SIAM J. CONTROL OPTIM.
Vol. 47, No. 1, pp. 509–534

ADAPTIVE FINITE ELEMENTS FOR ELLIPTIC OPTIMIZATION PROBLEMS WITH CONTROL CONSTRAINTS*

B. VEXLER† AND W. WOLLNER‡

Abstract. In this paper we develop a posteriori error estimates for finite element discretization of elliptic optimization problems with pointwise inequality constraints on the control variable. We derive error estimators for assessing the discretization error with respect to the cost functional as well as with respect to a given quantity of interest. These error estimators provide quantitative information about the discretization error and guide an adaptive mesh refinement algorithm allowing for substantial saving in degrees of freedom. The behavior of the method is demonstrated on numerical examples.

Key words. mesh adaptivity, optimal control, a posteriori error estimates, finite element method, quantity of interest, pointwise inequality constraints

AMS subject classifications. 65N50, 65N30, 65K10

DOI. 10.1137/070683416

1. Introduction.

In this paper we develop a posteriori error estimates for finite element approximations of optimization problems governed by elliptic partial differential equations. We discuss this question in a general manner, including the consideration of optimal control and parameter identification problems with control constraints given through a closed convex admissible set. The derived error estimates have the goal of guiding an adaptive mesh refinement algorithm for finding economical meshes for the optimization problem under consideration.

The use of adaptive techniques based on a posteriori error estimation is well accepted in the context of finite element discretization of partial differential equations; see, e.g., [6, 13, 35]. To our knowledge there are only a few results published on adaptive finite elements for optimization problems; see [2, 17, 20, 23, 25, 27, 4, 7, 8, 30].

In articles [17, 20, 23, 25, 27] the authors provide a posteriori error estimates for elliptic optimal control problems with distributed or Neumann control subject to box constraints. These estimates assess the error in the control, state, and the adjoint variable with respect to the natural norms of the corresponding spaces. In [2] another approach for the estimation of the error with respect to the norm of the control space is presented. In [17] convergence of an adaptive algorithm for a control constrained optimal control problem is shown.

However, in many applications, the error in global norms does not provide a useful error bound for the error in the quantity of physical interest. The a posteriori estimators derived in this paper grant access to the error with respect to given functionals.

*Received by the editors February 22, 2007; accepted for publication (in revised form) August 10, 2007; published electronically February 1, 2008.
http://www.siam.org/journals/sicon/47-1/68341.html

†Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenberger Straße 69, 4040 Linz, Austria (boris.vexler@oeaw.ac.at). This author's research was partially supported by the Austrian Science Fund FWF project P18971-N18 "Numerical analysis and discretization strategies for optimal control problems with singularities."

‡Institut für Angewandte Mathematik, Ruprecht-Karls-Universität Heidelberg, INF 294, 69120 Heidelberg, Germany (winnifried.wollner@iwr.uni-heidelberg.de). This author's research was supported by the DFG priority program "Optimization with Partial Differential Equations."

In [4, 6] the authors present a general concept for a posteriori estimation of the discretization error with respect to the cost functional in the context of optimal control problems. In articles [7, 8] the authors have extended this approach to the estimation of the discretization error with respect to an arbitrary functional depending on both the control and the state variable, so-called *quantity of interest*. This allowed, among other things, the treatment of parameter identification and model calibration problems. However, in all these publications, the control variable was searched for in a Hilbert space Q without additional (inequality) constraints. Therefore the main contribution of this work is the extension of these techniques to the case of optimization problems with additional control constraints given through a closed convex admissible set $Q_{\mathrm{ad}} \subset Q$. In the majority of practical cases this admissible set is described by inequality control constraints of box type $q_- \leq q(x) \leq q_+$. Therefore we will concentrate on this case, although our techniques may also be extended to the consideration of more general admissible sets Q_{ad}.

In this paper we consider optimization problems governed by (nonlinear) partial differential equations. The aim is to minimize a given cost functional $J(q,u)$ which depends on the state variable $u \in V$ and the control variable $q \in Q$, with Hilbert spaces V and Q. These variables have to satisfy the state equation

$$A(q,u) = f, \tag{1.1}$$

where A denotes a (nonlinear) differential operator and f represents the given data. The optimization problem is then formulated as follows:

$$\begin{cases} \text{Minimize } J(q,u), \quad u \in V,\, q \in Q_{\mathrm{ad}}, \\ A(q,u) = f. \end{cases} \tag{1.2}$$

Constraints on the control are incorporated via the definition of the closed and convex set Q_{ad} representing the set of admissible controls.

For numerical treatment this infinite dimensional optimization problem is discretized in virtue of finite element methods; see the discussion in section 3. Let the solution to the discretized problem be denoted by (q_h, u_h). Our aim is to derive a posteriori error estimates for the error between the solutions to the continuous and the discrete problem. A crucial point for our error analysis is the choice of a quantity, which describes the goal of the computation. If this quantity coincides with the cost functional, we have to estimate the error

$$J(q,u) - J(q_h, u_h).$$

In a more general case, we suppose $I\colon Q \times V \to \mathbb{R}$ to be a given functional describing the quantity of interest. Then the error to be estimated is

$$I(q,u) - I(q_h, u_h).$$

The consideration of quantities of interest is important, for instance, in the context of parameter identification and model calibration problem; see [8] for an application of this concept to an optimization problem from computational fluid dynamics.

To the authors' knowledge this is the first article providing a posteriori error estimates with respect to a given functional for optimization problems with partial differential equations and subject to control constraints.

The paper is organized as follows. In the next section we describe the optimization problem under consideration, discuss necessary optimality conditions, and sketch the solution algorithm on the continuous level. In section 3 we describe the discretization of the optimization problem in virtue of finite element methods. Section 4 is devoted to a posteriori error estimation. In sections 4.1 and 4.2 we derive two different error estimates for the error with respect to the cost functional J. The first error estimator is based on the optimality system involving a variational inequality, whereas the second one exploits Lagrange multipliers for the treatment of inequality constraints. Due to the fact that the optimal control q is not expected to be sufficiently smooth (due to inequality constraints), the approximation of (interpolation) weights involved in the error estimator cannot be treated in a usual way. To overcome this difficulty we exploit the projection formula (2.9) from the optimality conditions and propose an approximation on the (interpolation) weights using a postprocessing step (4.7), which is motivated by the considerations in [31]. In section 4.3 we provide an error estimator with respect to a given quantity of interest. To this end we utilize an additional (dual) linear-quadratic optimal control problem describing the sensitivity with respect to the quantity of interest. In the last section we present numerical examples to illustrate the behavior of our method.

2. Optimization problem. In this section we give a precise formulation of the optimization problem under consideration and describe necessary optimality conditions and the solution algorithm.

In order to deal with different types of optimization problems simultaneously, we seek the control variable q in the Hilbert space $Q = L^2(\omega)$ with scalar product $(\cdot,\cdot)$ and norm $\|\cdot\|$. Typically, ω is a subset of the computational domain Ω or a subset of its boundary $\partial\Omega$. The case of finite dimensional controls is realized by choosing $\omega = \{1, 2, \dots, n\}$ resulting in $Q \cong \mathbb{R}^n$.

Throughout this paper we suppose that the state equation (1.1) for $u \in V$ is given in a weak form:

$$a(q,u)(\varphi) = f(\varphi) \quad \forall \varphi \in V, \tag{2.1}$$

where $a\colon Q \times V \times V \to \mathbb{R}$ is a four times directional differentiable form which is linear in the third argument and f is in the dual space V'. A possible choice for this space is $V = H^1(\Omega)$, or $V = H_0^1(\Omega)$, or a direct product of such spaces. In the presence of inhomogeneous Dirichlet boundary conditions, one seeks the state variable u in $\hat{u} + V$, where $\hat{u}$ represents the boundary data. However, for clarity of notation, we assume throughout that $\hat{u} = 0$.

Remark 2.1. Throughout this paper we use two pairs of parentheses after a form to indicate that the form is linear in all variables enclosed by the second pair of parentheses, as seen in (2.1) for $a(\cdot,\cdot)(\cdot)$.

The cost function is given by

$$J(q,u) = J_1(u) + \frac{\alpha}{2}\|q\|^2\,, \tag{2.2}$$

where J_1 is a four times directionally differentiable operator on V and $\alpha > 0$. Let the admissible set Q_{ad} be given through box constraints on q, i.e.,

$$Q_{\text{ad}} = \{q \in Q \mid q_- \le q(x) \le q_+ \text{ a.e. on } \omega\}, \tag{2.3}$$

with bounds $q_-, q_+ \in \mathbb{R} \cup \{\pm\infty\}$ and $q_- < q_+$.

Now we are able to formulate the optimization problem as

$$\text{Minimize } J(q,u)\,, \quad u \in V,\, q \in Q_{\text{ad}}\,, \quad \text{subject to (2.1).} \tag{2.4}$$

Remark 2.2. The choice of constant bounds $q_-, q_+ \in \mathbb{R}\cup\{\pm\infty\}$ is not a limitation, since one can transform an optimal control problem with bounds $q_-, q_+ \in Q$ into an equivalent one with constant bounds for the control.

To shorten notation we introduce the space $\mathcal{X}$ and the admissible set $\mathcal{X}_{\text{ad}}$ by

$$\mathcal{X} = Q \times V \times V, \tag{2.5}$$

$$\mathcal{X}_{\text{ad}} = Q_{\text{ad}} \times V \times V. \tag{2.6}$$

In addition we shall write $\xi = (q, u, z)$ for a vector in $\mathcal{X}$ or $\mathcal{X}_{\text{ad}}$, where z will denote an adjoint state.

Throughout the paper we assume that the problem (2.4) admits a solution. Conditions ensuring the existence of solutions to optimal control problems may, for instance, be found in [16, 26, 34]. We shall especially assume that the primal and dual equations associated with (2.4) are solvable for every given $q \in Q$.

To establish an optimality system, we introduce the Lagrangian $\mathcal{L}\colon \mathcal{X} \to \mathbb{R}$ as follows:

$$\mathcal{L}(\xi) = J_1(u) + \frac{\alpha}{2}\|q\|^2 + f(z) - a(q,u)(z),$$

where z denotes the dual variable. Due to the convexity of the admissible set Q_{ad}, the first-order necessary optimality condition for $(q, u) \in Q_{\text{ad}} \times V$ reads as follows:

There exists $z \in V$ such that the triple $\xi = (q, u, z) \in \mathcal{X}_{\text{ad}}$ satisfies

$$\mathcal{L}'_u(\xi)(\delta u) = 0 \quad \forall \delta u \in V, \tag{2.7a}$$

$$\mathcal{L}'_q(\xi)(\delta q - q) \geq 0 \quad \forall \delta q \in Q_{\text{ad}}, \tag{2.7b}$$

$$\mathcal{L}'_z(\xi)(\delta z) = 0 \quad \forall \delta z \in V. \tag{2.7c}$$

This system can be stated explicitly in the following form:

$$J_1'(u)(\delta u) - a'_u(q,u)(\delta u, z) = 0 \quad \forall \delta u \in V, \tag{2.8a}$$

$$\alpha(q, \delta q - q) - a'_q(q,u)(\delta q - q, z) \geq 0 \quad \forall \delta q \in Q_{\text{ad}}, \tag{2.8b}$$

$$f(\delta z) - a(q,u)(\delta z) = 0 \quad \forall \delta z \in V. \tag{2.8c}$$

We introduce a projection operator $\mathcal{P}_{Q_{\text{ad}}}\colon Q \to Q_{\text{ad}}$ by

$$\mathcal{P}_{Q_{\text{ad}}}(p) = \max\big(q_-, \min(p, q_+)\big)$$

pointwise a.e. This allows us to rewrite variational inequality (2.8b) (see, e.g., [34]) as

$$q = \mathcal{P}_{Q_{\text{ad}}}\left(\frac{1}{\alpha} a'_q(q,u)(\cdot, z)\right), \tag{2.9}$$

where $a'_q(u, q)(\cdot, z)$ is understood as a Riesz representative of a linear functional on Q.

