $I$  denotes the identity operator and we use the identification described at the end of the introduction. Hypothesis **(H1)** amounts to defining  $\gamma$  by

$$\gamma := \lambda_{\min}(G^*G) + \alpha. \quad (2.8)$$

Especially, no regularization is needed ( $\alpha = 0$ ) in the finite-dimensional case, if  $G$  has no zero singular value.

**Example 2.2.** As a special case of Example 2.1 we consider the following distributed control problem. The state equation is the Poisson equation on a bounded domain  $\Omega \subset \mathbb{R}^2$  with homogenous Dirichlet boundary conditions:

$$-\Delta u = q \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \partial\Omega. \quad (2.9)$$

The cost functional is defined with a given function  $d \in L^2(\Omega)^2$  by

$$J(q, u) := \frac{1}{2} \int_{\Omega} |\nabla u - d|^2 dx + \frac{\alpha}{2} \|q\|_Q^2. \quad (2.10)$$

It is now clear that Hypothesis **(H1)** is related to the choice of the space  $Q$ . We choose  $Q = H^{-1}(\Omega)$ ,  $\alpha = 0$  and find, since

$$\|q\|_{H^{-1}(\Omega)} = \sup_{v \in H_0^1(\Omega)} \frac{q(v)}{\|v\|_{H_0^1(\Omega)}} = \sup_{v \in H_0^1(\Omega)} \frac{(\nabla S(q), \nabla v)}{\|v\|_{H_0^1(\Omega)}} = \|S(q)\|_{H_0^1(\Omega)}, \quad (2.11)$$

that

$$j''(q)(p, p) = \|p\|_{H^{-1}(\Omega)}^2 \quad (\gamma = 1). \quad (2.12)$$

For illustration of the estimator we introduce two additional variants of Example 2.1.

**Example 2.3.** We present a variant of Example 2.2. The state equation is again the Poisson equation on a bounded domain  $\Omega \subset \mathbb{R}^2$  with homogenous Dirichlet

boundary conditions (2.9). The cost functional is defined with a given function  $d \in L^2(\Omega)$  by

$$J(q, u) := \frac{1}{2} \int_{\Omega} |u - d|^2 dx + \frac{\alpha}{2} \|q\|_Q^2. \quad (2.13)$$

We choose  $Q = L^2(\Omega)$  and well-posedness of the problem requires now  $\alpha > 0$ , since  $G : Q \rightarrow V$  is a compact operator. We obtain a variant with control constraints by setting  $Q_{ad} = \{q \in Q : \underline{q} \leq q(x) \leq \bar{q} \text{ a.e. in } \Omega\}$  with real numbers  $\underline{q} \leq \bar{q}$ .

For illustration of the estimator we introduce two further variants of Example 2.1.

**Example 2.4.** Let  $Q = H_0^1(\Omega)$  and  $V = \emptyset$ . Let us define for given  $f \in L^2(\Omega)$

$$J(q, u) := \frac{1}{2} \|q\|_Q^2 - (f, q). \quad (2.14)$$

Since the constraint is trivial we end up with the Dirichlet problem for the control,

$$-\Delta q = f \quad \text{in } \Omega, \quad q = 0 \quad \text{on } \partial\Omega. \quad (2.15)$$

Here the choice of norms is trivial and we have  $\gamma = 1$  again. This example shows that the estimator to be developed is a generalization of the classical energy error estimator for the Ritz projection.

Let  $\psi$  be a smooth function satisfying  $\psi|_{\partial\Omega} \leq 0$ . Choosing  $Q_{ad} = \{q \in Q : q \geq \psi\}$  leads to the obstacle problem with obstacle  $\psi$ . We will later consider the case of bilateral constraints,  $Q_{ad} = \{q \in Q : \underline{q} \leq q(x) \leq \bar{q} \text{ a.e. in } \Omega\}$  with real numbers  $\underline{q} \leq 0 \leq \bar{q}$ .

**Example 2.5.** We define  $V = (H_0^1(\Omega))^n$ ,  $n$  denoting the space dimension, and  $Q = L^2(\Omega)/\mathbb{R}$ . We consider the state equation

$$-\Delta u + \nabla q = f, \quad (A = -\Delta, B = \nabla), \quad (2.16)$$

where the control is used to minimize the cost functional

$$J(q, u) := \frac{1}{2} \|\operatorname{div} u\|_{L^2(\Omega)}^2. \quad (2.17)$$

This minimization problem leads to the Stokes equations for a Lipschitz domain  $\Omega$  (since we can impose  $\operatorname{div} u = 0$ ). From the analysis of the Stokes problem we know that the operator  $G = CA^{-1}B = \operatorname{div}(-\Delta)^{-1}\nabla$  is spectrally equivalent to the identity operator in  $Q$ , see [10].

Next, we consider the approximation of the infinite-dimensional minimization problem (1.4) by the Galerkin method (1.6), where the unknowns  $q_h$  and  $u_h$  are

sought in finite-dimensional subspaces  $V_h \subset V$  and  $Q_h \subset Q$ . We suppose existence of the discrete solutions  $q_h, u_h$  and existence of the *discrete solution operator*:

$$S_h : U \cap Q_h \rightarrow V_h : \quad q_h \mapsto u_h \quad u_h \text{ the solution of (1.5)}. \quad (2.18)$$

In analogy to the continuous level we define the *discrete reduced functional*:

$$j_h : U \cap Q_h \rightarrow \mathbb{R} : \quad j_h(q_h) := J(q_h, S_h(q_h)). \quad (2.19)$$

The discrete constraints set  $Q_{ad,h}$  is supposed to satisfy  $Q_{ad,h} \subset Q_{ad} \cap Q_h$ . The discrete optimal control  $q_h \in Q_{ad,h}$  is solution of the variational inequality

$$j'_h(q_h)(p_h - q_h) \geq 0 \quad \forall p_h \in Q_{ad,h}. \quad (2.20)$$

Our next hypothesis is that the discrete optimal control is not too far from to the continuous solution in order **(H1)** to be applicable. We therefore require that

$$\|q - q_h\|_Q \leq \xi \quad (\xi \text{ is defined in (H1)}). \quad (\mathbf{H2})$$

We remark that for linear-quadratic problems, **(H2)** is void.

We have the following abstract error bound, which reflects a basic result of convex analysis.

