Comput Optim Appl (2011) 48: 581–600
DOI 10.1007/s10589-009-9278-x

# Discretization of interior point methods for state constrained elliptic optimal control problems: optimal error estimates and parameter adjustment

**Michael Hinze · Anton Schiela**

Received: 31 December 2008 / Published online: 28 July 2009


**Abstract** An adjustment scheme for the relaxation parameter of interior point approaches to the numerical solution of pointwise state constrained elliptic optimal control problems is introduced. The method is based on error estimates of an associated finite element discretization of the relaxed problems and optimally selects the relaxation parameter in dependence on the mesh size of discretization. The finite element analysis for the relaxed problems is carried out and a numerical example is presented which confirms our analytical findings.

**Keywords** Elliptic optimal control problem · Error estimates · Interior point method · Pointwise state constraints

## 1 Introduction and basic results

Let $\Omega \subset \mathbb{R}^d$ $(d = 2, 3)$ be a bounded domain with a Lipschitz boundary $\partial\Omega$. In this note, we are interested in the following control problem:

$$\begin{aligned} &\min_{w \in U} J(w) = \frac{1}{2}\int_\Omega \left|\mathcal{G}(Bw) - y_0\right|^2 + \frac{\alpha}{2}\|w\|_U^2 \\ &\text{subject to } \mathcal{G}(Bw) \le \overline{y} \quad \text{a.e. in } \Omega. \end{aligned} \tag{1.1}$$

A. Schiela is supported by the DFG Research Center MATHEON "Mathematics for key technologies".

M. Hinze (✉)
Schwerpunkt Optimierung und Approximation, Universität Hamburg, Bundesstraße 55, 20146 Hamburg, Germany
e-mail: michael.hinze@uni-hamburg.de

A. Schiela
Konrad Zuse Zentrum Berlin, Takustraße 7, 14195 Berlin, Germany

Here and throughout, "a.e." stands for "almost everywhere". We suppose that $\alpha > 0$, $y_0 \in L_2(\Omega)$ and $\overline{y} \in W^{2,\infty}(\Omega)$ are given. Further $(U, (\cdot,\cdot)_U)$ denotes a Hilbert space which we identify with its dual, and $V$ is a closed subspace of $H^1(\Omega)$ with $H_0^1(\Omega) \subset V$. $B : U \to V'$ a linear, continuous operator. For given $f \in V'$ the function $y = \mathcal{G}(f) \in V$ denotes the unique weak solution to the elliptic boundary value problem

$$a(y, v) = \langle f, v\rangle_{V',V} \quad \forall v \in V$$

$$a(y, v) := \int_\Omega \left( \sum_{i,j=1}^d a_{ij}(x) y_{x_i} v_{x_j} + \sum_{i=1}^d b_i(x) y_{x_i} v + c(x) y v \right) dx, \qquad y, v \in V, \tag{1.2}$$

where subsequently we assume that, the coefficients $a_{ij}, b_i$ and $c$ are bounded functions in $\overline{\Omega}$, and that there exists $c_0 > 0$ such that

$$\sum_{i,j=1}^d a_{ij}(x)\xi_i\xi_j \ge c_0|\xi|^2 \quad \text{for all } \xi \in \mathbb{R}^d \text{ and all } x \in \Omega.$$

Furthermore we suppose that $a$ is coercive on $V$, i.e., there exists $c_1 > 0$ such that

$$a(v, v) \ge c_1\|v\|_{H^1}^2 \quad \text{for all } v \in V. \tag{1.3}$$

Since (1.1) represents a minimization problem over a closed and convex set with quadratic objective it admits a unique solution $u \in U$ with corresponding unique state $y = \mathcal{G}(Bu)$. This follows from a standard result of optimization theory (cf. e.g. [10, Proposition II.1.2]).

In order to discuss first order optimality conditions for our problem, which are the basis for all our further considerations, we have to impose a finer topological framework for the states, induced by the maximum-norm.

Let now $\infty > q > d \ge 2$, which implies that the embedding $W^{1,q}(\Omega) \hookrightarrow C(\overline{\Omega})$ is continuous. Since $W^{1,q}(\Omega)$ is continuously embedded into $H^1(\Omega)$, and $V$ is closed in $H^1(\Omega)$, the space $V_q := W^{1,q}(\Omega) \cap V$ (the preimage of $V$ with respect to this embedding) is a closed subspace of $W^{1,q}(\Omega)$. Let $q' := q/(q-1)$ be the conjugate exponent, and let $V_{q'}$ be the closure of $V_q$ in $W^{1,q'}(\Omega)$. Then both $V_q$ and $V_{q'}$ are closed subspaces of reflexive spaces and hence reflexive (and in particular complete).

Moreover, for $y \in V_q$ the expression $a(y, \cdot)$ is a continuous linear functional on $V_{q'}$, since

$$\|a(y, \cdot)\|_{V'_{q'}} = \sup_{v \in V_{q'}} \frac{a(y, v)}{\|v\|_{V_{q'}}} \le C\|y\|_{V_q}$$

by boundedness of $a_{ij}, b_i$, and $c$ and by the Hölder inequality. Hence, the mapping

$$A : V_q \to V'_{q'},$$
$$y \to a(y, \cdot)$$

is continuous. It is injective, because $y \in V_q \subset V$ and $a(y, y) = 0$ implies $y = 0$ by coercivity of $a$.

**Assumption 1.1** For given $q > d$, $A : V_q \to V'_{q'}$ is surjective, and $B : U \to V'_{q'}$ is continuous.

Then, by the open mapping theorem, $A$ has a continuous inverse $A^{-1} : V'_{q'} \to V_q$, which is clearly a restriction of $\mathcal{G}$.

*Remark 1.2* Under very mild assumptions on the boundary conditions (which hold in particular for pure Neumann and for pure Dirichlet conditions) it follows that for given $a_{ij}$ there is some $q > 2$, such that $A$ is surjective (cf. [11]). Then for $d = 2$ our assumptions on $A$ hold as long as $q$ is chosen small enough. For $d = 3$ the smoothness requirements are a little bit stronger in order to achieve surjectivity for some $q > 3$. Here the results are relatively diverse (cf. [1, Theorem 9.2] for continuous coefficients and smooth domains, [12] for certain problems with jumping coefficients, and [19] for the Dirichlet problem for the Laplacian on Lipschitz domains).