For a solution (q, u) of (2.4) we introduce active sets ω_- and ω_+ as follows:

$$\omega_- = \{x \in \omega \mid q(x) = q_-\}, \tag{2.10}$$

$$\omega_+ = \{x \in \omega \mid q(x) = q_+\}. \tag{2.11}$$

Let $\xi \in \mathcal{X}$ be a solution to (2.7); then we introduce an additional Lagrange multiplier $\mu \in Q$ by the following identification:

$$(\mu, \delta q) = -\alpha(q, \delta q) + a'_q(q, u)(\delta q, z) = -\mathcal{L}'_q(\xi)(\delta q) \quad \forall \delta q \in Q\,. \tag{2.12}$$

The variational inequality (2.8b) or the projection formula (2.9) are known to be equivalent to the following conditions:

$$\mu(x) \le 0 \quad \text{a.e. on } \omega_-\,, \tag{2.13a}$$

$$\mu(x) \ge 0 \quad \text{a.e. on } \omega_+\,, \tag{2.13b}$$

$$\mu(x) = 0 \quad \text{a.e. on } \omega \setminus (\omega_- \cup \omega_+)\,. \tag{2.13c}$$

Using this representation of the optimality condition (2.8b) we apply nonlinear primal dual active set strategy (see, e.g., [9, 24]) to solve (2.4). In the following we sketch the corresponding algorithm on the continuous level.

Nonlinear primal-dual active set strategy

1. Choose initial guess q^0, μ^0 and $c > 0$ and set $n = 1$.
2. While not converged
3. Determine the active sets ω^n_+ and ω^n_-:

$$\omega^n_- = \{x \in \omega \,|\, q^{n-1}(x) + \mu^{n-1}(x)/c - q_- \le 0\},$$
$$\omega^n_+ = \{x \in \omega \,|\, q^{n-1}(x) + \mu^{n-1}(x)/c - q_+ \ge 0\}.$$

4. Solve the equality-constrained optimization problem

$$\text{Minimize} \quad J_1(u^n) + \frac{\alpha}{2}\|q^n\|^2, \quad u^n \in V,\, q^n \in Q,$$

subject to (2.1) and

$$q^n(x) = q_- \text{ on } \omega^n_-\,, \qquad q^n(x) = q_+ \text{ on } \omega^n_+\,.$$

5. Set

$$\mu^n = -\alpha q^n + a'_q(q^n, u^n)(\cdot, z^n)$$

with adjoint variable z^n.

6. Set $n = n + 1$ and go to 2.

Remark 2.3. The convergence in step 2 can be determined conveniently from agreement of the active sets in two consecutive iterations.

Remark 2.4. The algorithm above is known to be globally convergent for a class of optimal control problems if α is sufficiently large; see, e.g., [9, 24]. Moreover, local superlinear convergence can be shown; see, e.g., [21].

In our practical realization, the equality-constrained optimization problem in step 4 is solved by Newton's method on the control space without assembling the Hessian. The finite element discretization of the optimization problem, described in the next section, allows us to directly translate these algorithms onto the discrete level.

As we will encounter some trouble with the variational inequality in the necessary optimality condition (2.8) due to missing Galerkin orthogonality, we consider in addition the full Lagrangian $\tilde{\mathcal{L}}\colon \mathcal{X} \times Q \times Q \to \mathbb{R}$ which is given by

$$\tilde{\mathcal{L}}(\chi) = \mathcal{L}(\xi) + (\mu^-, q_- - q) + (\mu^+, q - q_+),$$

with $\chi = (\xi, \mu^-, \mu^+) = (q, u, z, \mu^-, \mu^+) \in \mathcal{X} \times Q \times Q$, where μ^- and μ^+ denote the variables corresponding to Lagrange multipliers for the inequality constraints. To shorten notation we introduce the abbreviation

$$\mathcal{Y} = \mathcal{X} \times Q \times Q. \tag{2.14}$$

Using the subspaces

$$\begin{aligned} Q_- &= \{r \in Q \,|\, r = 0 \text{ a.e. on } \omega \setminus \omega_-\}, \\ Q_+ &= \{r \in Q \,|\, r = 0 \text{ a.e. on } \omega \setminus \omega_+\}, \end{aligned}$$

we introduce

$$\mathcal{Y}_{\text{ad}} = \mathcal{X}_{\text{ad}} \times Q_- \times Q_+, \tag{2.15}$$

$$\tilde{\mathcal{Y}}_{\text{ad}} = \mathcal{X} \times Q_- \times Q_+ \tag{2.16}$$

and see that the following equality holds for all $\chi \in \mathcal{Y}_{\text{ad}}$:

$$\mathcal{L}(\xi) = \tilde{\mathcal{L}}(\chi). \tag{2.17}$$

We can rewrite the first-order necessary optimality condition for $(q, u) \in Q_{\text{ad}} \times V$ equivalently as follows (cf. [34]):

There exist $z \in V$, $\mu^- \in Q_-$, $\mu^+ \in Q_+$ such that the following conditions hold for $\chi = (q, u, z, \mu^-, \mu^+) \in \mathcal{Y}_{\text{ad}}$:

$$\tilde{\mathcal{L}}'_u(\chi)(\delta u) = 0 \quad \forall \delta u \in V, \tag{2.18a}$$

$$\tilde{\mathcal{L}}'_q(\chi)(\delta q) = 0 \quad \forall \delta q \in Q, \tag{2.18b}$$

$$\tilde{\mathcal{L}}'_z(\chi)(\delta z) = 0 \quad \forall \delta z \in V, \tag{2.18c}$$

$$\tilde{\mathcal{L}}'_{\mu^-}(\chi)(\delta \mu^-) = 0 \quad \forall \delta \mu^- \in Q_-, \tag{2.18d}$$

$$\tilde{\mathcal{L}}'_{\mu^+}(\chi)(\delta \mu^+) = 0 \quad \forall \delta \mu^+ \in Q_+, \tag{2.18e}$$

$$\mu^+, \mu^- \geq 0 \quad \text{a.e. on } \omega. \tag{2.18f}$$

It is easy to verify that the Lagrange multipliers μ^+ and μ^- are given as the positive and negative part of the Lagrange multiplier μ from (2.12); cf. [34].

Note that (2.18d), (2.18e) are equivalent to the complementarity conditions

$$\mu^-(q_- - q) = \mu^+(q - q_+) = 0 \quad \text{a.e. on } \omega. \tag{2.19}$$

For later use we recall a second-order sufficient optimality condition.

LEMMA 2.1 (sufficient optimality condition). *Let $\xi = (q, u, z) \in \mathcal{X}_{ad}$ satisfy the first-order necessary condition* (2.7a)–(2.7c) *of optimization problem* (2.4)*. Moreover, let $z \mapsto a'_u(q, u)(\cdot, z) : V \to V'$ be surjective. If there exists $\rho > 0$ such that*

$$(\delta q \,,\, \delta u) \begin{bmatrix} \mathcal{L}''_{qq}(\xi)(\cdot,\cdot) & \mathcal{L}''_{qu}(\xi)(\cdot,\cdot) \\ \mathcal{L}''_{uq}(\xi)(\cdot,\cdot) & \mathcal{L}''_{uu}(\xi)(\cdot,\cdot) \end{bmatrix} \begin{pmatrix} \delta q \\ \delta u \end{pmatrix} \geq \rho \left(\|\delta u\|_V^2 + \|\delta q\|_Q^2 \right) \tag{2.20}$$

holds for all $(\delta q, \delta u)$ satisfying the linear (tangent) partial differential equation

$$a'_u(q, u)(\delta u, \varphi) + a'_q(q, u)(\delta q, \varphi) = 0 \quad \forall \varphi \in V, \tag{2.21}$$

then (q, u) is a (strict) local solution to the optimization problem (2.4).

We refer the reader to [29] for the proof.

Remark 2.5. Throughout the paper we exploit only first-order information. This means that the error estimators proposed in section 4 are applicable to all solutions of the optimality system (2.8) or (2.18), respectively.

For the convenience of the reader we list the assumptions made in the preceding section.

Assumption 1. The optimization problem (2.4) possesses a solution (q, u). In addition there exists $z \in V$ such that the first-order necessary conditions (2.8) are fulfilled by the triple (q, u, z).

Remark 2.6. It is sufficient for the existence of z in the preceding assumption if the mapping $z \mapsto a'_u(q, u)(\cdot, z)$ is surjective onto V'. This is one of the requirements in Lemma 2.1 and is fulfilled by all examples given in this article.

Assumption 2. The functional $a(\cdot,\cdot)(\cdot) : Q \times V \times V \to \mathbb{R}$ defined in (2.1) is assumed to be four times directional differentiable.

Assumption 3. The functional $J(\cdot,\cdot) : Q \times V \to \mathbb{R}$ defined in (2.2) is assumed to be four times directional differentiable.

Assumption 4. The functional $I(\cdot,\cdot) : Q \times V \to \mathbb{R}$ mentioned in the introduction (see also (4.20)) is assumed to be three times directional differentiable.

3. Finite element discretization. In this section we discuss finite element discretization of the optimization problem (2.4).

To keep the following sections simple we restrain ourselves to the case of problems where H^1-conforming finite elements are satisfactory. However, the ideas can be adapted to other problems.

Let $\mathcal{T}_h$ be a triangulation (mesh) of the computational domain Ω consisting of closed cells K which are either triangles or quadrilaterals. The straight parts which make up the boundary ∂K of a cell K are called *faces*. The mesh parameter h is defined as a cellwise constant function by setting $h\big|_K = h_K$, and h_K is the diameter of K. The mesh $\mathcal{T}_h$ is assumed to be shape regular. In order to ease the mesh refinement we allow the cells to have nodes, which lie on midpoints of faces of neighboring cells. But at most one of such *hanging nodes* is permitted per face.

On the mesh $\mathcal{T}_h$ we define a finite element space $V_h \subset V$ consisting of linear or bilinear shape functions; see, e.g., [14] or [10]. The case of hanging nodes requires some additional remarks. There are no degrees of freedom corresponding to these irregular nodes, and therefore the value of the finite element function is determined by pointwise interpolation. This implies continuity and therefore global conformity.

For the discretization of the optimization problem (2.4) we introduce an additional finite dimensional subspace $Q_h \subset Q$ of the control space. Depending on the concrete situation there are different possible ways to choose the space Q_h. It is reasonable to set $Q_h = Q$ if Q is finite dimensional. In the case where the control variable is a distributed function on the computational domain Ω, i.e., $Q = L^2(\Omega)$, one may choose Q_h as an analogue to V_h or consider Q_h as a space of cellwise constant functions on the mesh $\mathcal{T}_h$. A priori error analysis for the last two choices in the context of distributed (or boundary) elliptic optimal control problems can be found, e.g., in [1, 11, 15, 18, 28] for cellwise constant control or in [12, 32, 33] for continuous cellwise linear control. An approach without discretization of the control variable is presented in [22].

We denote a basis of Q_h by

$$\mathcal{B} = \{\psi_i\}, \text{ with } \psi_i \geq 0, \quad \sum_i \psi_i = 1, \quad \max_{x \in \omega} \psi_i(x) = 1. \tag{3.1}$$

Remark 3.1. It might be desirable to use different meshes for the control and the state variable in the case of distributed control. The error estimator presented below can provide information for separate refinement of the control and state meshes. One can split the error estimator into two parts, one containing the functionals on the space V which give information for the refinement of the state mesh and one part consisting of the functionals defined on the control space Q which give information for the refinement of the control mesh. The refinement then follows an equilibration strategy for both estimators; cf. [30].

The discrete admissible set $Q_{\mathrm{ad},h}$ is defined as

$$Q_{\mathrm{ad},h} = Q_h \cap Q_{\mathrm{ad}}\,,$$

and the discretized optimization problem is formulated as follows:

$$\text{Minimize } J(q_h, u_h)\,, \quad u_h \in V_h,\, q_h \in Q_{\mathrm{ad},h}\,, \tag{3.2}$$

subject to

$$a(q_h, u_h)(v_h) = f(v_h) \quad \forall v_h \in V_h. \tag{3.3}$$

We introduce the discretized versions of (2.5) and (2.6) by

$$\mathcal{X}_h = Q_h \times V_h \times V_h, \tag{3.4}$$

$$\mathcal{X}_{\mathrm{ad},h} = Q_{\mathrm{ad},h} \times V_h \times V_h \tag{3.5}$$

and denote a vector from these sets by $\xi_h = (q_h, u_h, z_h)$. The optimality system for the discretized optimization problem is formulated as follows:

$$J_1'(u_h)(\delta u_h) - a_u'(q_h, u_h)(\delta u_h, z_h) = 0 \quad \forall \delta u_h \in V_h, \tag{3.6a}$$

$$\alpha(q_h, \delta q_h - q_h) - a_q'(q_h, u_h)(\delta q_h - q_h, z_h) \geq 0 \quad \forall \delta q_h \in Q_{\mathrm{ad},h}, \tag{3.6b}$$

$$f(\delta z_h) - a(q_h, u_h)(\delta z_h) = 0 \quad \forall \delta z_h \in V_h. \tag{3.6c}$$

The nonlinear primal dual active set strategy, described in the previous section, can be translated directly into the discrete level to solve (3.6a)–(3.6c).