**Proposition 2.1.** *Let assumptions **(H1)** and **(H2)** be satisfied. Then, the difference between a solution  $q$  of (2.4) and a solution  $q_h$  of (2.20) can be estimated by*

$$\frac{\gamma}{2} \|q - q_h\|_Q^2 \leq j(q_h) - j(q). \quad (2.21)$$

**Proof.** Let  $q_s := (1-s)q + sq_h$  for  $0 \leq s \leq 1$  and define the smooth function  $\varphi(s) := j(q_s) + (1-s)j'(q_s)(q_h - q)$ . We find that  $\varphi'(s) = (1-s)j''(q_s)(q_h - q, q_h - q)$ . Therefore

$$\begin{aligned} \frac{\gamma}{2} \|q - q_h\|_Q^2 &\leq \int_0^1 \varphi'(s) ds = \varphi(1) - \varphi(0) \\ &= j(q_h) - j(q) - j'(q)(q_h - q) \leq j(q_h) - j(q). \end{aligned}$$

In the last step we have used that  $q_h \in Q_{ad,h} \subset Q_{ad}$  and  $j'(q)(q_h - q) \geq 0$  due to (2.4).  $\square$

Proposition 2.1 tells us that the error can be bounded by the difference in the value of the continuous reduced functional. In the context of Example 2.4 we recover the well-known fact that for a Galerkin discretization the error in energy norm can be bounded by the difference in energy.

Any approximation of  $j(q)$ , e.g. the value of the cost functional corresponding to better approximation spaces  $V_{h'} \supset V_h$  and  $Q_{h'} \supset Q_h$  leads to an error estimator,



which can be used as a stopping criterion (supposed the stability constant  $\gamma$  has been estimated correctly). However, such an error estimator does not tell us how to improve the approximation spaces. Before going to such an estimator, we present a variant of Proposition 2.1. Using the local convexity of the reduced functional, we derive the following alternative estimate:

**Proposition 2.2.** *Let assumptions (H1) and (H2) be satisfied. Then, the difference between a solution  $q$  of (2.4) and a solution  $q_h$  of (2.20) can be estimated by*

$$\frac{\gamma}{2} \|q - q_h\|_Q^2 \leq j'(q_h)(q_h - q). \quad (2.22)$$

**Proof.** It is sufficient to recall that for a smooth convex function there holds:

$$j(q) \geq j(p) + j'(p)(q - p),$$

and choose  $p = q_h$ . □

Estimate (2.22) can be used to define a stopping criterion for an iteration as before. In addition, it can be related to the residual of the equations defining  $q$  and  $q_h$ .

We end up this section with the computation of the first and second derivatives of the reduced functional. We introduce the Lagrange functional of the optimization problem:

$$\mathcal{L}(q, u, z) := J(q, u) + f(z) - a(q, u)(z). \quad (2.23)$$

We will use the well-known fact that the gradient and Hessian of the reduced functional can be computed by means of the Lagrange functional. These relations are usually used in second-order optimization algorithms. They also form the basis for our approach to a posteriori error estimation. For completeness and preparation of the error estimation, we give the proofs of these facts.

**Proposition 2.3.** *Suppose  $a$  and  $J$  are differentiable. Let the state equation (1.1) and the following adjoint equation be satisfied:*

$$\text{Find } z \in V \text{ such that } a'_u(q, u)(v, z) = J'_u(q, u)(v) \quad \forall v \in V. \quad (2.24)$$

*Then the gradient of the reduced functional is related to the Lagrangian by*

$$j'(q)(p) = \mathcal{L}'_q(q, u, z)(p). \quad (2.25)$$

**Proof.** Let to  $q(\varepsilon) = q + \varepsilon p$  correspond the state  $u(\varepsilon) = S(q(\varepsilon))$ . For arbitrary  $z$  we have that

$$\begin{aligned} j(q(\varepsilon)) - j(q) &= \mathcal{L}(q(\varepsilon), u(\varepsilon), z) - \mathcal{L}(q, u, z) \\ &= \varepsilon \{ \mathcal{L}'_q(q, u, z)(p) + \mathcal{L}'_u(q, u, z)(u'(0)) \} + o(\varepsilon), \end{aligned}$$

where  $u'(0)$  denotes the derivative of  $\varepsilon \mapsto u(\varepsilon)$  for  $\varepsilon = 0$ . If  $z$  satisfies (2.24), the second term in braces vanishes, and we obtain the stated result.  $\square$

By Proposition 2.3 the necessary first-order condition for the constrained minimization problem (1.4) is the system of equations consisting of (1.1,2.24) and the variational inequality

$$\mathcal{L}'_q(q, u, z)(p - q) \geq 0 \quad \forall p \in Q_{ad}.$$

This system is called *optimality system*.

**Proposition 2.4.** *Suppose  $a$  and  $J$  are twice differentiable. Let  $u$  and  $z$  satisfy equations (1.1) and (2.24), respectively. In addition let  $w$  and  $y$  be the solution of the following system:*

$$\begin{cases} a'_u(q, u)(w, v) = -a'_q(q, u)(p, v) & \forall v \in V, \\ a'_u(q, u)(v, y) = -a''_{qu}(q, u)(p, v, z) - a''_{uu}(q, u, z)(v, w, z) \\ \quad - J''_{qu}(q, u)(p, v) - J''_{uu}(q, u)(w, v) & \forall v \in V. \end{cases} \quad (2.26)$$

Then the Hessian of the reduced functional is related to the Lagrangian by

$$j''(q)(p, r) = \mathcal{L}''_{qq}(q, u, z)(p, r) + \mathcal{L}''_{qu}(q, u, z)(r, w) + \mathcal{L}''_{qz}(q, u, z)(r, y). \quad (2.27)$$

**Proof.** We introduce the following auxiliary functional:

$$\mathcal{M}(q, u, z, w, y) := \mathcal{L}'_q(q, u, z)(p) + \mathcal{L}'_u(q, u, z)(w) + \mathcal{L}'_z(q, u, z)(y). \quad (2.28)$$

From Proposition 2.3 we know that  $\mathcal{M}(q, u, z, w, y) = j'(q)(p)$  if  $\nabla_{(w,y)} \mathcal{M} = 0$ . We apply the same proposition again to obtain that  $j''(q)(p, r) = \mathcal{M}'_q(q, u, z, w, y)(r)$  if  $\nabla_{(u,z,w,y)} \mathcal{M} = 0$ . The equations corresponding to  $\nabla_{(w,y)} \mathcal{M} = 0$  define  $w$  and  $y$ . The equations corresponding to  $\nabla_{(u,z)} \mathcal{M} = 0$  read:

$$\begin{cases} \mathcal{L}''_{uu}(q, u, z)(w, v) + \mathcal{L}''_{zu}(q, u, z)(y, v) = -\mathcal{L}''_{qu}(q, u, z)(p, v), \\ \mathcal{L}''_{uz}(q, u, z)(w, v) + \mathcal{L}''_{zz}(q, u, z)(y, v) = -\mathcal{L}''_{qz}(q, u, z)(p, v). \end{cases} \quad (2.29)$$

First we note that  $\mathcal{L}''_{zz} = 0$  since the Lagrangian is linear with respect to  $z$ . Furthermore, the cost functional  $J$  does not depend on  $z$ . We solve first the second equation in (2.29) which is the first equation of (2.26). The first equation of (2.29) becomes the second equation of (2.26).  $\square$