*Example 1.3* There are several examples for the choice of $B$ and $U$, for which Assumption 1.1 holds.

(i) Distributed control: $U = L^2(\Omega)$, $B = Id : L^2(\Omega) \to V'_{q'}$.
(ii) Boundary control: $U = L^2(\partial\Omega)$, $Bu(\cdot) = \int u\gamma_0(\cdot)\, dx : L^2(\Omega) \to V'_{q'}$, where $\gamma_0$ is the boundary trace operator in $V$. In this case Assumption 1.1 holds only in the case $d = 2$.
(iii) Linear combinations of input fields: $U = \mathbb{R}^n$, $Bu = \sum_{i=1}^n u_i f_i$, $f_i \in V'_{q'}$.

Since, by our choice of $q$, $V_q$ is continuously embedded into $C(\overline{\Omega})$ the mapping $\mathcal{G} \circ B : U \to V_q \subset C(\overline{\Omega})$ is continuous in this case.

In addition to these minimal regularity requirements, we will impose assumptions later to conclude a-priori error estimates. We note already here that uniform error estimates for the state will play a central in this context.

To discuss first order optimality conditions, we suppose a Slater condition of the type

**Assumption 1.4**

$$\text{There exist } \hat{u} \in U,\ \tau > 0: \quad \mathcal{G}(B\hat{u}) \le \overline{y} - \tau.$$

Further, let $\mathcal{M}(\overline{\Omega})$ denote the space of Radon measures, which is the representation of the dual space of $C(\overline{\Omega})$. It is endowed with the norm

$$\|\lambda\|_{\mathcal{M}(\overline{\Omega})} = \sup_{f \in C(\overline{\Omega}), |f| \le 1} \int_{\overline{\Omega}} f\, d\lambda.$$

First order optimality conditions for state constrained problems are well investigated [5, 6], where elliptic PDEs are considered under assumptions which slightly differ from those taken here.

**Theorem 1.5** *Let Assumption 1.1 and 1.4 be satisfied. There exists a positive measure $\lambda \in \mathcal{M}(\overline{\Omega})$ and a function $p \in V_{q'}$ which together with $y = \mathcal{G}(Bu)$ satisfies the dual system*

$$a(v, p) = \int_\Omega (y - y_0)v + \int_{\overline{\Omega}} v\, d\lambda \quad \forall v \in V_q, \tag{1.4}$$

$$B^* p + \alpha u = 0 \quad \text{in } U, \tag{1.5}$$

$$\lambda \ge 0, \qquad y \le \overline{y} \quad \text{a.e. in } \Omega \text{ and } \int_{\overline{\Omega}} (\overline{y} - y)\, d\lambda = 0. \tag{1.6}$$

*Proof* Let $C_V$ be the closure of $V_q$ in $C(\overline{\Omega})$. Then $\mathcal{G} \circ B \to C_V$ is continuous. Denote by $\iota_F$ the indicator function of the feasible set $F = \{y \in C_V : y \le \overline{y}\}$. This means $\iota_F(y) = 0$ for $y \in F$, and $\iota_F(y) = +\infty$ otherwise.

Let $\partial(\cdot)$ denote the subdifferential of convex analysis. Then by the well known relation $u \in \operatorname{argmin} f \Leftrightarrow 0 \in \partial f(u)$, the minimizer $u$ satisfies

$$0 \in \partial(J + \iota_F \circ \mathcal{G} \circ B)(u).$$

By continuity of $J$ we can apply the sum-rule of convex analysis (cf. e.g. [10, Proposition I.5.6]) to compute

$$\partial(J + \iota_F \circ \mathcal{G} \circ B) = \partial J + \partial(\iota_F \circ \mathcal{G} \circ B),$$

and by our Slater condition we obtain by the chain-rule (cf. e.g. [10, Proposition I.5.7]), setting $y = \mathcal{G} \circ Bu$,

$$\partial(\iota_F \circ \mathcal{G} \circ B)(u) = (\mathcal{G}B)^* \partial\iota_F(y) = B^* \mathcal{G}^* \partial\iota_F(y).$$

Here $\partial\iota_F(y) \subset C_V'$, $\mathcal{G}^* : C_V' \to V_{q'}$ and $B^* : V_{q'} \to U$. Similarly, by continuity of $J$ we obtain

$$\partial J(u) = \partial\big(1/2\|\mathcal{G}Bu - y_0\|^2 + \alpha/2\|u\|^2\big) = B^*\mathcal{G}^*(y - y_0) + \alpha u.$$

Hence, there is $m \in \partial\iota_F(y)$, such that $0 = B^*\mathcal{G}^*(y - y_0 + m) + \alpha u$ holds in $U' = U$. Setting $r := y - y_0 + m$ and $p := \mathcal{G}^* r$ we obtain $0 = B^* p + \alpha u$ and $\langle f, p\rangle = \langle r, \mathcal{G} f\rangle$ for all $f \in V_{q'}'$. Setting $v = \mathcal{G} f \in V_q$ such that $a(v, p) = \langle f, p\rangle$, we obtain

$$a(v, p) = \langle f, p\rangle = \langle r, \mathcal{G} f\rangle = \langle r, v\rangle = \langle y - y_0 + m, v\rangle.$$

Since $C_V$ is a closed subspace of $C(\overline{\Omega})$, $m$ has a representation as a measure $\lambda$. Since $m \in \partial\iota_F$, it follows that $\langle m, z - y\rangle \le 0$ for all $z \in F$. From this, positivity of $\lambda$ and (1.6) are easily derived. □

The assertion $\lambda \in \mathcal{M}(\overline{\Omega})$ can be improved in concrete examples. If, for example $V = H_0^1(\Omega)$, then the subspace $C_V$ in the proof is the space of continuous functions on $\Omega$ that vanish on $\partial\Omega$. Then every measure that vanishes on $\Omega$ is equivalent to the zero-functional in $C_V$, and thus $\lambda$ can be identified with an element of $\mathcal{M}(\Omega)$ in this case.