In order to formulate the analog system to (2.18a)–(2.18f) we introduce discrete active sets $\omega_{-,h}$ and $\omega_{+,h}$ for a solution (q_h, u_h) to (3.2), (3.3) by

$$\omega_{-,h} = \{x \in \omega \,|\, q_h(x) = q_-\}, \tag{3.7}$$

$$\omega_{+,h} = \{x \in \omega \,|\, q_h(x) = q_+\} \tag{3.8}$$

and define a Lagrange multiplier $\mu_h \in Q_h$ via

$$(\mu_h, \delta q_h) = -\mathcal{L}_q'(q_h, u_h, z_h)(\delta q_h) \quad \forall\, \delta q_h \in Q_h. \tag{3.9}$$

Moreover, we introduce $\mu_h^- \in Q_h$ and $\mu_h^+ \in Q_h$ by

$$\mu_h^+ - \mu_h^- = \mu_h, \;\; (\mu_h^-, \psi_i) \geq 0, \;\; (\mu_h^+, \psi_i) \geq 0 \quad \forall \psi_i \in \mathcal{B} \tag{3.10}$$

by which $\mu_h^\pm$ are uniquely determined if in addition the following complementarity conditions hold:

$$(\mu_h^-, q_h - q_-) = (\mu_h^+, q_+ - q_h) = 0. \tag{3.11}$$

Remark 3.2. This definition corresponds to the Lagrange multipliers obtained for the inequality constraints if the discrete optimization problem (3.2), (3.3) is considered a finite dimensional optimization problem for $q_h = \sum_i q_i \psi_i \in Q_h$ with the following restrictions:

$$q_- \leq q_i \leq q_+ \quad \forall i.$$

Note that due to the choice of the basis $\mathcal{B}$ in (3.1) this is equivalent to $q_- \leq q_h(x) \leq q_+$ for all $x \in \omega$. Utilizing this fact, the discrete active sets $\omega_{-,h}$, $\omega_{+,h}$ are completely determined by the values of the coordinate vector of q_h. In particular they consist only of whole cells, edges, and nodes.

To obtain the complementarity conditions with respect to the $Q = L^2(\omega)$-inner product (3.11) one requires

$$(\mu_h^+, \psi_i) = 0 \text{ if } \; q_i < q_+ \text{ and } \quad (\mu_h^-, \psi_i) = 0 \text{ if } q_i > q_- \,.$$

We now define the discretized versions of (2.14), (2.16), and (2.15) by

$$\mathcal{Y}_h = \mathcal{X}_h \times Q_h \times Q_h, \tag{3.12}$$

$$\mathcal{Y}_{\mathrm{ad},h} = \mathcal{X}_{\mathrm{ad},h} \times Q_{-,h} \times Q_{+,h}, \tag{3.13}$$

$$\tilde{\mathcal{Y}}_{\mathrm{ad},h} = \mathcal{X}_h \times Q_{-,h} \times Q_{+,h}, \tag{3.14}$$

where

$$\begin{aligned} Q_{-,h} &= \{ r \in Q_h \,|\, r(x) = 0 \text{ a.e. on } \omega \setminus \omega_h^- \}, \\ Q_{+,h} &= \{ r \in Q_h \,|\, r(x) = 0 \text{ a.e. on } \omega \setminus \omega_h^+ \}. \end{aligned}$$

A vector from these spaces will be abbreviated by $\chi_h = (q_h, u_h, z_h, \mu_h^-, \mu_h^+)$.

Using the definitions above we have the first-order necessary optimality condition for $(q_h, u_h) \in Q_{\mathrm{ad},h} \times V_h$:

There exist $z_h \in V_h$, $\mu_h^- \in Q_{-,h}$, $\mu_h^+ \in Q_{+,h}$ such that for $\chi_h = (q_h, u_h, z_h, \mu_h^-, \mu_h^+) \in \mathcal{Y}_{\mathrm{ad}}$ the following conditions hold:

$$\tilde{\mathcal{L}}'_u(\chi_h)(\delta u) = 0 \quad \forall \delta u \in V_h, \tag{3.15a}$$

$$\tilde{\mathcal{L}}'_q(\chi_h)(\delta q) = 0 \quad \forall \delta q \in Q_h, \tag{3.15b}$$

$$\tilde{\mathcal{L}}'_z(\chi_h)(\delta z) = 0 \quad \forall \delta z \in V_h, \tag{3.15c}$$

$$\tilde{\mathcal{L}}'_{\mu^-}(\chi_h)(\delta \mu^-) = 0 \quad \forall \delta \mu^- \in Q_{-,h}, \tag{3.15d}$$

$$\tilde{\mathcal{L}}'_{\mu^+}(\chi_h)(\delta \mu^+) = 0 \quad \forall \delta \mu^+ \in Q_{+,h}, \tag{3.15e}$$

$$\mu_h^+ - \mu_h^- = \mu_h, \quad (\mu_h^-, \psi_i) \geq 0, \quad (\mu_h^+, \psi_i) \geq 0 \quad \forall \psi_i \in \mathcal{B}. \tag{3.15f}$$

Here again (3.15d), (3.15e) are equivalent to the complementarity condition

$$(\mu_h^-, q_- - q_h) = (\mu_h^+, q_h - q_+) \,=\, 0. \tag{3.16}$$

Finally we state the following assumption concerning our discretization which is the analogue to Assumption 1.

Assumption 5. The optimization problem (3.2), (3.3) possesses a solution (q_h, u_h). In addition there exists $z_h \in V_h$ such the first-order necessary conditions (3.6) are fulfilled by the triple (q_h, u_h, z_h).

4. A posteriori error estimation. The aim of this section is to derive a posteriori error estimates for the error with respect to the cost functional and to an arbitrary quantity of interest. These error estimates extend the results from [4, 6, 7, 8] to the case of optimization problems with control constraints. The provided estimators will be used within the following adaptive algorithm for error control and mesh refinement: We start on a coarse mesh, solve the discretized optimization problem, and evaluate the error estimator. Thereafter we refine the current mesh using local information obtained from the error estimator, allowing for efficient reduction of the discretization error with respect to the quantity of interest. This procedure is iterated until the value of the error estimator is below a given tolerance; see, e.g., [7] for a detailed description of this algorithm.

The section is structured as follows: First we will derive two a posteriori error estimators for the error with respect to the cost functional. The first one is based on the first-order necessary condition (2.8), which involves a variational inequality, and the second estimator uses the information obtained from the Lagrange multipliers for the inequality constraints. Both estimators can be evaluated in terms of the solution to the discretized optimization problem (3.2), (3.3). Then we will proceed with the error estimator with respect to an arbitrary quantity of interest, which requires the solution to an auxiliary linear-quadratic optimization problem. Even though the idea behind the estimators remains unchanged, the latter estimators require a more technical discussion.

Throughout this section we shall denote a solution to the optimization problem (2.4) by (q, u) and the corresponding solution to the optimality system (2.7) by $\xi = (q, u, z) \in \mathcal{X}_{\mathrm{ad}}$ and its discrete counterpart (3.6) by $\xi_h = (q_h, u_h, z_h) \in \mathcal{X}_{\mathrm{ad},h}$. The corresponding solution to (2.18) and its discrete counterpart (3.15) will be abbreviated as $\chi = (q, u, z, \mu^-, \mu^+) \in \mathcal{Y}_{\mathrm{ad}}$ and $\chi_h = (q_h, u_h, z_h, \mu_h^-, \mu_h^+) \in \mathcal{Y}_{\mathrm{ad},h}$.

4.1. Error in the cost functional. For the derivation of the error estimator with respect to the cost functional, we introduce the residual functionals $\rho_u(\xi_h)(\cdot)$, $\rho_z(\xi_h)(\cdot) \in V'$, and $\rho_q(\xi_h)(\cdot) \in Q'$ by

$$\rho_u(\xi_h)(\cdot) = f(\cdot) - a(q_h, u_h)(\cdot), \tag{4.1}$$

$$\rho_z(\xi_h)(\cdot) = J_1'(u_h)(\cdot) - a_u'(q_h, u_h)(\cdot, z_h), \tag{4.2}$$

$$\rho_q(\xi_h)(\cdot) = \alpha(q_h, \cdot) - a_q'(u_h, q_h)(\cdot, z_h). \tag{4.3}$$

The following theorem is an extension of the result from [6].

THEOREM 4.1. *Let $\xi \in \mathcal{X}_{ad}$ be a solution to the first-order necessary system* (2.7) *and $\xi_h \in \mathcal{X}_{ad,h}$ be its Galerkin approximation* (3.6). *Then the following estimate holds:*

$$J(q,u) - J(q_h,u_h) \le \frac{1}{2}\rho_u(\xi_h)(z - \tilde{z}_h) + \frac{1}{2}\rho_z(\xi_h)(u - \tilde{u}_h) + \frac{1}{2}\rho_q(\xi_h)(q - q_h) + R_1, \tag{4.4}$$

where $\tilde{u}_h, \tilde{z}_h \in V_h$ are arbitrarily chosen and R_1 is a remainder term given by

$$R_1 = \frac{1}{2}\int_0^1 \mathcal{L}'''(\xi_h + s(\xi - \xi_h))(\xi - \xi_h, \xi - \xi_h, \xi - \xi_h)s(s-1)\, ds. \tag{4.5}$$

Proof. From optimality system (2.7a)–(2.7c) we obtain that

$$J(q, u) = \mathcal{L}(\xi).$$

A similar equality holds on the discrete level. Therefore we have

$$J(q,u) - J(q_h,u_h) = \mathcal{L}(\xi) - \mathcal{L}(\xi_h) = \int_0^1 \mathcal{L}'(\xi_h + s(\xi - \xi_h))(\xi - \xi_h)\, ds.$$

We approximate this integral by the trapezoidal rule and obtain

$$J(q,u) - J(q_h,u_h) = \frac{1}{2}\mathcal{L}'(\xi)(\xi - \xi_h) + \frac{1}{2}\mathcal{L}'(\xi_h)(\xi - \xi_h) + R_1, \tag{4.6}$$

with the reminder term R_1 as in (4.5). For the first term we have

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

Using optimality system (2.7a)–(2.7c) and the fact that $q_h \in Q_{\mathrm{ad},h} \subset Q_{\mathrm{ad}}$, we deduce that

$$\mathcal{L}'(\xi)(\xi - \xi_h) = -\mathcal{L}'_q(\xi)(q_h - q) \leq 0.$$

Rewriting the second term in (4.6) we obtain

$$\mathcal{L}'(\xi_h)(\xi - \xi_h) = \rho_u(\xi_h)(z - z_h) + \rho_z(\xi_h)(u - u_h) + \rho_q(\xi_h)(q - q_h).$$

Due to the Galerkin orthogonality for the state and adjoint equations, we have for arbitrary $\tilde{u}_h, \tilde{z}_h \in V_h$

$$\rho_u(\xi_h)(z - z_h) = \rho_u(\xi_h)(z - \tilde{z}_h) \quad \text{and} \quad \rho_z(\xi_h)(u - u_h) = \rho_z(\xi_h)(u - \tilde{u}_h).$$

This completes the proof. $\square$

Remark 4.1. We note that, in contrast to the terms involving the residuals of state and the adjoint equations, the error $q - q_h$ in the term $\rho_q(\xi_h)(q - q_h)$ in (4.4) cannot be replaced by $q - \tilde{q}_h$ with an arbitrary $\tilde{q}_h \in Q_{\mathrm{ad},h}$. This fact is caused by the control constraints. However, we may replace $\rho_q(\xi_h)(q - q_h)$ by $\rho_q(\xi_h)(q - q_h + \tilde{q}_h)$ with arbitrary $\tilde{q}_h$ fulfilling $\operatorname{supp}(\tilde{q}_h) \subset \omega \setminus (\omega_{-,h} \cup \omega_{+,h})$ due to the structure of $\rho_q(\xi_h)(\cdot)$.

In order to use the estimate from the theorem above for computable error estimation we proceed as follows: First we choose $\tilde{u}_h = i_h u$, $\tilde{z}_h = i_h z$, with an interpolation operator $i_h \colon V \to V_h$; then we have to approximate the corresponding interpolation errors $u - i_h u$ and $z - i_h z$. There are several heuristic techniques to do this; see, for instance, [6, 7]. Assume we have an operator $\pi \colon V_h \to \tilde{V}_h$, with $\tilde{V}_h \neq V_h$, such that $u - \pi u_h$ has a better local asymptotical behavior as $u - i_h u$. Then we approximate

$$\rho_u(\xi_h)(z - i_h z) \approx \rho_u(\xi_h)(\pi z_h - z_h) \quad \text{and} \quad \rho_z(\xi_h)(u - i_h u) \approx \rho_z(\xi_h)(\pi u_h - u_h).$$

Such an operator can be constructed, for example, by the interpolation of the computed bilinear finite element solution in the space of biquadratic finite elements on patches of cells. For this operator the improved approximation property relies on local smoothness of u and superconvergence properties of the approximation u_h. The use of such "local higher-order approximation" is observed to work very successfully in the context of a posteriori error estimation; see, e.g., [6, 7].