### 3. THE CASE WITHOUT CONSTRAINTS

In this section we consider the discretized optimization problem

$$\begin{cases} \inf_{q \in Q_h} J(q, u) & \text{subject to:} \\ a(q, u_h)(v_h) = l(v_h) & \forall v_h \in V_h. \end{cases} \quad (3.1)$$

We suppose that  $Q_{ad} = Q$  and  $Q_{ad,h} = Q_h$ . In order to measure how well the three equations forming the optimality system are satisfied, we introduce the residual functional  $\rho \in (Q \times V \times V)^*$  defined by

$$\begin{aligned} \rho(p, y, w) := & \mathcal{L}'_q(q_h, u_h, z_h)(p) + \{l(y) - a(q_h, u_h)(y)\} \\ & + \{J'_u(q_h, u_h)(w) - a'_u(q_h, u_h)(z_h, w)\}. \end{aligned} \quad (3.2)$$

The following result is motivated by the aim to derive local mesh refinement criteria. Only the residuals of the state and adjoint equations and the first-order optimality condition are involved. They are weighted by projection errors of additional functions.

**Theorem 3.1.** *Suppose  $a$  and  $J$  are three-times differentiable. Let assumptions (H1) and (H2) be satisfied and suppose that  $Q_{ad} = Q$  and  $Q_{ad,h} = Q_h$ . Let  $\tilde{u} \in V$  and  $\tilde{z} \in V$  be solutions of the following auxiliary problems at fixed control:*

$$\begin{cases} a(q_h, \tilde{u})(v) = l(v) & \forall v \in V, \\ a'_u(q_h, \tilde{u})(\tilde{z}, v) = J'_u(q_h, \tilde{u})(v) & \forall v \in V. \end{cases} \quad (3.3)$$

Further, we set for abbreviation

$$\delta u := \tilde{u} - u, \quad \delta z := \tilde{z} - z. \quad (3.4)$$

Then, the difference between a solution  $q$  of (1.4) and a solution  $q_h$  of (3.1) can be estimated by:

$$\gamma \|q - q_h\|_Q^2 \leq \inf_{(p_h, y_h, w_h) \in Q_h \times V_h \times V_h} \rho(p_h - q, \delta z - y_h, \delta u - w_h) + \mathcal{R}, \quad (3.5)$$

with the residual functional  $\rho$  defined in (3.2) and a cubic remainder term  $\mathcal{R}$  depending on the third derivatives of  $J$  and  $a$ .

In addition, there is a constant  $C$  such that the following lower bound holds:

$$\inf_{(p_h, y_h, w_h) \in Q_h \times V_h \times V_h} \rho(p_h - q, \delta z - y_h, \delta u - w_h) \leq C \|q - q_h\|_Q^2 + \mathcal{R}, \quad (3.6)$$

with a cubic remainder term  $\mathcal{R}$  depending on the third derivatives of  $J$  and  $a$ .

**Proof.** First, by Proposition 2.1 we have

$$\begin{aligned} \frac{\gamma}{2} \|q - q_h\|_Q^2 & \leq j(q_h) - j(q) = \{j(q_h) - j_h(q_h)\} - \{j(q) - j_h(q_h)\} \\ & = \{\varphi(1) - \varphi(0)\} - \{\psi(1) - \psi(0)\}, \end{aligned}$$

supposed we define (with  $q_s := q_h + s(q - q_h)$ ,  $\tilde{u}_s := u_h + s(\tilde{u} - u_h)$ , etc.)

$$\varphi(s) := \mathcal{L}(q_h, \tilde{u}_s, \tilde{z}_s), \quad \psi(s) := \mathcal{L}(q_s, u_s, z_s).$$

We then use the trapezoidal rule to find that

$$\begin{aligned} \frac{\gamma}{2} \|q - q_h\|_Q^2 &\leq \int_0^1 \{\varphi'(s) - \psi'(s)\} ds \\ &= \frac{1}{2} \{\varphi'(1) + \varphi'(0) - \psi'(1) - \psi'(0)\} + \mathcal{R}, \end{aligned}$$

with  $\mathcal{R} := \sup_{0 \leq s \leq 1} |\varphi'''(s) - \psi'''(s)|$ . Let us now examine the different terms. First, we have

$$\varphi'(1) = \mathcal{L}'_u(q_h, \tilde{u}, \tilde{z})(\tilde{u} - u_h) + \mathcal{L}'_z(q_h, \tilde{u}, \tilde{z})(\tilde{z} - z_h) = 0,$$

due to the definition of  $\tilde{u}$  and  $\tilde{z}$  as solutions to (3.3). Similarly we find

$$\psi'(1) = \mathcal{L}'_{(q,u,z)}(q, u, z)(q - q_h, u - u_h, z - z_h) = j'(q)(q - q_h)$$

due to Proposition 2.3. The continuous optimality condition yields  $\psi'(1) = 0$ . Next we have

$$\varphi'(0) = \mathcal{L}'_u(q_h, u_h, z_h)(\tilde{u} - u_h) + \mathcal{L}'_z(q_h, u_h, z_h)(\tilde{z} - z_h),$$

where the test functions  $u_h$  and  $z_h$  can be replaced by any discrete function. Finally we have

$$\psi'(0) = \mathcal{L}'_q(q_h, u_h, z_h)(q - q_h) + \mathcal{L}'_u(q_h, u_h, z_h)(u - u_h) + \mathcal{L}'_z(q_h, u_h, z_h)(z - z_h).$$

The same remark concerning the test functions, including this time  $q_h$  applies and we find (3.5).

For the lower bound, we remark that by convexity  $j(q) \geq j(q_h) + j'(q_h)(q - q_h)$ . Therefore, we have:

$$\begin{aligned} \frac{1}{2} \rho(q - p_h, \delta z - y_h, \delta u - w_h) &\leq j(q_h) - j(q) + \mathcal{R} \\ &\leq j'(q_h)(q_h - q) - j'(q)(q_h - q) + \mathcal{R}. \end{aligned}$$

Let  $C$  be the Lipschitz constant of  $\nabla j$  in a neighborhood containing  $q$  and  $q_h$ . Since  $j'(q_h)(q - q_h) - j'(q)(q - q_h) \leq C \|q - q_h\|^2$ , we find (3.6).  $\square$

We remark that there are two possible reasons for the weights  $\delta z - y_h$  and  $\delta u - w_h$  to be small: first  $\tilde{z}$  is close to  $z$  ( $\tilde{u}$  is close to  $u$ ), and second, the interpolation error is small.