A finite element analysis of problem (1.1) within a slightly different framework is carried out in e.g. [7–9] for $\operatorname{Im} B \subseteq L^2(\Omega)$. There the following error bounds are derived;

$$\|u-u_h\|_U,\ \|y-y_h\|_{H^1} = \begin{cases} O(h^{\frac{1}{2}}), & \text{if } d=2,\\ O(h^{\frac{1}{4}}), & \text{if } d=3, \end{cases} \tag{1.7}$$

where $u_h$ and $y_h$ denote the discrete optimal control and state, respectively. Moreover,

$$\begin{aligned} &\|u-u_h\|_U,\ \|y-y_h\|_{H^1} \\ &\quad \le C \begin{cases} h^{\frac{3}{2}-\frac{d}{2q'}}\sqrt{|\log h|}, & \text{if } Bu \in W^{1,q'}(\Omega),\\ h|\log h|, & \text{if } Bu,\, Bu_h \in L^\infty(\Omega) \text{ uniformly for } d=2,3. \end{cases} \end{aligned} \tag{1.8}$$

In the present paper, our aim is to investigate a finite element approximation of an interior point technique for the numerical solution of (1.1) and to provide optimal adjustment strategies for the relaxation parameter with respect to the finite element mesh size.

The regularized version of (1.1) considered in this paper reads

$$\min_{w\in U} J_\mu(w) = \frac{1}{2}\int_\Omega \left|\mathcal{G}(Bw)-y_0\right|^2 + \frac{\alpha}{2}\|w\|_U^2 + \int_\Omega l\big(y(x);\mu;k\big)dx. \tag{1.9}$$

Here $l$ is a function, chosen according to the following definition:

**Definition 1.6** For all $k \ge 1$ and $\mu > 0$ the functions $l(y;\mu;k) : \mathbb{R}_+ \to \mathbb{R}\cup\{+\infty\}$ defined by

$$l(y;\mu;k) := \begin{cases} -\mu \ln(\overline{y}-y): & k=1,\\ \frac{\mu^k}{(k-1)(\overline{y}-y)^{k-1}}: & k>1 \end{cases}$$

are called *barrier functions of order* $k$. We extend their domain of definition to $\mathbb{R}$ by setting $l(y;\mu;k)=\infty$ for $z \le 0$.

Further, we define a *barrier functional* by writing:

$$b(y;\mu;k) := \int_\Omega l(y(x);\mu;k)\,dx.$$

In the following we will drop the argument $k$ for the sake of brevity.

The rest of the paper is organized as follows: In Sect. 2 we collect basic results on (1.9). In Sect. 3 we present the finite element analysis of problem (1.9), deriving error bounds that depend on $L_\infty$ approximation errors for the states. In Sect. 4 we apply these results to a couple of specific settings. Among other aspects we prove the error bounds

$$\begin{aligned} &\left\|y^\mu - y_h^\mu\right\|_{H^1} + \left\|u^\mu - u_h^\mu\right\|_U \\ &\quad \le C \begin{cases} h^{1-\frac{d}{4}}, & \text{if } Bu^\mu \in L^2(\Omega),\\ h^{\frac{3}{2}-\frac{d}{2q'}}\sqrt{|\log h|}, & \text{if } Bu^\mu \in W^{1,q'}(\Omega),\\ h|\log h|, & \text{if } Bu^\mu,\, Bu_h^\mu \in L^\infty(\Omega) \text{ uniformly for } d=2,3, \end{cases} \end{aligned}$$

where $y_h^\mu, u_h^\mu$ denote the finite element approximations to $y^\mu$ and $u^\mu$, respectively and $C > 0$ is independent of $\mu$ and $h$. We note that these estimates are in the spirit of (1.7), (1.8). In Sect. 5 we discuss the overall errors

$$\|y - y_h^\mu\|_{H^1} \sim \|y - y^\mu\|_{H^1} + \|y^\mu - y_h^\mu\|_{H^1} \quad \text{and}$$
$$\|u - u_h^\mu\|_U \sim \|u - u^\mu\|_U + \|u^\mu - u_h^\mu\|_U$$

and propose a-priori strategies for balancing $\mu$ and $h$. In Sect. 6 we present numerical results which confirm our theoretical findings.

Let us comment on further approaches that tackle optimization problems for pdes with control and state constraints. In [20] Meyer considers a fully discrete strategy to approximate an elliptic control problem with pointwise state and control constraints. He obtains the approximation order $\mathcal{O}(h^{2-d/2-\epsilon})$ for the state in $H^1$ and for the control in $L^2$, where $d$ denotes the spatial dimension and $\epsilon > 0$ can be chosen arbitrarily. His results confirm those obtained by Deckelnick and the first author in [8] for the purely state constrained case. A *Lavrentiev-type regularization* of problem (1.1) is investigated in [22]. In this approach the state constraint $y \le b$ in (1.1) is replaced by the mixed constraint $\epsilon u + y \le b$, with $\epsilon > 0$ denoting a regularization parameter. It turns out that the associated Lagrange multiplier $\lambda_\epsilon$ belongs to $L^2(\Omega)$. Numerical analysis for this approach with emphasis on the coupling of gridsize and regularization parameter $\epsilon$ is presented by the first author and Meyer in [17]. The resulting optimization problems are solved either by interior-point methods or primal-dual active set strategies, compare [21]. Parameter adjustment strategies for Moreau-Yosida regularization of problem (1.1) are proposed by Hintermüller and the first author in [13]. The development of numerical approaches to tackle (1.1) is ongoing. A path-following method, based on quadratic penalization and semi-smooth Newton methods is proposed by Hintermüller and Kunisch in [14, 15]. An introductory text is provided by Tröltzsch with [30].