The approximation of the term $\rho_q(\xi_h)(q - q_h)$ requires more care. In contrast to the state u and the adjoint state z, the control variable q can generally not be approximated by "local higher-order approximation" for the following reasons:

- In the case of finite dimensional control space Q, there is no "patch-like" structure allowing for "local higher-order approximation."

- If q is a distributed control, it typically does not possess sufficient smoothness (due to the inequality constraints) for the improved approximation property.

We therefore suggest another approximation of $\rho_q(\xi_h)(q-q_h)$ based on the projection formula (2.9). To this end we introduce $\tilde{q} \in Q_{\text{ad}}$ by

$$\tilde{q} = \mathcal{P}_{Q_{\text{ad}}}\left(\frac{1}{\alpha}a_q'(q_h, \pi u_h)(\cdot, \pi z_h)\right). \tag{4.7}$$

In some cases one can show better approximation behavior of $q - \tilde{q}$ in comparison with $q - q_h$; see [31] and [22] for similar considerations in the context of a priori error analysis.

This construction results in the following computable a posteriori error estimator:

$$\eta_1 = \frac{1}{2}\big(\rho_u(\xi_h)(\pi z_h - z_h) + \rho_z(\xi_h)(\pi u_h - u_h) + \rho_q(\xi_h)(\tilde{q} - q_h)\big).$$

Remark 4.2. In order to use this error estimator as an indicator for mesh refinement, we have to localize it to cellwise or nodewise contributions. A direct localization of the terms like $\rho_u(\xi_h)(\pi z_h - z_h)$ leads, in general, to the local contributions of wrong order (overestimation) due to oscillatory behavior of the residual terms. To overcome this, one may integrate the residual terms by part (see, e.g., [6]) or use a filtering operator; see [36] for details.

We should note that (4.4) does not provide an estimate for the absolute value of $J(q,u) - J(q_h, u_h)$, which is due to the inequality sign in (4.4). In the next section we will overcome this difficulty utilizing the alternative optimality system (2.18a)–(2.18f).

4.2. Error in the cost functional reviewed. In order to derive an error estimator for the absolute value of $J(q,u) - J(q_h,u_h)$ we introduce the additional residual functionals $\tilde{\rho}_q(\chi_h)(\cdot)$, $\tilde{\rho}_{\mu^-}(\chi_h)(\cdot)$, $\tilde{\rho}_{\mu^+}(\chi_h)(\cdot) \in Q'$ by

$$\tilde{\rho}_q(\chi_h)(\cdot) = \alpha(q_h, \cdot) - a_q'(q_h, u_h)(\cdot, z_h) + (\mu_h^+ - \mu_h^-, \cdot), \tag{4.8}$$

$$\tilde{\rho}_{\mu^-}(\chi_h)(\cdot) = (\cdot, q_- - q_h), \tag{4.9}$$

$$\tilde{\rho}_{\mu^+}(\chi_h)(\cdot) = (\cdot, q_h - q_+). \tag{4.10}$$

In what follows, the last two residual functional will also be evaluated in the point χ where they read as follows:

$$\tilde{\rho}_{\mu^-}(\chi)(\cdot) = (\cdot, q_- - q), \quad \tilde{\rho}_{\mu^+}(\chi)(\cdot) = (\cdot, q - q_+).$$

Analogous to Theorem 4.1 we obtain the following theorem.

THEOREM 4.2. *Let $\chi \in \mathcal{Y}_{ad}$ be a solution to the first-order necessary condition* (2.18a)–(2.18f) *and $\chi_h \in \mathcal{Y}_{ad,h}$ be its Galerkin approximation* (3.15a)–(3.16). *Then the following estimate holds:*

$$\begin{aligned} J(q,u) - J(q_h,u_h) = {} & \frac{1}{2}\rho_u(\chi_h)(z - \tilde{z}_h) + \frac{1}{2}\rho_z(\chi_h)(u - \tilde{u}_h) + \frac{1}{2}\tilde{\rho}_q(\chi_h)(q - \tilde{q}_h) \\ & + \frac{1}{2}\tilde{\rho}_{\mu^-}(\chi_h)(\mu^- - \tilde{\mu}_h^-) + \frac{1}{2}\tilde{\rho}_{\mu^+}(\chi_h)(\mu^+ - \tilde{\mu}_h^+) \\ & + \frac{1}{2}\tilde{\rho}_{\mu^-}(\chi)(\tilde{\mu}^- - \mu_h^-) + \frac{1}{2}\tilde{\rho}_{\mu^+}(\chi)(\tilde{\mu}^+ - \mu_h^+) + R_2, \end{aligned} \tag{4.11}$$

where $\tilde{u}_h, \tilde{z}_h \in V_h$, $\tilde{q}_h \in Q_h$, $\tilde{\mu}_h^- \in Q_{-,h}$, $\tilde{\mu}_h^+ \in Q_{+,h}$, $\tilde{\mu}^- \in Q_-$, $\tilde{\mu}^+ \in Q_+$ are arbitrarily chosen and R_2 is a remainder term given by

$$R_2 = \frac{1}{2}\int_0^1 \tilde{\mathcal{L}}'''(\chi_h + s(\chi - \chi_h))(\chi - \chi_h, \chi - \chi_h, \chi - \chi_h)s(s-1)\,ds. \tag{4.12}$$

Proof. From (2.17) and optimality system (2.8a)–(2.8c) we obtain

$$J(q,u) = \mathcal{L}(\xi) = \tilde{\mathcal{L}}(\chi).$$

The analog result holds on the discrete level. We therefore have

$$J(q,u) - J(q_h,u_h) = \tilde{\mathcal{L}}(\chi) - \tilde{\mathcal{L}}(\chi_h) = \int_0^1 \tilde{\mathcal{L}}'(\chi_h + s(\chi-\chi_h))(\chi-\chi_h)\,ds.$$

As in the proof of Theorem 4.1 we approximate this integral by the trapezoidal rule and obtain

$$J(q,u) - J(q_h,u_h) = \frac{1}{2}\tilde{\mathcal{L}}'(\chi)(\chi-\chi_h) + \frac{1}{2}\tilde{\mathcal{L}}'(\chi_h)(\chi-\chi_h) + R_2, \tag{4.13}$$

with the remainder term R_2 as in (4.12). For the first term we have

$$\begin{aligned}\tilde{\mathcal{L}}'(\chi)(\chi-\chi_h) =& \tilde{\mathcal{L}}'_u(\chi)(u-u_h) + \tilde{\mathcal{L}}'_z(\chi)(z-z_h) + \tilde{\mathcal{L}}'_q(\chi)(q-q_h)\\ &+ \tilde{\mathcal{L}}'_{\mu^-}(\chi)(\mu^- - \mu_h^-) + \tilde{\mathcal{L}}'_{\mu^+}(\chi)(\mu^+ - \mu_h^+).\end{aligned}$$

Using optimality system (2.18a)–(2.18f) we deduce that

$$\tilde{\mathcal{L}}'(\chi)(\chi-\chi_h) = \tilde{\mathcal{L}}'_{\mu^-}(\chi)(\mu^- - \mu_h^-) + \tilde{\mathcal{L}}'_{\mu^+}(\chi)(\mu^+ - \mu_h^+).$$

From (2.18d) and (2.18e) together with linearity of $\tilde{\mathcal{L}}'_{\mu^-}(\chi)(\cdot)$ and $\tilde{\mathcal{L}}'_{\mu^+}(\chi)(\cdot)$ we obtain that for arbitrary $\tilde{\mu}^- \in Q_-$ and $\tilde{\mu}^+ \in Q_+$

$$\tilde{\mathcal{L}}'_{\mu^-}(\chi)(\mu^- - \mu_h^-) = \tilde{\mathcal{L}}'_{\mu^-}(\chi)(\tilde{\mu}^- - \mu_h^-), \quad \tilde{\mathcal{L}}'_{\mu^+}(\chi)(\mu^+ - \mu_h^+) = \tilde{\mathcal{L}}'_{\mu^+}(\chi)(\tilde{\mu}^+ - \mu_h^+)$$

holds, and thus we obtain

$$\tilde{\mathcal{L}}'(\chi)(\chi-\chi_h) = \tilde{\rho}_{\mu^-}(\chi)(\tilde{\mu}^- - \mu_h^-) + \tilde{\rho}_{\mu^+}(\chi)(\tilde{\mu}^+ - \mu_h^+).$$

Rewriting the second term in (4.13) we obtain

$$\begin{aligned}\tilde{\mathcal{L}}'(\chi_h)(\chi-\chi_h) =& \rho_u(\chi_h)(u-u_h) + \rho_z(\chi_h)(z-z_h) + \tilde{\rho}_q(\chi_h)(q-q_h)\\ &+ \tilde{\rho}_{\mu^-}(\chi_h)(\mu^- - \mu_h^-) + \tilde{\rho}_{\mu^+}(\chi_h)(\mu^+ - \mu_h^+),\end{aligned}$$

where we can use linearity of the residual functionals in the second argument and (3.15a)–(3.15c) to obtain the following equalities:

$$\rho_u(\chi_h)(u-u_h) = \rho_u(\chi_h)(u-\tilde{u}_h), \tag{4.14}$$

$$\rho_z(\chi_h)(z-z_h) = \rho_z(\chi_h)(z-\tilde{z}_h), \tag{4.15}$$

$$\tilde{\rho}_q(\chi_h)(q-q_h) = \tilde{\rho}_q(\chi_h)(q-\tilde{q}_h) \tag{4.16}$$

for arbitrary $\tilde{u}_h, \tilde{z}_h \in V_h$, $\tilde{q}_h \in Q_h$. Additionally we gain from (3.15d) and (3.15e) that for arbitrary $\tilde{\mu}_h^- \in Q_{-,h}$ and $\tilde{\mu}_h^+ \in Q_{+,h}$

$$\tilde{\rho}_{\mu^-}(\chi_h)(\mu^- - \mu_h^-) = \tilde{\rho}_{\mu^-}(\chi_h)(\mu^- - \tilde{\mu}_h^-), \tag{4.17}$$

$$\tilde{\rho}_{\mu^+}(\chi_h)(\mu^+ - \mu_h^+) = \tilde{\rho}_{\mu^+}(\chi_h)(\mu^+ - \tilde{\mu}_h^+) \tag{4.18}$$

holds. This completes the proof. ∎

To gain a computable error estimator we proceed as in the previous section. In order to deal with the new residual functionals we utilize (2.12) and construct an approximation for μ by

$$\tilde{\mu} = -\alpha\tilde{q} + a'_q(\tilde{q}, \pi u_h)(\cdot, \pi z_h), \tag{4.19}$$

where $\tilde{q}$ is given by (4.7). This leads to a computable a posteriori error estimator:

$$\begin{aligned}\eta_2 = \frac{1}{2}\big(&\rho_u(\chi_h)(\pi z_h - z_h) + \rho_z(\chi_h)(\pi u_h - u_h) + \tilde{\rho}_q(\chi_h)(\tilde{q} - q_h),\\ &\tilde{\rho}_{\mu^-}(\chi_h)(\tilde{\mu}^- - \mu_h^-) + \tilde{\rho}_{\mu^+}(\chi_h)(\tilde{\mu}^+ - \mu_h^+),\\ &\tilde{\rho}_{\mu^-}(\tilde{\chi})(\tilde{\mu}^- - \mu_h^-) + \tilde{\rho}_{\mu^+}(\tilde{\chi})(\tilde{\mu}^+ - \mu_h^+)\big).\end{aligned}$$

Remark 4.3. We note that the a posteriori error estimates derived in Theorems 4.1 and 4.2 coincide if the control constraints are inactive, e.g., if $Q_{\mathrm{ad}} = Q$. Moreover, if the active sets are approximated from outside, i.e., $\omega_- \subset \omega_{-,h}$ and $\omega_+ \subset \omega_{+,h}$, these error estimators coincide as well.