Under appropriate assumptions on the state equation, Theorem 3.1 leads to an error estimate for all unknowns of the optimization problem. To this end we make the following regularity and continuity assumptions. We suppose that there exist  $\alpha > 0$  and  $C$  such that

$$\sup_{v \in V \setminus \{0\}} \frac{a(q, u_1)(v) - a(q, u_2)(v)}{\|v\|_V} \geq \alpha \|u_1 - u_2\|_V, \quad (3.7)$$

$$a(q_1, u)(v) - a(q_2, u)(v) \leq C \|q_1 - q_2\|_Q \|v\|_V, \quad (3.8)$$

$$a(q, u_1)(v) - a(q, u_2)(v) \leq C \|u_1 - u_2\|_V \|v\|_V. \quad (3.9)$$

We define the residual functional for the state equation  $\rho_{state} \in V^*$  by

$$\rho_{state}(v) := l(v) - a(q_h, u_h)(v). \quad (3.10)$$

**Theorem 3.2.** *In addition to the assumptions of Theorem 3.1, suppose that (3.7,3.8) are satisfied. Set for abbreviation*

$$\begin{aligned} \eta_q &:= \inf_{(p_h, y_h, w_h) \in Q_h \times V_h \times V_h} \sqrt{|\rho(p_h - q, \delta z - y_h, \delta u - w_h)|}, \\ \eta_u &:= \sup_{v \in V} \inf_{v_h \in V_h} \frac{\rho_{state}(v - v_h)}{\|v\|_V}. \end{aligned} \quad (3.11)$$

Then we have the following error bound

$$\|q - q_h\|_Q + \|u - u_h\|_V \leq C(\eta_q + \eta_u) + \mathcal{R}, \quad (3.12)$$

with a cubic remainder term  $\mathcal{R}$  and a constant  $C$  depending on  $\alpha$  and  $\gamma$ .

Suppose in addition (3.9). Then there exists a constant  $C$  such that

$$\eta_q + \eta_u \leq C(\|q - q_h\|_Q + \|u - u_h\|_V) + \mathcal{R}, \quad (3.13)$$

with a cubic remainder term  $\mathcal{R}$ .

**Proof.** First we note that by (3.7)

$$\begin{aligned} \alpha \|\tilde{u} - u_h\| &\leq \sup_{v \in V} \frac{a(q_h, \tilde{u})(v) - a(q_h, u_h)(v)}{\|v\|_V} \\ &= \sup_{v \in V} \inf_{v_h \in V_h} \frac{l(v - v_h) - a(q_h, u_h)(v - v_h)}{\|v\|_V} = \sup_{v \in V} \inf_{v_h \in V_h} \frac{\rho_{state}(v - v_h)}{\|v\|_V}. \end{aligned}$$

Next, we have using (3.8)

$$a(q, \tilde{u})(v) - a(q, u)(v) = a(q, \tilde{u})(v) - a(q_h, \tilde{u})(v) \leq C \|q - q_h\|_Q \|v\|_V,$$

such that again by (3.7)

$$\begin{aligned} \alpha \|\tilde{u} - u\| &\leq \sup_{v \in V} \frac{a(q, \tilde{u})(v) - a(q, u)(v)}{\|v\|_V} \\ &\leq C \|q - q_h\|_Q. \end{aligned}$$

For the lower bound, we first remark that by Theorem 3.1  $\eta_q \leq C \|q - q_h\|_Q$ . Next, since  $\rho_{state}(v) = a(q, u)(v) - a(q_h, u_h)(v) = a(q, u)(v) - a(q_h, u)(v) + a(q_h, u)(v) - a(q_h, u_h)(v)$ , we have with (3.9):

$$\rho_{state}(v) \leq C \|q - q_h\|_Q \|v\|_V + C \|u - u_h\|_V \|v\|_V.$$

The estimate (3.13) follows by the definition of  $\eta_u$ .  $\square$

The result of Theorem 3.2 is different from the estimates presented in [13], since the error in the Lagrange multiplier  $z$  is not involved in the right-hand side of the lower bound (3.13). This might be of interest in cases where regularity estimates for  $z$  are difficult to obtain.

The following corollaries are simple consequences of Theorem 3.1. They illustrate how the estimator may be used to control the two spaces individually. For brevity, we skip the statements of the corresponding lower bounds.

**Corollary 3.1.** (discrete control) Let the assumptions of Theorem 3.1 be fulfilled. In addition, assume that  $Q_h = Q$ . Then we have the error estimate:

$$\begin{cases} \gamma \|q - q_h\|_Q^2 \leq \{l(\delta z - y_h) - a(q_h, u_h)(\delta z - y_h)\} \\ + \{J'_u(q_h, u_h)(\delta u - w_h) - a'_u(q_h, u_h)(z_h, \delta u - w_h)\} + \mathcal{R}, \end{cases} \quad (3.14)$$

with arbitrary  $w_h, y_h \in V_h$  and a cubic remainder term  $\mathcal{R}$  depending on the third derivative of  $J$  and  $a$ .

**Corollary 3.2.** (discrete state) Let the assumptions of Theorem 3.1 be fulfilled. In addition, assume that  $V_h = V$ . Then we have the error estimate:

$$\gamma \|q - q_h\|_Q^2 \leq \mathcal{L}'_q(q_h, u_h, z_h)(q - p_h) + \mathcal{R}, \quad (3.15)$$

with arbitrary  $p_h \in Q_h$  and a cubic remainder term  $\mathcal{R}$  depending on the third derivative of  $J$  and  $a$ .

**Remark 3.1.** Let us write the error estimate (3.15) in the case of the Poisson problem (Example 2.4). Since  $\gamma = 1$  and  $\mathcal{R} = 0$ , we obtain

$$\|q - q_h\|_{H_0^1(\Omega)}^2 \leq \mathcal{L}'_q(q_h)(q - p_h) = \langle \Delta q_h + f, q - p_h \rangle_{H^{-1}(\Omega) \times H_0^1(\Omega)} \quad (3.16)$$

Now the techniques developed for a posteriori error estimation as residual error estimators or hierarchical estimators can be used to bound, respectively, approximate the term  $\|\Delta q_h + f\|_{H^{-1}(\Omega)}$ .

The error estimator presented in Theorem 3.1 requires the unknown variables  $u, z$ , as well as the solution of the continuous problem (3.3). In the algorithm we develop below, we approximate these functions by solutions of the same equations on a finer mesh, which we choose as a global refinement of the given one. We remark that the work required for solution of the additional equations (3.3) is low, since they correspond to the solution of the state and co-state equations on the new mesh with interpolated control. In an outer loop of mesh adaptation, they are computed in the first step of the inner optimization iteration on the new mesh.