## 2 The regularized problem

Problems of the type (1.9) are analyzed in [27]. Let us start with extending this analysis to the more general case considered here, where we use Assumption 1.1 and Assumption 1.4. For the convenience of the reader we recall the definition of the subdifferential $\partial f(w)$ of a function $f: W \to \overline{\mathbb{R}}$ at a point $w$ as the subset of all $w' \in W'$ which satisfy $\langle w', w - \tilde{w}\rangle_{W',W} \le f(w) - f(\tilde{w})\ \forall \tilde{w} \in W$. Here, $W$ denotes a normed space with its normed dual $W'$. If $w' \in \partial f(w)$ and $\tilde{w}' \in \partial f(\tilde{w})$, then

$$\langle w' - \tilde{w}', w - \tilde{w}\rangle_{W',W} \le \big[f(w) - f(\tilde{w})\big] - \big[f(\tilde{w}) - f(w)\big] = 0. \tag{2.1}$$

**Theorem 2.1** *For each $\mu \ge 0$ problem (1.9) admits a unique solution $u^\mu \in U$ with associated state $y^\mu = \mathcal{G}(Bu^\mu)$, which is strictly feasible a.e. in $\Omega$. Furthermore there exists a function $p^\mu \in V_{q'}$ and a regular Borel measure $\lambda^\mu \in M(\overline{\Omega})$ which satisfy the system*

$$a\big(v, p^\mu\big) = \int_\Omega \big(y^\mu - y_0\big)v + \int_\Omega v\, d\lambda^\mu \quad \forall v \in V_q, \tag{2.2}$$

$$B^* p^\mu + \alpha u^\mu = 0 \quad \text{in } U. \tag{2.3}$$

*Here* $\lambda^\mu \in \partial b(y^\mu; \mu)$ *is an element of the subdifferential of* $b(y^\mu; \mu)$.

*Proof* After having established lower-semi continuity of $b(\cdot; \mu)$ (cf. [27, Lemma 3.2]), existence of minimizers is readily established as in the unregularized limit case, see e.g. [5]. Also the derivation of first order optimality conditions is performed just as in the proof of Theorem 1.5 with the only difference that $\iota_F(y)$ is now replaced by $b(y; \mu)$. Strict feasibility of $y^\mu$ a.e. in $\Omega$ follows from $l(y; \mu) \to +\infty$ for $y \to \overline{y}$. □

The subdifferential $\partial b(y; \mu)$ has been analyzed in [27, Proposition 3.5] and exhibits the following structure. The measure $\lambda^\mu \in M(\overline{\Omega})$ is non-negative, and splits into two parts:

$$\int_{\overline{\Omega}} v \, d\lambda^\mu = -\int_\Omega l'(y^\mu; \mu) v \, dx + \int_{y^\mu = \overline{y}} v \, d\lambda, \tag{2.4}$$

with $l'(y^\mu; \mu) \in L_1(\Omega)$. The measure $\lambda$ is non-negative and vanishes if $y^\mu$ is strictly feasible. In this case $\lambda^\mu$ and $p^\mu$ are uniquely defined. As a consequence the following complementarity condition

$$\int_{\overline{\Omega}} y^\mu - \overline{y} \, d\lambda = 0 \tag{2.5}$$

holds. Furthermore, from [27, Proposition 4.5] we deduce

$$\|\lambda^\mu\|_{M(\overline{\Omega})}, \quad \|l'(y^\mu; \mu)\|_{L_1(\Omega)}, \|\lambda\|_{M(\overline{\Omega})}, \quad \text{and} \quad \|p^\mu\|_{W^{1,q'}(\Omega)} \le C \tag{2.6}$$

with some positive constant $C$ that is independent of $\mu$.

*Remark 2.2* The potential occurrence of the additional measure $\lambda$ in (2.4) is the main motivation to consider rational barrier functionals. In [27] it is shown that for sufficiently high *order* $k$, the non-regular part $\lambda$ vanishes for all $\mu > 0$. This order $k$ depends on the dimension of the problem and on the regularity of $y^\mu$, and can be chosen a-priori. Our framework asserts that $y^\mu$ is bounded in $W^{1,q}(\Omega)$, and thus is Hölder-continuous, i.e. $y^\mu \in C^{0,\beta}(\overline{\Omega})$ for some $\beta > 0$. This ensures that choosing $k$ large enough guarantees strict feasibility of $y^\mu$. In this work, however, we can dispense with a strict feasibility assumption.

The convergence analysis of the regularization path is also covered by the results of [27]. In Lemma 5.1 there it is proven the function values converge linearly in $\mu$:

$$J(u^\mu) - J(u) \le C\mu.$$

The convergence results for $u$ in [27, Theorem 5.3]:

$$\|u^\mu - u\|_U \le C\sqrt{\mu}$$

is then an easy conclusion from the strict convexity of $J$. From

$$c_1\|y-y^\mu\|^2_{H^1} \le a\big(y-y^\mu, y-y^\mu\big) = \big\langle B\big(u-u^\mu\big), y-y^\mu\big\rangle \\ \le C\|u-u^\mu\|_U\|y^\mu-y\|_{H^1}$$

we immediately infer

$$\|y^\mu-y\|_{H^1} + \|u^\mu-u\|_U \le C\|u^\mu-u\|_U. \tag{2.7}$$

## 3 Error analysis for the barrier problem

We will perform a finite element error analysis in this and in the next section. This is done in two steps. In the first step, which works under very general conditions, we show that approximation errors for our optimal control problem can be estimated in terms of $L_\infty$ errors for the Galerkin approximation to the optimal state $y$.

In the second step we discuss various settings, in which $L_\infty$-error estimates are available, and derive corresponding error estimates in terms of powers of the mesh-size $h$.

Denote for a "mesh-parameter" $h$ by $V_h \subset W^{1,\infty}(\Omega)$ a sequence of finite dimensional subspaces of $V_q$.