4.3. Error in the quantity of interest. The aim of this section is the derivation of an error estimator for the error

$$I(q, u) - I(q_h, u_h) \tag{4.20}$$

with a given functional $I : Q \times V \to \mathbb{R}$ describing the quantity of interest which we require to be three times directional differentiable. To this end we consider an additional Lagrangian $\mathcal{M} : \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}$ defined by

$$\mathcal{M}(\chi)(\psi) = I(q, u) + \tilde{\mathcal{L}}'(\chi)(\psi), \tag{4.21}$$

where we abbreviate $\chi = (q, u, z, \mu^-, \mu^+)$ and $\psi = (p, v, y, \nu^-, \nu^+)$. Here (p, v, y, ν^-, ν^+) will be variables dual to (q, u, z, μ^-, μ^+). Note that for the solution χ to the optimality system (2.18a)–(2.18f) of the optimization problem (2.4) the identity

$$\mathcal{M}(\chi)(\psi) = I(q, u) \tag{4.22}$$

holds for all $\psi \in \tilde{\mathcal{Y}}_{\mathrm{ad}}$. To proceed as in the proof of Theorem 4.2 it remains to find $\psi \in \tilde{\mathcal{Y}}_{\mathrm{ad}}$ such that (χ, ψ) is a stationary point of $\mathcal{M}$ on $\tilde{\mathcal{Y}}_{\mathrm{ad}} \times \tilde{\mathcal{Y}}_{\mathrm{ad}}$.

Therefore we consider the auxiliary (linear-quadratic) optimization problem

$$\text{Minimize } K(\chi, p, v), \quad p \in P_{\mathrm{ad}},\ v \in V, \tag{4.23}$$

$$\text{subject to } \tilde{\mathcal{L}}''_{uz}(\chi)(v, \varphi) + \tilde{\mathcal{L}}''_{qz}(\chi)(p, \varphi) = 0 \quad \forall\, \varphi \in V \tag{4.24}$$

for given $\chi \in \mathcal{Y}$. The admissible set P_{ad} is given as

$$P_{\mathrm{ad}} = \{p \in Q \,|\, p_-(x) \le p(x) \le p_+(x) \text{ a.e. on } \omega\}, \tag{4.25}$$

with the bounds

$$p_-(x) = \begin{cases} 0, & \mu(x) \neq 0 \text{ or } q(x) = q_-(x), \\ -\infty & \text{else,} \end{cases}$$

$$p_+(x) = \begin{cases} 0, & \mu(x) \neq 0 \text{ or } q(x) = q_+(x), \\ +\infty & \text{else,} \end{cases}$$

and the cost functional $K : \mathcal{Y} \times Q \times V \to \mathbb{R}$ is defined via

$$K(\chi, p, v) = I'_u(q,u)(v) + I'_q(q,u)(p) + \tilde{\mathcal{L}}''_{uq}(\chi)(v,p) + \frac{1}{2}\tilde{\mathcal{L}}''_{uu}(\chi)(v,v) + \frac{1}{2}\tilde{\mathcal{L}}''_{qq}(\chi)(p,p). \tag{4.26}$$

We introduce the following abbreviation for later use:

$$\bar{\mathcal{Y}}_{\text{ad}} \; = \; P_{\text{ad}} \times V \times V \times Q_- \times Q_+. \tag{4.27}$$

Remark 4.4. Consideration of the auxiliary optimization problem (4.23), (4.24) is motivated by the unconstrained case $Q_{\text{ad}} = Q$. There the stationary point of $\mathcal{M}$ is given as the solution to (4.23), (4.24) with $P_{\text{ad}} = Q$. A similar linear-quadratic optimization problem is considered in [19] in the context of sensitivity analysis.

Remark 4.5. If we assume that the second-order sufficient condition from Lemma 2.1 holds, the linear-quadratic optimization problem (4.23) possesses a solution. This is the case, as the quadratic part $\tilde{\mathcal{L}}''_{uq}(\chi)(v,p) + \frac{1}{2}\tilde{\mathcal{L}}''_{uu}(\chi)(v,v) + \frac{1}{2}\tilde{\mathcal{L}}''_{qq}(\chi)(p,p)$ of $K(p,v)$ is positive definite (see (2.20)) for all solutions to the linear equation (2.21), which is exactly the same as (4.24).

We introduce an auxiliary Lagrangian $\mathcal{N} : \mathcal{Y} \times \mathcal{X} \to \mathbb{R}$ for (4.23), (4.24) by

$$\mathcal{N}(\chi, p, v, y) = K(\chi, p, v) + \tilde{\mathcal{L}}''_{uz}(\chi)(v,y) + \tilde{\mathcal{L}}''_{qz}(\chi)(p,y). \tag{4.28}$$

For a solution (p, v) to (4.23), (4.24) the following first-order necessary condition holds:

There exists $y \in V$ such that

$$\mathcal{N}'_y(\chi, p, v, y)(\delta y) \; = \; 0 \quad \forall \delta y \in V, \tag{4.29a}$$

$$\mathcal{N}'_v(\chi, p, v, y)(\delta v) \; = \; 0 \quad \forall \delta v \in V, \tag{4.29b}$$

$$\mathcal{N}'_p(\chi, p, v, y)(\delta p - p) \; \geq \; 0 \quad \forall \delta p \in P_{\text{ad}} \tag{4.29c}$$

or, if written more explicitly,

$$\tilde{\mathcal{L}}''_{uz}(\chi)(v, \delta y) + \tilde{\mathcal{L}}''_{qz}(\chi)(p, \delta y) \; = \; 0 \quad \forall \delta y \in V, \tag{4.30a}$$

$$I'_u(q,u)(\delta v) + \tilde{\mathcal{L}}''_{uq}(\chi)(\delta v, p) + \tilde{\mathcal{L}}''_{uu}(\chi)(\delta v, v) + \tilde{\mathcal{L}}''_{uz}(\chi)(\delta v, y) \; = \; 0 \quad \forall \delta v \in V, \tag{4.30b}$$

$$I'_q(q,u)(\delta p) + \tilde{\mathcal{L}}''_{uq}(\chi)(v, \delta p) + \tilde{\mathcal{L}}''_{qq}(\chi)(\delta p, p) + \tilde{\mathcal{L}}''_{qz}(\chi)(\delta p, y) \; \geq \; 0 \quad \forall \delta p \in P_{\text{ad}} - p. \tag{4.30c}$$

Again we can introduce the full Lagrangian $\tilde{\mathcal{N}} : \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}$ by

$$\tilde{\mathcal{N}}(\chi, \psi) = \mathcal{N}(\chi, p, v, y) + (\nu^-, p_- - p) + (\nu^+, p - p_+). \tag{4.31}$$

As in (2.18a)–(2.18f) we can rewrite the necessary optimality condition for $\psi \in \bar{\mathcal{Y}}_{\mathrm{ad}}$ as

$$\begin{aligned}
&\text{(4.32a)} && \tilde{\mathcal{N}}'_v(\chi,\psi)(\delta v) = 0 \quad \forall \delta v \in V,\\
&\text{(4.32b)} && \tilde{\mathcal{N}}'_p(\chi,\psi)(\delta p) = 0 \quad \forall \delta p \in Q,\\
&\text{(4.32c)} && \tilde{\mathcal{N}}'_y(\chi,\psi)(\delta y) = 0 \quad \forall \delta y \in V,\\
&\text{(4.32d)} && \tilde{\mathcal{N}}'_{\nu^-}(\chi,\psi)(\delta \nu^-) = 0 \quad \forall \delta \nu^- \in Q_-,\\
&\text{(4.32e)} && \tilde{\mathcal{N}}'_{\nu^+}(\chi,\psi)(\delta \nu^+) = 0 \quad \forall \delta \nu^+ \in Q_+,\\
&\text{(4.32f)} && \nu^+ - \nu^- = \nu, \quad \nu^-(p_- - p) = \nu^+(p - p_+) = 0 \quad \text{a.e. on}\, \omega,\\
&\text{(4.32g)} && \operatorname{supp} \nu^+ \subseteq \omega \setminus \{x \in \omega \,|\, q = q_- \text{ and } \mu \neq 0\}, \quad \nu^+ \geq 0, \quad \text{a.e. where } \mu = 0,\\
&\text{(4.32h)} && \operatorname{supp} \nu^- \subseteq \omega \setminus \{x \in \omega \,|\, q = q_+ \text{ and } \mu \neq 0\}, \quad \nu^- \geq 0, \quad \text{a.e. where } \mu = 0,
\end{aligned}$$

where ν^- and ν^+ are given by the following relations depending on $\nu = -\mathcal{N}'_p(\chi,p,v,y)(\cdot)$:

$$\nu^+(x) = \begin{cases} \nu, & q(x) = q_+ \text{ and } \mu(x) \neq 0,\\ 0, & q(x) = q_- \text{ and } \mu(x) \neq 0,\\ \max(0,\nu) & \text{else,} \end{cases}$$

$$\nu^-(x) = \begin{cases} \nu, & q(x) = q_- \text{ and } \mu(x) \neq 0,\\ 0, & q(x) = q_+ \text{ and } \mu(x) \neq 0,\\ \max(0,-\nu) & \text{else.} \end{cases}$$

Note that due to the choice of p_- and p_+ the Lagrange multipliers are contained in the desired spaces, e.g., $\nu^- \in Q_-$ and $\nu^+ \in Q_+$.

Remark 4.6. It should be noted that we use the convention $\pm\infty \cdot 0 = 0$ in (4.31), (4.32f) to ease notation. The same convention will be used throughout this section.

Remark 4.7. The condition (4.32g) arises naturally, as ν^+ is the Lagrange multiplier which corresponds to the equality and inequality constraints for p that are induced by the active upper control bound q_+. Similarly (4.32h) arises from the active lower control bound q_-.

We introduce

$$\bar{\mathcal{Y}}_{\mathrm{ad},h} = P_{\mathrm{ad},h} \times V_h \times V_h \times Q_{-,h} \times Q_{+,h} \tag{4.33}$$

to shorten notation. This is discretized using the discretized admissible set

$$P_{\mathrm{ad},h} = \{p \in Q_h \,|\, p_{h,-}(x) \leq p(x) \leq p_{h,-}(x) \text{ a.e. on } \omega\}, \tag{4.34}$$

with the bounds

$$p_{h,-}(x) = \begin{cases} 0, & \mu_h(x) \neq 0 \text{ or } q_h(x) = q_-(x),\\ -\infty & \text{else,} \end{cases}$$

$$p_{h,+}(x) = \begin{cases} 0, & \mu_h(x) \neq 0 \text{ or } q_h(x) = q_+(x),\\ +\infty & \text{else.} \end{cases}$$

Then the following first-order condition holds with the discretized full Lagrangian:

$$\tilde{\mathcal{N}}_h(\chi,\psi) = \mathcal{N}(\chi,p,v,y) + (\nu^-, p_{h,-} - p) + (\nu^+, p - p_{h,+}),$$

where $\tilde{\mathcal{N}}_h : \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}$. There exist $y_h \in V_h$, $\nu_h^+, \nu_h^- \in Q_h$ such that for $\psi_h = (p_h, v_h, y_h, \nu_h^-, \nu_h^+) \in \bar{\mathcal{Y}}_{\mathrm{ad},h}$ the following hold:

$$
\begin{aligned}
&(4.35\text{a}) && \tilde{\mathcal{N}}'_{h,v}(\chi_h, \psi_h)(\delta v) = 0 \quad \forall \delta v \in V_h,\\
&(4.35\text{b}) && \tilde{\mathcal{N}}'_{h,p}(\chi_h, \psi_h)(\delta p) = 0 \quad \forall \delta p \in Q_h,\\
&(4.35\text{c}) && \tilde{\mathcal{N}}'_{h,y}(\chi_h, \psi_h)(\delta y) = 0 \quad \forall \delta y \in V_h,\\
&(4.35\text{d}) && \tilde{\mathcal{N}}'_{h,\nu^-}(\chi_h, \psi_h)(\delta \nu^-) = 0 \quad \forall \delta \nu^- \in Q_{-,h},\\
&(4.35\text{e}) && \tilde{\mathcal{N}}'_{h,\nu^+}(\chi_h, \psi_h)(\delta \nu^+) = 0 \quad \forall \delta \nu^+ \in Q_{+,h},\\
&(4.35\text{f}) && \nu_h^+ - \nu_h^- = \nu_h \quad (\nu_h^-, p_{h,-} - p_h) = (\nu_h^+, p_h - p_{h,+}) = 0,\\
&(4.35\text{g}) && (\nu_h^+, \psi_i) = 0 \;\; \forall i : (\mu_h, \psi_i) \neq 0 \text{ and } q_i = q_-,\\
&(4.35\text{h}) && (\nu_h^+, \psi_i) \geq 0 \;\; \forall i : (\mu_h, \psi_i) = 0,\\
&(4.35\text{i}) && (\nu_h^-, \psi_i) = 0 \;\; \forall i : (\mu_h, \psi_i) \neq 0 \text{ and } q_i = q_+,\\
&(4.35\text{j}) && (\nu_h^-, \psi_i) \geq 0 \;\; \forall i : (\mu_h, \psi_i) = 0.
\end{aligned}
$$