Let us be more precise concerning the construction of the spaces. Let  $V_k \subset V_{k+1} \subset \dots$  be the sequence of finite-element spaces to be generated. For each space

$V_k$  we define a finer space  $W_k$  which is obtained from  $V_k$  by global refinement. We end up with the following sequence of spaces on which the discrete optimization problem is solved

$$\begin{array}{ccccccc} \dots & V_k & & V_{k+1} & & V_{k+2} & \dots \\ & \downarrow & \nearrow & \downarrow & \nearrow & & \\ \dots & W_k & & W_{k+1} & & W_{k+2} & \dots, \end{array}$$

where the north-east arrow at step  $k$  corresponds to a local refinement step based on information from the nested spaces  $V_k \subset W_k$ .  $V_{k+1}$  is obtained from  $V_k$  by local refinement, such that in general  $W_k \not\subset V_{k+1}$ .

This scheme can be implemented in the following way. Depending on whether the iteration counter is even or odd, we perform a global refinement step (corresponding to  $V_k \rightarrow W_k$ ) or do a local refinement of the one but last mesh using information from the last step (corresponding to  $V_k \rightarrow V_{k+1}$ ).

In order to evaluate the error estimator, we define approximations  $\tilde{u}_k \in W_k$  and  $\tilde{z}_k \in W_k$  of  $\tilde{u}$  and  $\tilde{z}$  appearing in (3.5). They are available without extra computation from the first step of the Newton iteration. We approximate the weighting function  $\delta u - w_h$  in (3.5) by  $(I - I_{V_k})(\tilde{u}_k - u_k)$  with a local interpolation interpolator  $I_{V_k} : W_k \rightarrow V_k$ , and similar for  $\delta z - y_h$ .

The same procedure is applied to the sequence of control spaces  $\{Q_k\}$ .

A typical algorithm reads as follows:

1. Let  $V_0$  and  $Q_0$  be first finite-element spaces and  $q_0 \in Q_0$ . Set  $k = 1$ .
2. Use a stopping criterion based on (2.21) or (2.22).
3. Solve the optimization problem on space  $V_k$  by a Newton iteration:
  - (a) set  $q_k^0 = q_{k-1}$ ,  $l = 0$ ,
  - (b) compute the gradient  $\nabla j(q_k^l)$  (by Proposition 2.3 this requires solution of the state and adjoint equations),
  - (c) solve the system
$$\nabla^2 j(q_k^l) \delta^l = -\nabla j(q_k^l) \quad (3.17)$$
by a conjugate gradient iteration. By Proposition 2.4 this requires solution of the linear equations (2.26).
  - (d) Set  $q_k^{l+1} = q_k^l + \delta^l$ , increase  $l$  and go to (b), unless a stopping criterion is satisfied.
4. Use the conjugate gradient iteration in 3(c) to estimate  $\gamma_k = \lambda_{\min}(\nabla^2 j_k(q_k))$ . Stop the iteration if  $\gamma_k \rightarrow 0$  is detected.
5. Set  $q_k = q_k^{L+1}$  where  $L$  is the number of Newton iterations. If  $k$  is odd do a global mesh refinement. Otherwise evaluate the error estimator and perform a local mesh refinement step. Increase  $k$  and go to 2.

The purpose of step (4) is to prevent that  $\gamma_k \rightarrow 0$ . In practice, we need stopping criteria in order to minimize the number of Newton steps and the number of inner CG iterations.

We remark that the algorithm can be modified in order to separate adaptation of the discrete spaces  $Q_h$  and  $V_h$ , based on corollaries 3.1 and 3.2.

Convergence of such type of automatic adaptive algorithms is far from obvious. We remark that if no coarsening is allowed, convergence can in general be ensured by a priori error analysis. However, proving optimal speed of convergence - that is the asymptotic behavior of error with respect to the number of unknowns - is a difficult task; see [8] for such an analysis for linear-quadratic optimal control problems discretized by wavelets, and [4] for the Poisson problem discretized by finite elements on triangular meshes.

#### 4. THE CASE OF CONTROL CONSTRAINTS

We discuss generalization of the a posteriori error estimate of Theorem 3.1 to the case of control constraints  $q \in Q_{ad}$ . Our aim is again to outline an approach on the abstract level.

We now consider the constrained minimization problem:

$$\begin{cases} \inf_{q \in Q_{ad}} J(q, u) & \text{subject to:} \\ a(q, u)(v) = l(v) & \forall v \in V. \end{cases} \quad (4.1)$$

Since  $Q_{ad}$  is closed and convex, there exists the projection operator  $P : Q \rightarrow Q_{ad}$  defined by  $\|Pq - q\|_Q \leq \|p - q\|$  for all  $p \in Q_{ad}$ . Unless  $Q_{ad}$  is a closed linear subspace,  $P$  is not differentiable; however, generalized differentiability properties of projectors are well-studied, see for example [11] and [14] for application of a generalized mean-value theorem to a priori error analysis of constrained optimal control problems.

We suppose that there exists an operator  $P'(q)$  such that a necessary first-order condition for (4.1) reads:

$$j'(q)(P'(q)p) = 0 \quad \forall p \in Q_{ad}. \quad (4.2)$$

Condition (4.2) is closely related to the variational inequality

$$j'(q)(p - q) \geq 0 \quad \forall p \in Q_{ad} \quad (4.3)$$

in case  $P'(q)p$  is the directional derivative of  $P$  at  $q$  in direction  $p$ . To this end, we recall that (4.3) is equivalent to the equation

$$q - P(q - \alpha \nabla j(q)) = 0 \quad (4.4)$$

for any  $\alpha > 0$ . In order to see the equivalence of (4.3) with (4.4), multiply (4.3) by  $\alpha > 0$  and add  $(p - q, p - q) \geq 0$  to get:

$$(p - (q - \alpha \nabla j(q)), p - q) \geq 0 \quad \forall p \in Q_{ad}. \quad (4.5)$$



A variational characterization of  $P$  reads  $(p - r, p - P(r)) \geq 0$  for all  $p \in Q_{ad}$ , and we obtain (4.4).

Next we note that  $q = P(q)$ . Then (4.4) becomes for  $\alpha > 0$

$$\frac{P(q) - P(q - \alpha \nabla j(q))}{\alpha} = 0,$$

and  $\alpha \rightarrow 0$  leads to (4.2), supposed  $(P(q) - P(q - \alpha \nabla j(q)) - \alpha P'(q)(\nabla j(q)))/\alpha \rightarrow 0$  in a suitable topology.

We consider discretization of (4.1) by

$$\begin{cases} \inf_{q_h \in Q_{ad,h}} J(q_h, u_h) & \text{subject to:} \\ a(q_h, u_h)(v) = l(v) & \forall v \in V. \end{cases} \quad (4.6)$$

A solution to (4.6) satisfies the discrete variational inequality:

$$j'_h(q_h)(p_h - q_h) \geq 0 \quad \forall p_h \in Q_{ad,h}. \quad (4.7)$$

In order to deal with the constraints, we use a special interpolation operator. We say that an interpolation operator  $I_h : Q \rightarrow Q_h$  respects the constraints if

$$I_h q \in Q_{ad,h} \quad \text{and} \quad (I - P'(q))(q - I_h q) = 0 \quad \forall q \in Q_{ad}. \quad (4.8)$$

In the following we summarize our results for a posteriori error estimation.