In what follows it is convenient to introduce a discrete approximation of the operator $\mathcal{G}$. For a given right hand side $f \in V_q'$ we denote by $z_h = \mathcal{G}_h(f) \in V_h$ the solution of the discrete problem

$$a(z_h, v_h) = \langle f, v_h\rangle \quad \text{for all } v_h \in V_h.$$

**Assumption 3.1** Assume that for each $u \in U$

$$\lim_{h\to 0}\|\mathcal{G}_h(Bu) - \mathcal{G}(Bu)\|_{L_\infty} = 0. \tag{3.1}$$

### 3.1 Exact quadrature

Problem (1.9) is now approximated by the following sequence of control problems depending on the mesh parameter $h$:

$$\min_{u\in U} J_h(u) := \frac{1}{2}\int_\Omega \big|\mathcal{G}_h(Bu) - y_0\big|^2 + \frac{\alpha}{2}\|u\|_U^2 + \int_\Omega l\big(\overline{y} - \mathcal{G}_h(Bu); \mu\big). \tag{3.2}$$

Problem (3.2) represents a convex infinite-dimensional optimization problem of a similar structure as problem (1.9). It admits a unique solution $u_h^\mu \in U$ with corresponding state $y_h^\mu \in V_h$. Furthermore, in accordance with problem (1.9), there exist a unique function $p_h^\mu \in V_h$ and a regular, non–negative Borel measure $\lambda_h^\mu$ satisfying

$$a\big(v_h, p_h^\mu\big) = \int_\Omega \big(y_h^\mu - y_0\big)v_h + \int_{\overline{\Omega}} v_h\, d\lambda_h^\mu \quad \text{for all } v_h \in V_h, \tag{3.3}$$

and

$$\alpha u_h^\mu + B^* p_h^\mu = 0 \quad \text{in } U. \tag{3.4}$$

We note that the control is not discretized in (3.2), compare [8, 9, 16, 26] for a more detailed discussion of this discretization approach.

Next we prove an error estimate in $h$ which is independent of $\mu$. For this purpose we first prove uniform bounds w.r.t. $\mu$ for $\|\lambda_h^\mu\|_{M(\overline{\Omega})}$.

**Lemma 3.2** *Let Assumption 1.1, 1.4, and 3.1 be satisfied. Then there is $h_0 > 0$, such that for all $h \le h_0$*

$$\|\lambda_h^\mu\|_{M(\overline{\Omega})} \le C \tag{3.5}$$

*with some positive constant $C$ independent of $\mu$ and of $h$.*

*Proof* Using (3.1) we obtain for some small enough $0 < h_0$

$$\mathcal{G}_h(B\hat{u}) \le \overline{y} - \frac{\tau}{2} \quad \text{for all } 0 < h \le h_0.$$

Therefore,

$$\begin{aligned}
\frac{\tau}{2}\int_{\overline{\Omega}} d\lambda_h^\mu &\le \int_{\overline{\Omega}} \overline{y} - \mathcal{G}_h(B\hat{u})\, d\lambda_h^\mu = \int_\Omega \big(y_h^\mu - y_0\big)\mathcal{G}_h(B\hat{u})\, dx - a\big(\mathcal{G}_h(B\hat{u}), p_h^\mu\big) \\
&= \int_\Omega \big(y_h^\mu - y_0\big)\mathcal{G}_h(B\hat{u})\, dx - \langle \hat{u}, B^* p_h\rangle_U \\
&= \int_\Omega \big(y_h^\mu - y_0\big)\mathcal{G}_h(B\hat{u})\, dx + \alpha\big\langle \hat{u}, u_h^\mu\big\rangle_U \le C.
\end{aligned}$$

□

Let $y^h \in V_h$ denote the discrete functions defined by the Galerkin orthogonality relation

$$a\big(y^\mu - y^h, v_h\big) = 0 \quad \forall v_h \in V_h. \tag{3.6}$$

Further, define by $\tilde{y} \in V_q$ the function $\tilde{y} := \mathcal{G}(Bu_h^\mu)$, i.e. the continuous solution to the discrete optimal control. It holds

$$a\big(\tilde{y} - y_h^\mu, v_h\big) = 0 \quad \forall v_h \in V_h. \tag{3.7}$$

Finally, let $\tilde{p}$ defined as the continuous solution to the discrete right hand side of the adjoint equation:

$$\tilde{p} := \mathcal{G}^*\big(y_h^\mu - y_0 + \lambda_h^\mu\big), \tag{3.8}$$

which satisfies the Galerkin orthogonality relation

$$a\big(v_h, \tilde{p} - p_h^\mu\big) = 0 \quad \forall v_h \in V_h. \tag{3.9}$$

The crucial point is now that a-priori error estimates are available for $y^h - y^\mu$ and $y_h^\mu - \tilde{y}$, taking into account that by Lemma 3.2 $\tilde{p}$ is uniformly bounded in $W^{1,q'}(\Omega)$, as $h \to 0$. We will discuss these estimates for particular settings in Sect. 4, below. We are now prepared to prove our main result:

**Theorem 3.3** *Let Assumption 1.1, 1.4, and 3.1 be satisfied. Let $u^\mu$ denote the solution of (1.9) with $y^\mu = \mathcal{G}(Bu^\mu)$, and $u_h^\mu$ the solution to (3.2) with $y_h^\mu = \mathcal{G}_h(Bu_h^\mu)$. Then there exists $h_0 > 0$ and a constant $C$ independent of $\mu$ and $h$ such that for all $h \le h_0$ the estimates*

$$\|u^\mu - u_h^\mu\|_U + \|y^\mu - y_h^\mu\|_{L_2} \le C\sqrt{\|y^\mu - y^h\|_\infty + \|\tilde{y} - y_h^\mu\|_\infty}, \tag{3.10}$$

$$\|y^\mu - y_h^\mu\|_{H^1} \le C\|u^\mu - u_h^\mu\|_U + \|y^\mu - y^h\|_{H^1}, \tag{3.11}$$

$$\|y^\mu - y_h^\mu\|_\infty \le C\|u^\mu - u_h^\mu\|_U + \|y^\mu - y^h\|_{L_\infty} \tag{3.12}$$