For the error estimator with respect to the quantity of interest we introduce the residual functionals $\tilde{\rho}_v(\chi_h, \psi_h)(\cdot)$, $\tilde{\rho}_y(\chi_h, \psi_h)(\cdot) \in V'$ and $\tilde{\rho}_p(\chi_h, \psi_h)(\cdot)$, $\tilde{\rho}_{\nu^-}(\chi_h, \psi_h)(\cdot)$, $\tilde{\rho}_{\nu^+}(\chi_h, \psi_h)(\cdot) \in Q'$ by

$$
\begin{aligned}
&(4.36) && \tilde{\rho}_v(\chi_h, \psi_h)(\cdot) = \tilde{\mathcal{L}}''_{zu}(\chi_h)(\cdot, v_h) + \tilde{\mathcal{L}}''_{zq}(\chi_h)(\cdot, p_h),\\
&(4.37) && \tilde{\rho}_y(\chi_h, \psi_h)(\cdot) = I'_u(q_h, u_h)(\cdot) + \tilde{\mathcal{L}}''_{uu}(\chi_h)(\cdot, v_h) + \tilde{\mathcal{L}}''_{uz}(\chi_h)(\cdot, y_h)\\
&&& \qquad + \tilde{\mathcal{L}}''_{uq}(\chi_h)(\cdot, p_h),\\
&(4.38) && \tilde{\rho}_p(\chi_h, \psi_h)(\cdot) = I'_q(q_h, u_h)(\cdot) + \tilde{\mathcal{L}}''_{qu}(\chi_h)(\cdot, v_h) + \tilde{\mathcal{L}}''_{qz}(\chi_h)(\cdot, y_h)\\
&&& \qquad + \tilde{\mathcal{L}}''_{qq}(\chi_h)(\cdot, p_h) + (\cdot, \nu_h),\\
&(4.39) && \tilde{\rho}_{\nu^-}(\chi_h, \psi_h)(\cdot) = -(\cdot, p_h),\\
&(4.40) && \tilde{\rho}_{\nu^+}(\chi_h, \psi_h)(\cdot) = (\cdot, p_h),
\end{aligned}
$$

in addition to the already defined residual functionals (4.1)–(4.10). Again the last two residual functionals also have to be evaluated in the point (χ, ψ) where they read as follows:

$$\tilde{\rho}_{\nu^-}(\chi, \psi)(\cdot) = -(\cdot, p), \quad \tilde{\rho}_{\nu^+}(\chi, \psi)(\cdot) = (\cdot, p).$$

THEOREM 4.3. *Let $\chi \in \mathcal{Y}_{ad}$ be a solution to the necessary optimality condition* (2.18) *and $\chi_h \in \mathcal{Y}_{ad,h}$ be its Galerkin approximation* (3.15). *In addition let $\psi \in \bar{\mathcal{Y}}_{ad}$ be a solution to the necessary optimality condition* (4.32) *of the auxiliary optimization problem* (4.23), (4.24) *and $\psi_h \in \bar{\mathcal{Y}}_{ad,h}$ be its discrete approximation* (4.35). *Then the* *following estimate holds:*

(4.41)
$$\begin{aligned}
I(q,u) - I(q_h,u_h) =& \frac{1}{2}\rho_u(\chi_h)(y-\tilde{y}_h) + \frac{1}{2}\rho_z(\chi_h)(v-\tilde{v}_h) + \frac{1}{2}\tilde{\rho}_q(\chi_h)(p-\tilde{p}_h)\\
&+ \frac{1}{2}\tilde{\rho}_{\mu^-}(\chi_h)(\nu^- - \tilde{\nu}_h^-) + \frac{1}{2}\tilde{\rho}_{\mu^+}(\chi_h)(\nu^+ - \tilde{\nu}_h^+)\\
&+ \frac{1}{2}\tilde{\rho}_v(\chi_h,\psi_h)(z-\tilde{z}_h) + \frac{1}{2}\tilde{\rho}_y(\chi_h,\psi_h)(u-\tilde{u}_h) + \frac{1}{2}\tilde{\rho}_p(\chi_h,\psi_h)(q-\tilde{q}_h)\\
&+ \frac{1}{2}\tilde{\rho}_{\nu^-}(\chi_h,\psi_h)(\mu^- - \tilde{\mu}_h^-) + \frac{1}{2}\tilde{\rho}_{\nu^+}(\chi_h,\psi_h)(\mu^+ - \tilde{\mu}_h^+)\\
&+ \frac{1}{2}\tilde{\rho}_{\mu^-}(\chi)(\tilde{\nu}^- - \nu_h^-) + \frac{1}{2}\tilde{\rho}_{\mu^+}(\chi)(\tilde{\nu}^+ - \nu_h^+)\\
&+ \frac{1}{2}\tilde{\rho}_{\nu^-}(\chi,\psi)(\tilde{\mu}^- - \mu_h^-) + \frac{1}{2}\tilde{\rho}_{\nu^+}(\chi,\psi)(\tilde{\mu}^+ - \mu_h^+) + R_3,
\end{aligned}$$

where $\tilde{u}_h,\ \tilde{v}_h,\ \tilde{z}_h,\ \tilde{y}_h \in V_h,\ \tilde{q}_h,\ \tilde{p}_h \in Q_h,\ \tilde{\mu}_h^-,\ \tilde{\nu}_h^- \in Q_{-,h},\ \tilde{\mu}_h^+,\ \tilde{\nu}_h^+ \in Q_{+,h}$ *as well as* $\tilde{\mu}^-,\ \tilde{\nu}^- \in Q_-,\ \tilde{\mu}^+,\ \tilde{\nu}^+ \in Q_+$ *are arbitrarily chosen and* R_3 *is a remainder term given by*

$$R_3 = \frac{1}{2}\int_0^1 \mathcal{M}'''((\chi_h,\psi_h)+se)(e,e,e)s(s-1)\,ds, \tag{4.42}$$

with $e = (\chi-\chi_h, \psi-\psi_h)$.

Proof. From (4.22) and the analog discrete result we obtain

$$I(q,u) - I(q_h,u_h) = \mathcal{M}(\chi,\psi) - \mathcal{M}(\chi_h,\psi_h) = \int_0^1 \mathcal{M}'((\chi_h,\psi_h)+se)(e)\,ds.$$

Approximation by the trapezoidal rule gives

$$I(q,u) - I(q_h,u_h) = \frac{1}{2}\mathcal{M}'(\chi,\psi)(e) + \frac{1}{2}\mathcal{M}'(\chi_h,\psi_h)(e) + R_3, \tag{4.43}$$

with the remainder term R_3 as in (4.42). For the first term we have

$$\begin{aligned}
\mathcal{M}'(\chi,\psi)(e) =& \mathcal{M}'_u(\chi,\psi)(u-u_h) + \mathcal{M}'_v(\chi,\psi)(v-v_h)\\
&+ \mathcal{M}'_z(\chi,\psi)(z-z_h) + \mathcal{M}'_y(\chi,\psi)(y-y_h)\\
&+ \mathcal{M}'_q(\chi,\psi)(q-q_h) + \mathcal{M}'_p(\chi,\psi)(p-p_h)\\
&+ \mathcal{M}'_{\mu^-}(\chi,\psi)(\mu^- - \mu_h^-) + \mathcal{M}'_{\nu^-}(\chi,\psi)(\nu^- - \nu_h^-)\\
&+ \mathcal{M}'_{\mu^+}(\chi,\psi)(\mu^+ - \mu_h^+) + \mathcal{M}'_{\nu^+}(\chi,\psi)(\nu^+ - \nu_h^+).
\end{aligned}$$

Using the identities

$$\begin{aligned}
\mathcal{M}'_u(\chi,\psi)(\cdot) &= \tilde{\mathcal{N}}'_v(\chi,p,v,y)(\cdot), \quad \mathcal{M}'_v(\chi,\psi)(\cdot) = \tilde{\mathcal{L}}'_u(\chi)(\cdot),\\
\mathcal{M}'_z(\chi,\psi)(\cdot) &= \tilde{\mathcal{N}}'_y(\chi,p,v,y)(\cdot), \quad \mathcal{M}'_y(\chi,\psi)(\cdot) = \tilde{\mathcal{L}}'_z(\chi)(\cdot),\\
\mathcal{M}'_q(\chi,\psi)(\cdot) &= \tilde{\mathcal{N}}'_p(\chi,p,v,y)(\cdot), \quad \mathcal{M}'_p(\chi,\psi)(\cdot) = \tilde{\mathcal{L}}'_q(\chi)(\cdot),
\end{aligned}$$

we see that the first six terms on the right-hand side vanish due to (2.18a)–(2.18c) and (4.32a)–(4.32c). Furthermore we see from (2.18d), (2.18e) and (4.32d), (4.32e) that with arbitrary $\tilde\mu^-$, $\tilde\nu^- \in Q_-$ and $\tilde\mu^+$, $\tilde\nu^+ \in Q_+$ the following identities hold:

$$\mathcal{M}'_{\mu^-}(\chi,\psi)(\mu^- - \mu_h^-) = \mathcal{M}'_{\mu^-}(\chi,\psi)(\tilde\mu^- - \mu_h^-) = \tilde\rho_{\nu^-}(\chi,\psi)(\tilde\mu^- - \mu_h^-), \tag{4.44}$$

$$\mathcal{M}'_{\nu^-}(\chi,\psi)(\nu^- - \nu_h^-) = \mathcal{M}'_{\nu^-}(\chi,\psi)(\tilde\nu^- - \nu_h^-) = \tilde\rho_{\mu^-}(\chi)(\tilde\nu^- - \nu_h^-), \tag{4.45}$$

$$\mathcal{M}'_{\mu^+}(\chi,\psi)(\mu^+ - \mu_h^+) = \mathcal{M}'_{\mu^+}(\chi,\psi)(\tilde\mu^+ - \mu_h^+) = \tilde\rho_{\nu^+}(\chi,\psi)(\tilde\mu^+ - \mu_h^+), \tag{4.46}$$

$$\mathcal{M}'_{\nu^+}(\chi,\psi)(\nu^+ - \nu_h^+) = \mathcal{M}'_{\nu^+}(\chi,\psi)(\tilde\nu^+ - \nu_h^+) = \tilde\rho_{\mu^+}(\chi)(\tilde\nu^+ - \nu_h^+). \tag{4.47}$$

Thus we obtain

$$\begin{aligned}\mathcal{M}'(\chi,\psi)(e) &= \tilde\rho_{\mu^-}(\chi)(\tilde\nu^- - \nu_h^-) + \tilde\rho_{\mu^+}(\chi)(\tilde\nu^+ - \nu_h^+)\\ &\quad + \tilde\rho_{\nu^-}(\chi,\psi)(\tilde\mu^- - \mu_h^-) + \tilde\rho_{\nu^+}(\chi,\psi)(\tilde\mu^+ - \mu_h^+).\end{aligned}$$

For the second term we obtain from (3.15a)–(3.15e) and (4.32a)–(4.32e) that

$$\mathcal{M}'(\chi_h,\psi_h)(e) = \mathcal{M}'(\chi_h,\psi_h)(\chi - \tilde\chi_h, \psi - \tilde\psi_h)$$

for each $\tilde\chi_h, \tilde\psi_h \in \tilde{\mathcal{Y}}_{\mathrm{ad},h}$, which completes the proof. ∎

Remark 4.8. Note that in the case $I = J$ the solution (p, v, y) to (4.29) is given by $(0, 0, z)$, which can be seen after some calculations. Using this, one obtains that for $I = J$ the estimates in Theorems 4.2 and 4.3 coincide.