**Theorem 4.1.** *Same assumptions as Theorem 3.1, beside that  $Q_{ad} \subset Q$  is now a closed convex set with projector  $P$ . Suppose that (4.2) holds with an appropriate operator  $P'$ . Furthermore, we suppose that  $Q_{ad,h} \subset Q_{ad}$  and that there exists an interpolation operator  $I_h$  that respects the constraints.*

*Then, the difference between a solution  $q$  of (4.1) and a solution  $q_h$  of (4.6) can be estimated by:*

$$\gamma \|q - q_h\|_Q^2 \leq \inf_{(y_h, w_h) \in V_h \times V_h} \rho(P'(q)(I_h q - q), \delta z - y_h, \delta u - w_h) + \mathcal{R}, \quad (4.9)$$

*with the residual functional  $\rho$  defined in (3.2) and a cubic remainder term  $\mathcal{R}$  depending on the third derivatives of  $J$  and  $a$ .*

**Proof.** First we note that by the proof of Proposition 2.1 we have

$$\frac{\gamma}{2} \|q - q_h\|_Q^2 \leq j(q_h) - j(q) - j'(q)(q_h - q). \quad (4.10)$$

On the other hand the proof of Theorem 3.1 reveals that with arbitrary  $y_h, w_h \in V_h$

$$j(q_h) - j(q) - \frac{1}{2} j'(q)(q_h - q) = \rho(q_h - q), \delta z - y_h, \delta u - w_h) + \mathcal{R}. \quad (4.11)$$

Using the variational inequality (4.3) and its discrete analogon, we find:

$$\begin{aligned}
 \frac{\gamma}{2} \|q - q_h\|_Q^2 &\leq j(q_h) - j(q) - \frac{1}{2} j'(q)(q_h - q) - \frac{1}{2} j'(q)(q_h - q) \\
 &\leq \rho(q_h - q), \delta z - y_h, \delta u - w_h + \mathcal{R} \\
 &= \rho(p_h - q), \delta z - y_h, \delta u - w_h + j'_h(q_h)(q_h - p_h) + \mathcal{R} \\
 &\leq \rho(p_h - q), \delta z - y_h, \delta u - w_h + \mathcal{R}.
 \end{aligned}$$

It remains to choose  $p_h = I_h q$  and to use (4.8) in order to obtain the estimate (4.9).  $\square$

**Remark 4.1.** The additional assumptions on the projector and interpolation operator are not necessary, if we are only interested in the upper bound

$$\gamma \|q - q_h\|_Q^2 \leq \inf_{(p_h, y_h, w_h) \in Q_h \times V_h \times V_h} \rho(p_h - q, \delta z - y_h, \delta u - w_h) + \mathcal{R}. \quad (4.12)$$

This estimate is however not optimal since it does not use the constraints: residuals have also be evaluated on the 'active set' where (in case  $Q_{ad}$  is defined by inequalities) the inequality constraints become equalities.

In the following we consider briefly application of the estimate of Theorem 4.1 to the case of box-constraints  $Q_{ad} := \{q \in Q : \underline{q} \leq q \leq \bar{q}\}$ . Since the topology of  $Q$  is of great importance, we distinguish three further cases.

The first case is  $Q = \mathbb{R}^n$ . We denote by  $q(i)$  the  $i$ -th component of  $q \in Q$ . We have  $P(q)(i) = \max(\min(q(i), \underline{q}), \bar{q})$ . The inactive and active sets are defined as the set of indices  $\mathbb{I}_q := \{i \in \{1, \dots, n\} : \underline{q} < q(i) < \bar{q}\}$  and  $\mathbb{A}_q := \{1, \dots, n\} \setminus \mathbb{I}_q$ . Finally we set  $P'(q)p(i) = p(i)$  if  $i \in \mathbb{I}_q$  and  $P'(q)p(i) = 0$  otherwise. It is immediate to see that the variational inequality (4.3) implies (4.2). Let  $Q_h = \mathbb{R}^m$  with  $m < n$ . The construction of an interpolation operator which satisfies (4.8) is trivial.

The second case is  $Q = L^2(\Omega)$ , see Example 2.3. Now  $P$  is defined by  $P(q)(x) = \max(\min(q(x), \underline{q}), \bar{q})$ . The inactive and active sets are defined by  $\mathbb{I}_q := \{x \in \Omega : \underline{q} < q(x) < \bar{q}\}$  and  $\mathbb{A}_q := \Omega \setminus \mathbb{I}_q$ . We set  $P'(q)(p) = \chi_{\mathbb{I}_q} p$  where  $\chi_A$  is the characteristic function of the set  $A \subset \Omega$ . Again, (4.2) is satisfied. However, the construction of an interpolation operator respecting the constraints is more technical and largely depends on the choice of  $Q_h$  and  $Q_{ad,h}$ . We note that (4.8) reads in this case:

$$x \in \mathbb{A}_q \implies I_h q(x) = q(x). \quad (4.13)$$

For related questions in the context of a priori error analysis for constrained optimal control problems see [6,18].

The third case is the bilateral obstacle problem of Example 2.4. This allows us to compare our approach with the recent development of a posteriori error analysis for the obstacle problem, see [7,16,19] and the literature cited therein. Let  $\tilde{q} \in H_0^1(\Omega)$  be the solution of the Poisson problem without obstacle,  $-\Delta \tilde{q} = f$ . Then, the obstacle

problem is equivalent to minimization of  $\|q - \tilde{q}\|_Q^2$ ; therefore  $q$  is the projection of  $\tilde{q}$  in the  $H_0^1(\Omega)$  scalar product on  $Q_{ad}$ . We use the same definition of inactive and active sets as in the last example. In the context of the obstacle problem, the active set is called contact set. An appropriate candidate for  $P'(q)$  is the projector on the linear subspace  $Q_q := \{q \in Q : q(x) = 0 \text{ a.e. } x \in \mathbb{A}_q\}$ . The equation defining this linear projector reads: For given  $\tilde{p} \in Q$  find  $p = P'(q)\tilde{p} \in Q_q$  such that

$$(\nabla p, \nabla \delta) = (\nabla \tilde{p}, \nabla \delta) \quad \forall \delta \in Q_q. \quad (4.14)$$

The construction of an appropriate interpolation operator again leads to (4.13).

We rewrite now the estimator of Theorem 4.1:

$$\begin{aligned} j'(q_h)(P'(q)(q - I_h q)) &= (q_h - \tilde{q}, P'(q)(q - I_h q))_Q \\ &= -\langle \Delta q_h + f, P'(q)(q - I_h q) \rangle_{H^{-1}(\Omega) \times H^1(\Omega)}. \end{aligned}$$

The last identity could now be used to obtain a residual error estimator which vanishes on the contact set. There are similarities to the a posteriori error estimator of [19], which however deals in addition with variable obstacles, the case  $Q_{ad,h} \not\subset Q_{ad}$ , and does not use knowledge about the position of the continuous contact zone.