*hold.*

*Proof* We test the difference of (2.3) and (3.4) with $u^\mu - u_h^\mu$ and the difference of (2.2) and (3.3) with $y^\mu - y_h^\mu$ and add both. This yields, using the definitions of $\tilde{p}$ and $\tilde{y}$,

$$\begin{aligned}
&\alpha\|u^\mu - u_h^\mu\|_U^2 + \|y^\mu - y_h^\mu\|_{L_2}^2 + \int_{\overline{\Omega}} (y^\mu - y_h^\mu)\, d(\lambda^\mu - \lambda_h^\mu)\\
&\quad = a(y^\mu - y_h^\mu, p^\mu - \tilde{p}) - \langle B^*(p^\mu - p_h^\mu), u^\mu - u_h^\mu\rangle_U\\
&\quad = a(y^\mu - y_h^\mu, p^\mu - \tilde{p}) - \langle B(u^\mu - u_h^\mu), p^\mu - p_h^\mu\rangle_{(W^{1,q'})', W^{1,q'}}\\
&\quad = a(y^\mu - y_h^\mu, p^\mu - \tilde{p}) - a(y^\mu - \tilde{y}, p^\mu - p_h^\mu).
\end{aligned}$$

The right hand side can now be treated, using Galerkin orthogonality relations (3.9) (twice), (3.7), and (3.6):

$$\begin{aligned}
&a(y^\mu - y_h^\mu, p^\mu - \tilde{p}) - a(y^\mu - \tilde{y}, p^\mu - p_h^\mu)\\
&\quad = a(\tilde{y} - y_h^\mu, p^\mu) + a(y^\mu, p_h^\mu - \tilde{p}) + a(y_h^\mu, \tilde{p}) - a(\tilde{y}, p_h^\mu)\\
&\quad = a(\tilde{y} - y_h^\mu, p^\mu) + a(y^\mu - y^h, p_h^\mu - \tilde{p}) + a(y_h^\mu, p_h^\mu) - a(y_h^\mu, p_h^\mu)\\
&\quad = a(\tilde{y} - y_h^\mu, p^\mu) - a(y^\mu - y^h, \tilde{p}).
\end{aligned}$$

Now we reinsert the definition of $p^\mu$ and $\tilde{p}$ via (2.2) and (3.3) to obtain

$$\begin{aligned}
&a(\tilde{y} - y_h^\mu, p^\mu) - a(y^\mu - y^h, \tilde{p})\\
&\quad = \int_\Omega (y - y_0)(\tilde{y} - y_h^\mu)\, dx + \int_{\overline{\Omega}} (\tilde{y} - y_h^\mu)\, d\lambda^\mu - \int_\Omega (y_h - y_0)(y^\mu - y^h)\, dx\\
&\qquad - \int_{\overline{\Omega}} (y^\mu - y^h)\, d\lambda_h^\mu\\
&\quad \le \|\tilde{y} - y_h^\mu\|_\infty \|y - y_0 + \lambda^\mu\|_{M(\overline{\Omega})} + \|y^\mu - y^h\|_\infty \|y_h - y_0 + \lambda_h^\mu\|_{M(\overline{\Omega})}.
\end{aligned}$$

Using the uniform boundedness of $\lambda_h^\mu$ in $M(\overline{\Omega})$ and $y_h - y_0$ in $L_2(\Omega)$ we finally obtain:

$$\alpha\|u^\mu - u_h^\mu\|_U^2 + \|y^\mu - y_h^\mu\|_{L_2}^2 + \int_{\overline{\Omega}} (y^\mu - y_h^\mu)\, d(\lambda^\mu - \lambda_h^\mu) \\ \leq C\left(\|\tilde{y} - y_h^\mu\|_\infty + \|y^\mu - y^h\|_\infty\right). \tag{3.13}$$

Now observe that by monotonicity of the subdifferential (cf. (2.1))

$$\int_{\overline{\Omega}} (y^\mu - y_h^\mu)\, d(\lambda^\mu - \lambda_h^\mu) \geq 0, \tag{3.14}$$

which yields (3.10). Using this estimate and

$$\|y^\mu - y_h^\mu\|_{H^1} \leq \|y^\mu - y^h\|_{H^1} + \|y^h - y_h^\mu\|_{H^1} \leq \|y^\mu - y^h\|_{H^1} + C\|u^\mu - u_h^\mu\|_U$$

we finally get (3.11), and similarly (3.12). □

We observe in the proof that the same estimate holds for the quantity in (3.14).

## 3.2 The use of quadrature rules

So far we have assumed that the integrals $\int l'(y;\mu) v_h\, dx$ are evaluated exactly. Since $l'$ is a rational function this may be an involved computation. Now let us replace this term in (3.3) by a quadrature rule of the form

$$I_h[f] := \sum_i \omega_i^h f(x_i^h), \quad x_i^h \in \overline{\Omega},\ \omega_i^h \in \mathbb{R},$$

where $x_i^h$ are suitably chosen quadrature points with corresponding weights $\omega_i^h$.

The discrete optimality system (3.3)–(3.4) is then replaced by

$$a(v_h, p_h^\mu) = \int_\Omega (y_h^\mu - y_0) v_h + I_h[l'(y;\mu) v_h] \quad \text{for all } v_h \in V_h, \tag{3.15}$$

and

$$\alpha u_h^\mu + B^* p_h^\mu = 0 \quad \text{in } U. \tag{3.16}$$

Let $1_{\overline{\Omega}}$ be the function defined by $1_{\overline{\Omega}}(x) = 1\ \forall x \in \overline{\Omega}$. We will now study discretization errors for this case under the following assumptions:

- the quadrature rule yields positive values for positive functions;
- $I_h[1_{\overline{\Omega}}]$ is bounded, independently of $h$;
- the solution $y_h^\mu$ of (3.15)–(3.16) is feasible.

Note that assumptions on the error introduced by the quadrature rule are not needed.