We define the projection onto the admissible set by

$$\mathcal{P}_{P_{\mathrm{ad},h}}(p) = \max\bigl(p_{h,-}, \min(p, p_{h,+})\bigr).$$

To obtain a computable error estimator we introduce $\tilde p \in P_{\mathrm{ad}}$ as an approximation to p by

$$\tilde p = \mathcal{P}_{P_{\mathrm{ad},h}}\left(\frac{1}{\alpha}(a'_q()(\cdot,\pi y_h) + a''_{qu}()(\cdot,\pi v_h,\pi z_h) + a''_{qq}()(\cdot,p_h,\pi z_h) - I'_q(\tilde q,\pi u_h)(\cdot))\right), \tag{4.48}$$

where $()$ is an abbreviation for $(\tilde q, \pi u_h)$, and $\tilde\nu$ is introduced as an approximation to ν by

$$\tilde\nu = -\alpha\tilde p + a'_q()(\cdot,\pi y_h) + a''_{qu}()(\cdot,\pi v_h,\pi z_h) + a''_{qq}()(\cdot,p_h,\pi z_h) - I'_q(\tilde q,\pi u_h)(\cdot), \tag{4.49}$$

which is an analogue to the construction of the approximations $\tilde q$ and $\tilde\mu$ in (4.7) and (4.19).

Using these approximations we obtain the following computable error estimator:

$$\begin{aligned}\eta_{\mathrm{QI}} &= \frac12\rho_u(\chi_h)(\pi y - y_h) + \frac12\rho_z(\chi_h)(\pi v - v_h) + \frac12\tilde\rho_q(\chi_h)(\tilde p - p_h)\\
&\quad + \frac12\tilde\rho_{\mu^-}(\chi_h)(\tilde\nu^- - \nu_h^-) + \frac12\tilde\rho_{\mu^+}(\chi_h)(\tilde\nu^+ - \nu_h^+)\\
&\quad + \frac12\tilde\rho_v(\chi_h,\psi_h)(\pi z - z_h) + \frac12\tilde\rho_y(\chi_h,\psi_h)(\pi u - u_h) + \frac12\tilde\rho_p(\chi_h,\psi_h)(\tilde q - q_h)\\
&\quad + \frac12\tilde\rho_{\nu^-}(\chi_h,\psi_h)(\tilde\mu^- - \mu_h^-) + \frac12\tilde\rho_{\nu^+}(\chi_h,\psi_h)(\tilde\mu^+ - \mu_h^+)\\
&\quad + \frac12\tilde\rho_{\mu^-}(\tilde\chi)(\tilde\nu^- - \nu_h^-) + \frac12\tilde\rho_{\mu^+}(\tilde\chi)(\tilde\nu^+ - \nu_h^+)\\
&\quad + \frac12\tilde\rho_{\nu^-}(\tilde\chi,\tilde\psi)(\tilde\mu^- - \mu_h^-) + \frac12\tilde\rho_{\nu^+}(\tilde\chi,\tilde\psi)(\tilde\mu^+ - \mu_h^+),\end{aligned}$$

where $\tilde\chi = (\tilde q, \pi u, \pi z, \tilde\mu^-, \tilde\mu^+)$ and $\tilde\psi = (\tilde p, \pi v, \pi y, \tilde\nu^-, \tilde\nu^+)$.

Remark 4.9. We would like to point out that in case of strict complementarity, e.g., if the set

$$\{x \in \omega \mid q(x) = q_-(x) \text{ or } q(x) = q_+(x)\} \setminus \{x \in \omega \mid \mu(x) \neq 0\}$$

has zero measure, the auxiliary problem (4.23), (4.24) does not involve inequality constraints for the controls. In that case the set P_{ad} is not only convex but in fact a real subspace of Q.

Remark 4.10. The constrained linear-quadratic optimization problem (4.23), (4.24) can be solved using primal-dual active set strategy. In the case of strict complementarity the algorithm will converge in one step due to the fact that P_{ad} is a linear subspace of Q is this case.

Remark 4.11. Due to the definition of P_{ad} (4.25), the solution $p \in Q$ of auxiliary optimization problem (4.23)–(4.24) is usually discontinuous. Therefore, a cellwise constant discretization of the control space Q seems to be more suitable than a discretization with continuous trial functions if the error with respect to a quantity of interested is estimated.

5. Numerical examples. In this section we discuss two numerical examples illustrating the behavior of our method. For both examples we use bilinear (H^1-conforming) finite elements for the discretization of the state variable and cellwise constant discretization of the control space. The optimization problems are solved by primal-dual active set strategy as sketched in section 2, where the equality-constrained problems in the inner loop are solved using Newton's method for the reduced cost functional.

All examples have been computed using the optimization library RoDoBo [5] and the finite element toolkit Gascoigne [3].

5.1. Example 1. We consider the following nonlinear optimization problem:

$$\text{Minimize} \quad \frac{1}{2}\|u - u^d\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|q\|^2_{L^2(\Omega)}, \qquad u \in V,\ q \in Q_{\mathrm{ad}}, \tag{5.1}$$

subject to

$$\begin{aligned} -\Delta u + 30\, u^3 + u &= f + q \qquad && \text{in } \Omega, \\ u &= 0 && \text{on } \partial\Omega, \end{aligned} \tag{5.2}$$

where $\Omega = \omega = (0,1)^2 \setminus [0.4, 0.6]^2$, $V = H^1_0(\Omega)$, $Q = L^2(\Omega)$, and the admissible set Q_{ad} is given by

$$Q_{\mathrm{ad}} = \{q \in Q \mid -7 \le q(x) \le 20 \text{ a.e. on } \Omega\}.$$

The desired state u^d and the right-hand side f are defined as

$$u^d(x) = x_1 \cdot x_2, \quad f(x) = \left((x_1 - 0.5)^2 + (x_2 - 0.5)^2\right)^{-1},$$

and the regularization parameter is chosen as $\alpha = 10^{-4}$. We note that the state equation (5.2) is a monotone semilinear equation, which possesses a unique solution $u \in V$ for each $q \in Q$. The proof of the existence of a global solution as well as derivation of necessary and sufficient optimality conditions for the corresponding optimization problem (5.1)–(5.2) can be found, e.g., in [34].

In section 4 we derived two different error estimators for the error with respect to the cost functional and one error estimator with respect to a quantity of interest. In this example, we choose the quantity of interest as

$$I(q,u) = \frac{1}{2}\int_{(0.7,0.8)^2} |\nabla u(x)|^2\,dx + \int_{(0.2,0.3)^2} q(x)\,dx. \tag{5.3}$$

In order to check the quality of the error estimators, we define the following effectivity indices:

$$I_{\text{eff}}(\eta_1) = \frac{J(u)-J(u_h)}{\eta_1}, \quad I_{\text{eff}}(\eta_2) = \frac{J(u)-J(u_h)}{\eta_2}, \quad I_{\text{eff}}(\eta_{\text{QI}}) = \frac{I(q,u)-I(q_h,u_h)}{\eta_{\text{QI}}}. \tag{5.4}$$

In Table 5.1 these effectivity indices are listed for different types of mesh refinement: random refinement and refinement based on the error estimator η_{QI} for the quantity of interest.

TABLE 5.1
Effectivity indices.

N	$I_{\text{eff}}(\eta_1)$	$I_{\text{eff}}(\eta_2)$	$I_{\text{eff}}(\eta_{\text{QI}})$
432	1.1	1.1	1.2
906	1.1	1.1	1.1
2328	1.3	1.2	2.3
5752	1.2	1.2	1.4
13872	1.3	1.3	1.5
33964	1.3	1.3	1.4
83832	1.2	1.2	1.5

(a) Random refinement

N	$I_{\text{eff}}(\eta_1)$	$I_{\text{eff}}(\eta_2)$	$I_{\text{eff}}(\eta_{\text{QI}})$
432	1.1	1.1	1.1
824	1.1	1.1	1.4
1692	1.0	1.0	0.3
3992	1.0	1.0	0.2
11396	1.0	1.0	0.5
30604	1.0	1.0	1.0
80354	1.0	1.0	1.3

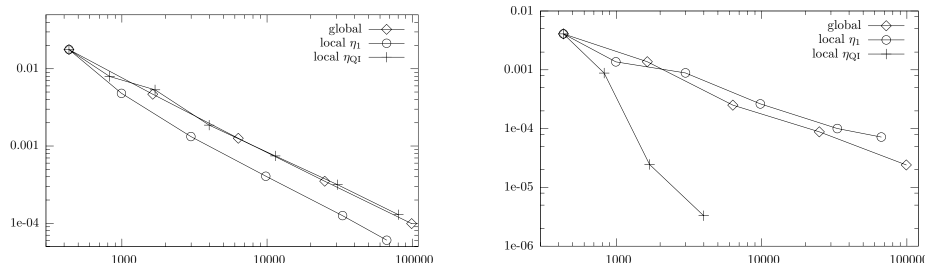
(b) Refinement according to η_{QI}

We observe that the error estimators provide quantitative information about the discretization error. We note that the results for η_1 and η_2 are very close to each other in this example; cf. Remark 4.3.

In addition, our results show that the local mesh refinement based on error estimators derived above leads to substantial saving in degrees of freedom for achieving a given level of the discretization error. In Figure 5.1 the dependence of discretization error on the number of degrees of freedom is shown for different refinement criteria: global (uniform) refinement, refinement based on the error estimator η_1 for the cost functional, and refinement based on the error estimator η_{QI} for the quantity of interest. In Figure 5.1(a) the error with respect to the cost functional (5.1) and in Figure 5.1(b) the error with respect to the quantity of interest (5.3) are considered, respectively.

We observe the best behavior of error with respect to the cost functional if the mesh is refined based on η_1 and the best behavior of error with respect to the quantity of interest for the refinement based on η_{QI}.

A series of meshes generated according to the information obtained from the error estimators are shown in Figure 5.2 together with the optimal control q and the corresponding state u.



(a) Error in J

(b) Error in I

FIG. 5.1. *Discretization error for different refinement criteria.*

5.2. Example 2. Our second example is motivated by a parameter identification problem. The minimization problem is given by

$$\text{Minimize} \quad \frac{1}{2}\|u - u^d\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|q\|^2_{L^2(\Omega)}, \qquad u \in V,\ q \in Q_{\text{ad}}, \tag{5.5}$$

subject to

$$\begin{aligned} -\Delta u + qu &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{5.6}$$

where $\Omega = \omega = (0, 0.5) \times (0, 1) \cup (0, 1) \times (0.5, 1)$, $V = H^1_0(\Omega)$, $Q = L^2(\Omega)$, and the admissible set Q_{ad} is given by

$$Q_{\text{ad}} = \{q \in Q \,|\, q_-(x) \leq q(x) \leq q_+(x) \text{ a.e. on } \Omega\}, \text{ with } q_-(x) = 0, \quad q_+(x) = 0.3\,.$$

The desired state u^d and the right-hand side f are defined as

$$u^d(x) = \frac{1}{8\pi^2}\sin(2\pi x_1)\sin(2\pi x_2), \quad f(x) = 1,$$

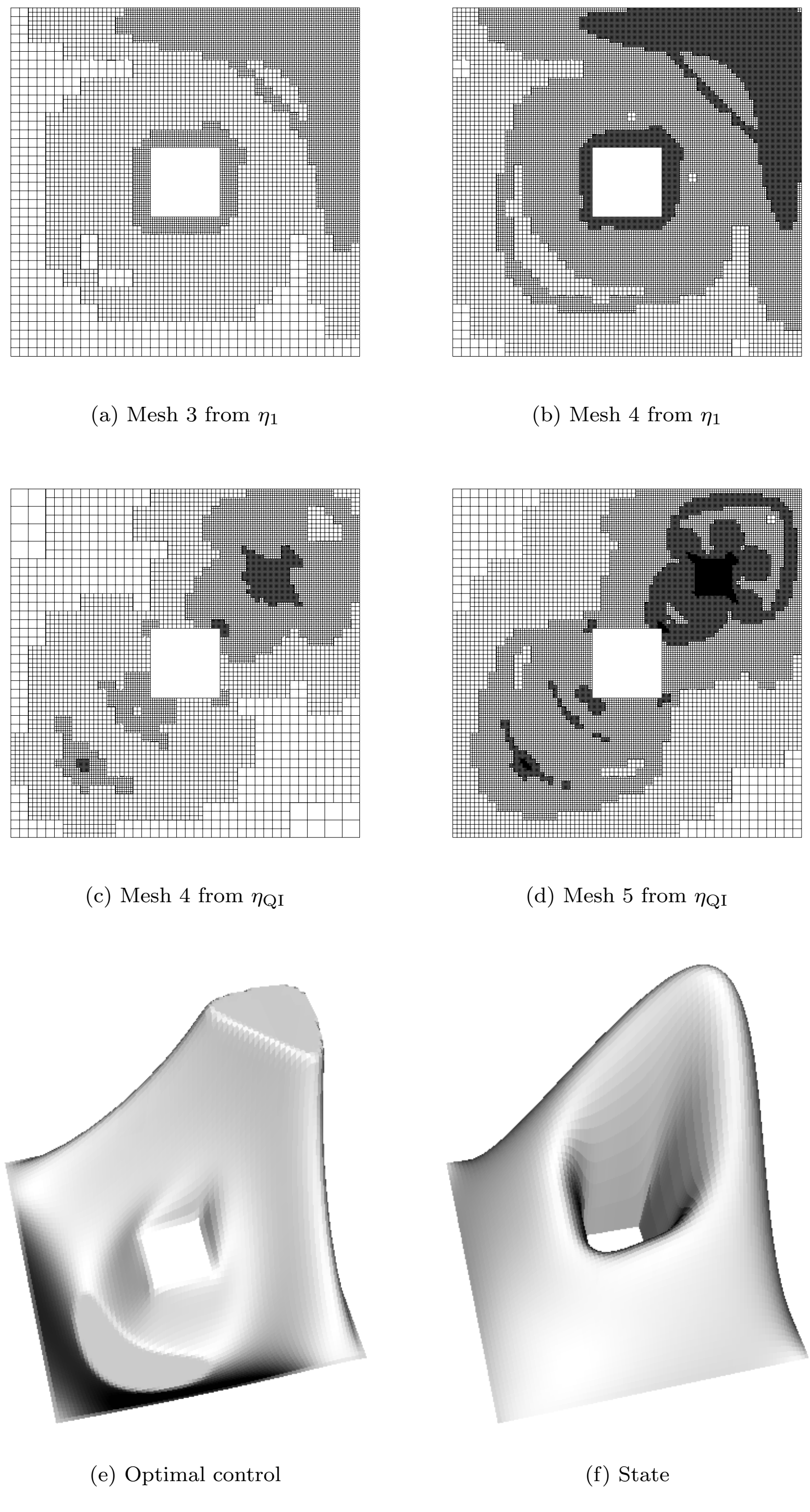
and the regularization parameter is chosen $\alpha = 10^{-4}$. Note that for any given $q \in Q_{\text{ad}}$ the state equation (5.6) possesses a unique solution $u \in V$ due to $q \geq 0$.