## 5. NUMERICAL EXAMPLE

We consider a linear-quadratic model problem. The cost is searched for in  $Q = H^1(\Omega)$ ,  $\Omega = ]-3, 3[ \times ]-6, 6[ \setminus ]-1, 1[ \times ]-2, 2[$ , and the cost functional is given by

$$J(q, u) = \frac{1}{2} \sum_{i=1}^4 |u(x_i) - \bar{u}_i|^2 + \frac{\alpha}{2} \|q\|_Q^2, \quad (5.1)$$

with  $\alpha = 10^{-4}$ ,  $\{x_i\} = \{(-2, -4), (-2, -2), (-2, 2), (-2, 4)\}$ , and  $\bar{u}_i = 0.5$  for  $1 \leq i \leq 4$ . The partial differential equation to be solved is:

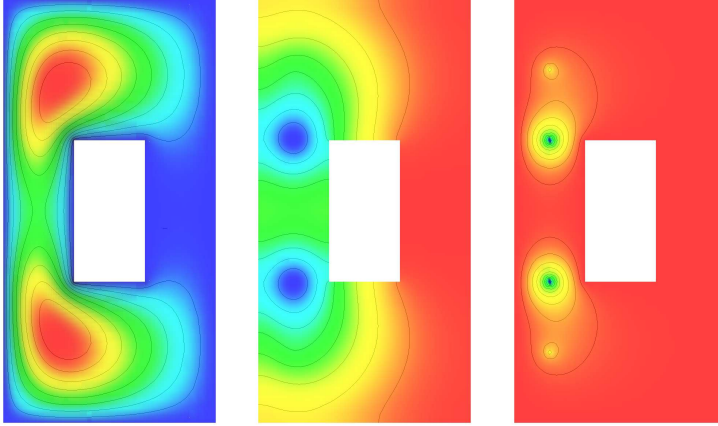
$$-\Delta u = -q \quad \text{in } \Omega, \quad u|_{\partial\Omega} = 0. \quad (5.2)$$

The value of the cost functional at the optimum of this largely underdetermined problem is  $J \approx 5.9e - 04$ .

The problem is interesting since its solution suffers from different singularities: first the domain has re-entrant corners such that  $u \notin H^2(\Omega)$ . The same singularity deteriorates the regularity of the Lagrange multiplier  $z$ . In addition, since  $z$  is the solution of

$$-\Delta z = \sum_{i=1}^4 |u(x_i) - \bar{u}_i| \delta_{x_i} \quad \text{in } \Omega, \quad z|_{\partial\Omega} = 0, \quad (5.3)$$

the Dirac measures on the right-hand side of (5.3) lead to additional singularities. However, due to the smallness of  $\alpha$  they have small coefficients.



**Figure 1.** Optimal solution:  $u$ ,  $q$ , and  $z$ .

For simplicity we use bilinear finite elements for the discretization of all three variables on the same mesh.

Pictures of the optimal displacement  $u$ , right-hand side  $q$  and Lagrange multiplier  $z$  are shown in Figure 1. We use this example to compare the error estimator developed in this paper with a possibly more direct error estimator which measures the residuals of each equation separately.

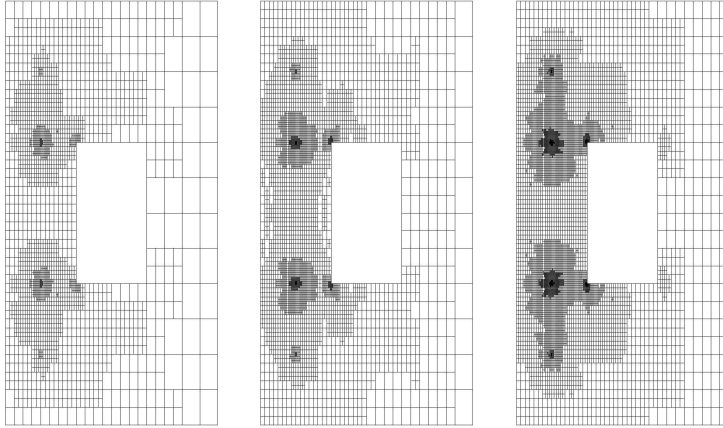
We next describe some details of the employed algorithm. As explained before, we use two nested meshes - which lead to two finite element spaces  $V_h$  and  $V_{h/2}$  (corresponding to  $V_k$  and  $W_k$  in the description of the algorithm in Section section:NoConstraints) of dimension approximately  $N$  and  $4N$  (in two space dimensions)- in order to evaluate the estimator (3.5). We define the residual vectors as

$$\begin{aligned}\rho_u(i) &:= l(\varphi_i) - a(q_h, u_h)(\varphi_i) & \varphi_i \in V_{h/2}, \\ \rho_z(i) &:= J'(q_h, u_h)(\varphi_i) - a'(q_h, u_h)(z_h, \varphi_i) & \varphi_i \in V_{h/2}, \\ \rho_q(i) &:= \mathcal{L}'_q(q_h, u_h, z_h)(\varphi_i) & \varphi_i \in V_{h/2}.\end{aligned}$$

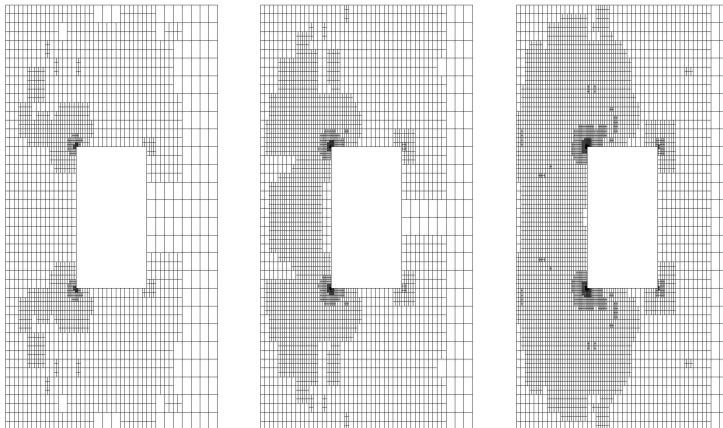
In order to obtain an approximation for  $\delta u$  and  $\delta z$  defined in (3.4) we take the difference between  $u_{h/2}$  (the final iterate of the optimization algorithm on  $V_{h/2}$ ) and the solution of the state equation corresponding to control  $q_h$  on mesh  $V_{h/2}$  (the first iterate of the optimization algorithm on  $V_{h/2}$ ). Further more we subtract its interpolation on  $V_h$ . Let us denote by  $du$  and  $dz$  the resulting vectors. In order to obtain  $dq$  we subtract the coarse grid interpolation from the optimal control  $q_{h/2}$ . Now we define the estimator  $\eta$  by

$$\eta(i) := |\rho_u(i)dz(i) + \rho_z(i)du(i) + \rho_q(i)dq(i)|. \quad (5.4)$$

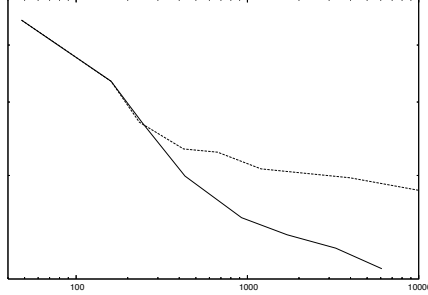
Next the vector is sorted and we generate a new mesh such that at least 90% of the total amount of corresponding vertices are flagged to be refined. A typical sequence of meshes generated in this way is shown in Figure 2.