We may as well interpret this quadrature rule as follows: we replace the Lebesgue measure on $\Omega$ by a bounded and positive measure with a support that consists of finitely many discrete points only. In particular, the space $L_1(I_h)$, by which we denote equivalence classes of all integrable functions with respect to this measure, is well defined.

It is easy to see that $y_h^\mu$ is strictly feasible at $x_i^h$. Otherwise the discrete barrier functional would be $+\infty$. Our assumptions are valid a-priori for linear finite elements, if we assume that $\overline{y} \in V_h$, the quadrature points are taken at the nodes of the discretization, and the weights are chosen appropriately (from the trapezoidal rule, say).

We apply the quadrature rule only for the evaluation of the barrier term, while all other integrals are assumed to be evaluated exactly.

Just as in the preceding section we prove the following lemma:

**Lemma 3.4** *Let Assumption 1.1, 1.4, and 3.1 be satisfied, let* $(u_h^\mu, y_h^\mu, p_h^\mu)$ *denote the unique solution to* (3.15)–(3.16). *Then there exists* $h_0 > 0$ *and a constant* $C$ *independent of* $\mu$ *and* $h$ *such that*

$$\big\|l'\big(y_h^\mu;\mu\big)\big\|_{L_1(I_h)} := I_h\big[\big|l'\big(y_h^\mu;\mu\big)\big|\big] \le C \quad \forall h \le h_0. \tag{3.17}$$

*Proof* Using (3.1) we obtain for some small enough $0 < h_0$

$$\mathcal{G}_h(B\hat{u}) \le \overline{y} - \frac{\tau}{2} \quad \text{for all } 0 < h \le h_0.$$

Therefore,

$$\begin{aligned}
&\frac{\tau}{2}\int I_h\big[\big|l'\big(y_h^\mu;\mu\big)\big|\big]\\
&\quad\le I_h\big[\big(\overline{y} - \mathcal{G}_h(B\hat{u})\big)l'\big(y_h^\mu;\mu\big)\big]\\
&\quad= \int_\Omega \big(y_h^\mu - y_0\big)\mathcal{G}_h(B\hat{u})\,dx - a\big(\mathcal{G}_h(B\hat{u}), p_h^\mu\big)\\
&\quad= \int_\Omega \big(y_h^\mu - y_0\big)\mathcal{G}_h(B\hat{u})\,dx - \big\langle\hat{u}, B^* p_h\big\rangle_U\\
&\quad= \int_\Omega \big(y_h^\mu - y_0\big)\mathcal{G}_h(B\hat{u})\,dx + \alpha\big\langle\hat{u}, u_h^\mu\big\rangle_U \le C.
\end{aligned}$$

□

**Lemma 3.5** *Consider either* $L_1(\Omega)$ *or* $L_1(I_h)$ *as defined above. Then for each* $0 < s \le 1$ *there is a constant* $c$, *independent of* $h$, *such that for all* $0 \le f \in L_1$

$$\big\|f^s\big\|_{L_1} \le c\|f\|_{L_1}^s.$$

*Proof* Clearly, $f^s$ is a measurable function, and thus it remains to show boundedness. Since $1/s > 1$ this follows from the computation

$$\big\|f^s\big\|_{L_1} \le c\big(|\Omega|, s\big)\big\|f^s\big\|_{L_{1/s}} = c\big(|\Omega|, s\big)\|f\|_{L_1}^s.$$

Here $|\Omega|$ denotes either the Lebesgue measure of $\Omega$, or $I_h[1_{\overline{\Omega}}]$, i.e., the discrete measure of $\Omega$. □

**Theorem 3.6** *Let Assumption 1.1, 1.4, and 3.1 be satisfied. Let $u^\mu$ denote the solution of (1.9) with $y^\mu = \mathcal{G}(Bu^\mu)$, and $(u_h^\mu, y_h^\mu, p_h^\mu)$ denote the unique solution to (3.15)–(3.16). Then there exists $h_0 > 0$ and a constant $C$ independent of $\mu$ and $h$ such that for all $h \le h_0$*

$$\|u^\mu - u_h^\mu\|_U + \|y^\mu - y_h^\mu\|_{L_2} \le C\sqrt{\|y^\mu - y^h\|_{L_\infty} + \|\tilde{y} - y_h^\mu\|_{L_\infty} + \mu}, \quad (3.18)$$

$$\|y^\mu - y_h^\mu\|_{H^1} \le C\|u^\mu - u_h^\mu\|_U + \|y^\mu - y^h\|_{H^1}, \quad (3.19)$$

$$\|y^\mu - y_h^\mu\|_\infty \le C\|u^\mu - u_h^\mu\|_U + \|y^\mu - y^h\|_{L_\infty}. \quad (3.20)$$

*Proof* The proof runs along the lines of the proof of Theorem 3.3 with the exception that every occurrence of $\int_{\overline{\Omega}} v\, d\lambda_h^\mu$ has to be replaced by its discrete counterpart $I_h[l'(y_h^\mu;\mu)v]$. Hence, instead of (3.13) we arrive at

$$\begin{aligned}
&\alpha\|u^\mu - u_h^\mu\|_U^2 + \|y^\mu - y_h^\mu\|_{L_2}^2 + I_h\big[l'\big(y_h^\mu;\mu\big)\big(y_h^\mu - y^\mu\big)\big] - \int_{\overline{\Omega}}\big(y_h^\mu - y^\mu\big)\,d\lambda^\mu \\
&\le \|\tilde{y} - y_h^\mu\|_\infty\big(\|y_h - y_0\|_{L_1(\Omega)} + \|l'\big(y_h^\mu;\mu\big)\|_{L_1(I_h)}\big) \\
&\quad + \|y^\mu - y^h\|_\infty\|y - y_0 + \lambda^\mu\|_{M(\overline{\Omega})} \\
&\le C\big(\|\tilde{y} - y_h^\mu\|_\infty + \|y^\mu - y^h\|_\infty\big). \qquad (3.21)
\end{aligned}$$