We are interested in the error in the unknown parameter, and thus we choose

$$I(q, u) = \int_{\Omega_O} q(x)\,dx,$$

where $\Omega_O = (0, 0.25) \times (0.75, 1)$.

In Table 5.2 the effectivity indices, defined as in (5.4), are listed for different types of mesh refinement: global (uniform) refinement, random refinement, refinement based on the error estimator η_1 for the cost functional, and refinement based on the error estimator η_{QI} for the quantity of interest. As in the first example we observe that the error estimators provide quantitative information on the discretization errors.



(a) Mesh 3 from η_1

(b) Mesh 4 from η_1

(c) Mesh 4 from η_{QI}

(d) Mesh 5 from η_{QI}

(e) Optimal control

(f) State

FIG. 5.2. *Locally refined meshes and solution.*

TABLE 5.2
Effectivity indices.

N	$I_{\text{eff}}(\eta_1)$	$I_{\text{eff}}(\eta_2)$	$I_{\text{eff}}(\eta_{\text{QI}})$
65	1.2	1.2	2.0
225	1.3	1.2	1.9
833	1.4	1.4	1.5
3201	1.5	1.5	1.7

(a) Global refinement

N	$I_{\text{eff}}(\eta_1)$	$I_{\text{eff}}(\eta_2)$	$I_{\text{eff}}(\eta_{\text{QI}})$
65	1.2	1.2	2.0
225	1.3	1.3	1.9
785	1.4	1.4	1.6
2705	1.5	1.5	1.7

(b) Refinement according to η_1

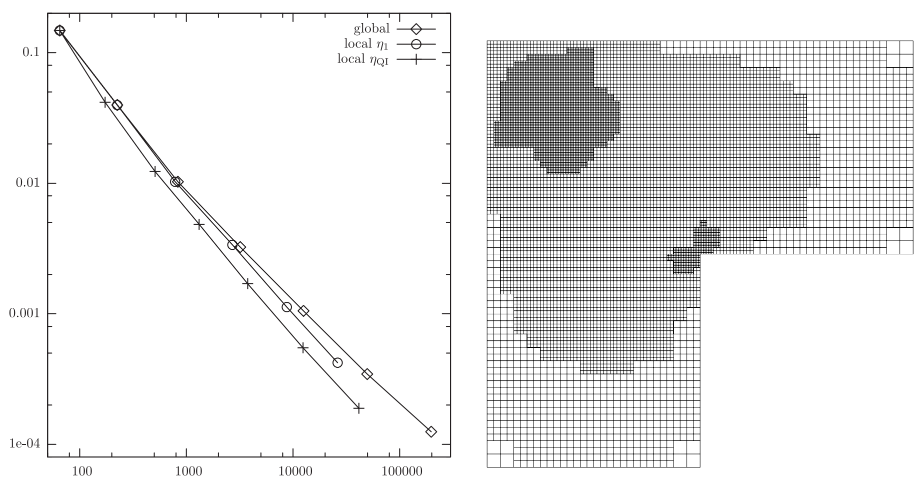
N	$I_{\text{eff}}(\eta_1)$	$I_{\text{eff}}(\eta_2)$	$I_{\text{eff}}(\eta_{\text{QI}})$
65	1.2	1.2	2.0
141	1.2	1.2	2.0
307	1.2	1.2	0.5
763	1.4	1.4	2.0

(c) Random refinement

N	$I_{\text{eff}}(\eta_1)$	$I_{\text{eff}}(\eta_2)$	$I_{\text{eff}}(\eta_{\text{QI}})$
65	1.2	1.2	2.0
173	1.2	1.2	1.8
509	1.2	1.2	1.3
1317	1.2	1.2	1.3

(d) Refinement according to η_{QI}

From Figure 5.3(a), where the discretization error with respect to the quantity of interest is plotted for different refinement criteria, we again observe that the local mesh refinement based on the appropriate error estimator leads to a certain saving in degrees of freedom for achieving a given tolerance for the discretization error. A typical mesh generated using the information obtained from η_{QI} is shown in Figure 5.3(b).



(a) Error in I (b) Mesh 5 from η_{QI}

FIG. 5.3. *Discretization error and mesh.*

REFERENCES

[1] N. ARADA, E. CASAS, AND F. TRÖLTZSCH, *Error estimates for the numerical approximation of a semilinear elliptic control problem*, Comput. Optim. Appl., 23 (2002), pp. 201–229.

[2] R. Becker, *Estimating the control error in the discretized PDE-constrained optimization*, J. Numer. Math., 14 (2006), pp. 163–185.
[3] *The Finite Element Toolkit* Gascoinge; http://www.gascoigne.uni-hd.de/.
[4] R. Becker, H. Kapp, and R. Rannacher, *Adaptive finite element methods for optimal control of partial differential equations: Basic concept*, SIAM J. Control Optim., 39 (2000), pp. 113–132.
[5] *A C++ Library for Optimization with Stationary and Nonstationary PDEs*; http://www.rodobo.uni-hd.de/.
[6] R. Becker and R. Rannacher, *An optimal control approach to a posteriori error estimation*, in Acta Numerica 2001, A. Iserles, ed., Cambridge University Press, Cambridge, UK, 2001, pp. 1–102.
[7] R. Becker and B. Vexler, *A posteriori error estimation for finite element discretizations of parameter identification problems*, Numer. Math., 96 (2004), pp. 435–459.
[8] R. Becker and B. Vexler, *Mesh refinement and numerical sensitivity analysis for parameter calibration of partial differential equations*, J. Comput. Phys., 206 (2005), pp. 95–110.
[9] M. Bergounioux, K. Ito, and K. Kunisch, *Primal-dual strategy for constrained optimal control problems*, SIAM J. Control Optim., 37 (1999), pp. 1176–1194.
[10] S. Brenner and R.L. Scott, *The Mathematical Theory of Finite Element Methods*, Springer-Verlag, Berlin, Heidelberg, New York, 1994.
[11] E. Casas, M. Mateos, and F. Tröltzsch, *Error estimates for the numerical approximation of boundary semilinear elliptic control problems*, Comput. Optim. Appl., 31 (2005), pp. 193–220.
[12] E. Casas and F. Tröltzsch, *Error estimates for linear-quadratic elliptic control problems*, in Analysis and Optimization of Differential Systems (Constanta, 2002), Kluwer Academic Publishers, Boston, MA, 2003, pp. 89–100.
[13] K. Eriksson, D. Estep, P. Hansbo, and C. Johnson, *Introduction to adaptive methods for differential equations*, in Acta Numerica 1995, A. Iserles, ed., Cambridge University Press, Cambridge, UK, 1995, pp. 105–158.
[14] K. Eriksson, D. Estep, P. Hansbo, and C. Johnson, *Computational Differential Equations*, Cambridge University Press, Cambridge, UK, 1996.
[15] R. S. Falk, *Approximation of a class of optimal control problems with order of convergence estimates*, J. Math. Anal. Appl., 44 (1973), pp. 28–47.
[16] A. V. Fursikov, *Optimal Control of Distributed Systems: Theory and Applications*, Transl. Math. Monogr. 187, AMS, Providence, RI, 2000.
[17] A. Gaevskaya, R. H. W. Hoppe, Y. Iliash, and M. Kieweg, *Convergence analysis of an adaptive finite element method for distributed control problems with control constraints*, in Control of Coupled Partial Differential Equations, Birkhäuser, Basel, 2007, pp. 47–68.
[18] T. Geveci, *On the approximation of the solution of an optimal control problem governed by an elliptic equation*, RAIRO Anal. Numér., 13 (1979), pp. 313–328.
[19] R. Griesse and B. Vexler, *Numerical sensitivity analysis for the quantity of interest in PDE-constrained optimization*, SIAM J. Sci. Comput., 29 (2007), pp. 22–48.
[20] M. Hintermüller, R. H. W. Hoppe, Y. Iliash, and M. Kieweg, *An a posteriori error analysis of adaptive finite element methods for distributed elliptic control problems with control constraints*, ESAIM Control Optim. Calc. Var., to appear.
[21] M. Hintermüller, K. Ito, and K. Kunisch, *The primal-dual active set strategy as a semismooth Newton method*, SIAM J. Optim., 13 (2003), pp. 865–888.
[22] M. Hinze, *A variational discretization concept in control constrained optimization: The linear-quadratic case*, Comput. Optim. Appl., 30 (2005), pp. 45–61.
[23] R. H. W. Hoppe, Y. Iliash, C. Iyyunni, and N. H. Sweilam, *A posteriori error estimates for adaptive finite element discretizations of boundary control problems*, J. Numer. Math., 14 (2006), pp. 57–82.
[24] K. Kunisch and A. Rösch, *Primal-dual active set strategy for a general class of constrained optimal control problems*, SIAM J. Optim., 13 (2002), pp. 321–334.
[25] R. Li, W. Liu, H. Ma, and T. Tang, *Adaptive finite element approximation for distributed elliptic optimal control problems*, SIAM J. Control Optim., 41 (2002), pp. 1321–1349.
[26] J.-L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Grundlehren Math. Wiss. 170, Springer-Verlag, Berlin, 1971.
[27] W. Liu and N. Yan, *A posteriori error estimates for distributed convex optimal control problems*, Adv. Comput. Math., 15 (2001), pp. 285–309.
[28] K. Malanowski, *Convergence of approximations versus regularity of solutions for convex, control-constrained optimal-control problems*, Appl. Math. Optim., 8 (1982), pp. 69–95.
[29] H. Maurer and J. Zowe, *First and second-order necessary and sufficient optimality conditions*

for infinite-dimensional programming problems, Math. Programming, 16 (1979), pp. 98–110.

[30] D. Meidner and B. Vexler, *Adaptive space-time finite element methods for parabolic optimization problems*, SIAM J. Control Optim., 46 (2007), pp. 116–142.

[31] C. Meyer and A. Rösch, *Superconvergence properties of optimal control problems*, SIAM J. Control Optim., 43 (2004), pp. 970–985.

[32] C. Meyer and A. Rösch, *L^∞-estimates for approximated optimal control problems*, SIAM J. Control Optim., 44 (2005), pp. 1636–1649.

[33] A. Rösch, *Error estimates for linear-quadratic control problems with control constraints*, Optim. Methods Softw., 21 (2006), pp. 121–134.

[34] F. Tröltzsch, *Optimale Steuerung partieller Differentialgleichungen*, Vieweg, Wiesbaden, Germany, 2005.

[35] R. Verfürth, *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*, Wiley/Teubner, New York, Stuttgart, 1996.

[36] B. Vexler, *Adaptive Finite Elements for Parameter Identification Problems*, Ph.D. thesis, Institut für Angewandte Mathematik, Universität Heidelberg, Heidelberg, Germany, 2004.