**Figure 2.** Meshes obtained with estimator  $\eta$  ( $N = 2934, 5588, 9948$ ).



**Figure 3.** Meshes obtained with estimator  $\tilde{\eta}$  ( $N = 2810, 4500, 6652$ ).



**Figure 4.** Error decay for the two estimators  $\eta$  (solid line) and  $\tilde{\eta}$  (dashed line) (log-log-scale).

As a competitor we consider the estimator  $\tilde{\eta}$  defined by

$$\tilde{\eta}(i) := |\rho_u(i)du(i)| + |\rho_z(i)dz(i)| + |\rho_q(i)dq(i)|. \quad (5.5)$$

Note that the first two terms correspond to an energy estimator of the state and of the adjoint equation, respectively. We use the same procedure for mesh refinement, except that values  $\eta(i)$  are replaced by  $\tilde{\eta}(i)$ . A sequence of meshes (iterates 6,7,8 in the mesh refinement iteration) generated by  $\tilde{\eta}$  is shown in Figure 3. As compared to Figure 2, it is clear that  $\tilde{\eta}$  is much more attracted by the corner singularities of the state equation, whereas the refinement according to  $\eta$  emphasizes the points of evaluation. Changing the position of the absolute values in (5.5) or suppressing the term involving  $\rho_q$  did not lead to any remarkable change.

Finally, the numerical efficiency of the two estimators is compared in Figure 4, where the dependence of estimated error on the number of unknowns is shown in log-log scale. The estimated error is the difference to a solution on the finest accessible grid using alternatively local and global refinement, measured in the  $L^2$ -norm. The adaptive refinement seems to cure the loss of regularity, since the error behaves close to a second-order method ( $e \sim N^{-1}$  in two dimensions).

## 6. CONCLUSIONS

We have developed an a posteriori error estimator for the control of the error in the control variable measured in a norm which is related to a strong form of a sufficient second-order conditions, hypothesis **(H1)**. In addition, we have required the discrete solution to be close enough to the continuous one, hypothesis **(H2)**. This assumption is automatically fulfilled in the convex case, for example linear-quadratic problems. Under these conditions we have derived an a posteriori error estimator, including the case of control constraints, which requires only marginal extra cost in an adaptive optimization algorithm. It has been shown that an additional assumption on the state operator leads to an a posteriori error estimate for the state variable,

too. Generalization of the error estimate to the control constrained case has been outlined.

## REFERENCES

1. I. Babuška and W.C. Rheinboldt. Error estimates for adaptive finite element computations. *SIAM J. Numer. Anal.*, 15:736–754, 1978.
2. R. Becker, H. Kapp, and R. Rannacher. Adaptive finite element methods for optimal control of partial differential equations: Basic concepts. *SIAM J. Control Optimization*, 39(1):113–132, 2000.
3. R. Becker and B. Vexler. A posteriori error estimation for finite element discretizations of parameter identification problems. *Numer. Math.*, 96(3):435–459, 2004.
4. P. Binev, W. Dahmen, and R. DeVore. Adaptive finite element methods with convergence rates. *Numerische Mathematik*, 2004.
5. J.F. Bonnans. Second order analysis for control constrained optimal control problems of semi-linear elliptic systems. *Applied Math. Optimization*, 38(3):303–325, 1998.
6. E. Casas and F. Tröltzsch. Error estimates for linear-quadratic elliptic control problems. In *Barbu, VioREL (ed.) et al., Analysis and optimization of differential systems. IFIP TC7/WG 7.2 international working conference, Constanta, Romania, September 10-14, 2002. Boston, MA: Kluwer Academic Publishers. 89-100*. 2003.
7. Z. Chen and R.H. Nochetto. Residual type a posteriori error estimates for elliptic obstacle problems. *Numer. Math.*, 84(4):527–548, 2000.
8. W. Dahmen and A. Kunoth. Adaptive wavelet methods for linear-quadratic elliptic control problems: convergence rates. *SIAM J. Control Optimization*, 43(5):1640–1675, 2005.
9. K. Eriksson, D. Estep, P. Hansbo, and C. Johnson. Introduction to adaptive methods for differential equations. In A. Iserles, editor, *Acta Numerica 1995*, pages 105–158. Cambridge University Press., 1995.
10. V. Girault and P.-A. Raviart. *Finite Elements for the Navier Stokes Equations*. Springer, Berlin, 1986.
11. A. Haraux. How to differentiate the projection on a convex set in hilbert space. some applications to variational inequalities. *J. Math. Soc. Japan*, 29:615–631, 1977.
12. R. Li, W. Liu, H. Ma, and T. Tang. Adaptive finite element approximation for distributed elliptic optimal control problems. *SIAM J. Control Optimization*, 41(5):1321–1349, 2002.
13. W. Liu and N. Yan. A posteriori error estimates for convex boundary control problems. *SIAM J. Numer. Anal.*, 39(1):73–99, 2001.
14. K. Malanowski. Convergence of approximations vs. regularity of solutions for convex, control-constrained optimal-control problems. *Appl. Math. Optimization*, 8:69–95, 1981.
15. H. Maurer and J. Zowe. First and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems. *Math. Program.*, 16:98–110, 1979.
16. R.H. Nochetto, K.G. Siebert, and A. Veiser. Fully localized a posteriori error estimators and barrier sets for contact problems. *SIAM J. Numer. Anal.*, 42(5):2118–2135, 2005.
17. E. Polak. *Optimization. Algorithms and consistent approximations*. Number 124 in Applied Mathematical Sciences. Springer, 1997.
18. A. Rösch. Error estimates for parabolic optimal control problems with control constraints. *ZAA*, 23:353–376, 2004.

19. A. Veiser. Efficient and reliable a posteriori error estimators for elliptic obstacle problems. *SIAM J. Numer. Anal.*, 39(1):146–167, 2001.
20. R. Verfürth. *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. John Wiley/Teubner, New York-Stuttgart, 1996.