For the last inequality we used Lemmas 3.4 and 3.2. At this point it is not possible to apply a monotonicity argument as before. Instead, we introduce the notation $\hat{y} := y - \overline{y} \le 0$ and compute

$$\begin{aligned}
&I_h\big[l'\big(y_h^\mu;\mu\big)\big(y_h^\mu - y^\mu\big)\big] - \int_{\overline{\Omega}}\big(y_h^\mu - y^\mu\big)\,d\lambda^\mu \\
&\quad = I_h\big[l'\big(y_h^\mu;\mu\big)\big(\hat{y}_h^\mu\big)\big] + \int_{\overline{\Omega}}\hat{y}^\mu\,d\lambda^\mu - \underbrace{\Big(I_h\big[l'\big(y_h^\mu;\mu\big)\hat{y}^\mu\big] + \int_{\overline{\Omega}}\hat{y}_h^\mu\,d\lambda^\mu\Big)}_{\le 0} \\
&\quad \ge I_h\big[l'\big(y_h^\mu;\mu\big)\hat{y}_h^\mu\big] + \int_{\overline{\Omega}}\hat{y}^\mu\,d\lambda^\mu.
\end{aligned}$$

To estimate the remaining terms on the right hand side, we first note that by (2.4),

$$\int_{\overline{\Omega}}\hat{y}^\mu\,d\lambda^\mu = \int_\Omega \hat{y}^\mu l'\big(y^\mu;\mu\big)\,dx,$$

which means that possibly measure valued parts disappear. Hence by non-negativity of $l'$,

$$I_h\big[l'\big(y_h^\mu;\mu\big)\hat{y}_h^\mu\big] + \int_{\overline{\Omega}}\hat{y}^\mu\,d\lambda^\mu = -\|\hat{y}_\mu^h l'\big(y_h^\mu;\mu\big)\|_{L_1(I_h)} - \|\hat{y}^\mu l'\big(y^\mu;\mu\big)\|_{L_1(\Omega)}.$$

Next, recall that $l'(y;\mu) = \mu^k \hat{y}^{-k}$ for $k = 1, 2, 3, \ldots$. Hence, $l'(y;\mu)\hat{y} = \mu(l'(y;\mu))^{(k-1)/k}$, (in particular $l'(y;\mu)\hat{y} = \mu$ for $k = 1$), and thus

$$\big\|l'(y_h^\mu;\mu)\hat{y}_h^\mu\big\|_{L_1(I_h)} = \mu\big\|l'(y_h^\mu;\mu)^{\frac{k-1}{k}}\big\|_{L_1(I_h)} \quad \text{and}$$
$$\big\|l'(y^\mu;\mu)\hat{y}^\mu\big\|_{L_1(\Omega)} = \mu\big\|l'(y^\mu;\mu)^{\frac{k-1}{k}}\big\|_{L_1(\Omega)}.$$

Now Lemma 3.4, Lemmas 3.2 and 3.5 yield the uniform bounds

$$\big\|l'(y_h^\mu;\mu)^{\frac{k-1}{k}}\big\|_{L_1(I_h)} \le C \quad \text{and} \quad \big\|l'(y^\mu;\mu)^{\frac{k-1}{k}}\big\|_{L_1(\Omega)} \le C,$$

and thus

$$\big\|\hat{y}_h^\mu l'(y_h^\mu;\mu)\big\|_{L_1(I_h)} + \big\|\hat{y}^\mu l'(y^\mu;\mu)\big\|_{L_1(\Omega)} \le C\mu.$$

Inserting this estimate into (3.21) yields finally

$$\alpha\big\|u^\mu - u_h^\mu\big\|_U^2 + \big\|y^\mu - y_h^\mu\big\|_{L_2}^2 \le C\big(\big\|\tilde{y} - y_h^\mu\big\|_\infty + \big\|y^\mu - y^h\big\|_\infty + \mu\big),$$

which implies (3.18). The estimates (3.19) and (3.20) now follow just as in Theorem 3.3 □

Hence, using this type of quadrature rules introduces an additional error, which is only of the order of the remaining length of the central path. It seems thus acceptable to use inexact quadrature in a numerical interior point algorithm in order to save computational effort.

Using the trapezoidal rule as a quadrature rule and a fixed grid we may interpret the resulting numerical scheme as an interior point method for solving the discrete optimization problem obtained by variational discretization in the spirit of [9].

## 4 A-priori error estimates for specific settings

Let us now apply Theorem 3.3 to some concrete settings in optimal control. So we impose Assumption 1.1 and 1.4 in the whole section. We will study settings, where Assumption 3.1 holds with a certain rate. We only consider the case of exact quadrature. In view of Theorem 3.6 the modifications for quadrature rules are straightforward.

Let $\mathcal{T}_h$ be a triangulation of $\Omega$ with maximum mesh size $h := \max_{T\in\mathcal{T}_h} \operatorname{diam}(T)$ and vertices $x_1, \ldots, x_m$. We suppose that $\overline{\Omega}$ is the union of the elements of $\mathcal{T}_h$ so that element edges lying on the boundary are curved in case that $\Omega$ has curved boundary. In addition, we assume that the triangulation is quasi-uniform in the sense that there exists a constant $\kappa > 0$ (independent of $h$) such that each $T \in \mathcal{T}_h$ is contained in a ball of radius $\kappa^{-1}h$ and contains a ball of radius $\kappa h$. Let us define the space of linear finite elements,

$$V_h := \big\{v_h \in C^0(\overline{\Omega}) \mid v_h \text{ is a linear polynomial on each } T \in \mathcal{T}_h\big\}$$

with the appropriate modification for boundary elements.

In the previous section we have derived a-priori error estimates for state constrained problems that depend on a-priori estimates with respect to the $L_\infty$-norm for $y^\mu$. There is plenty of literature on $L_\infty$-error estimates for finite element approximations. The quality of these estimates depends on the regularity of $y^\mu$. The regularity of $y^\mu$ in turn depends on the regularity of $Bu^\mu$, for which we have the following results available:

*Example 4.1* By (2.6), $p^\mu$ is uniformly bounded in $W^{1,q'}(\Omega)$. In view of Example 1.3 we thus have the following regularity results

(i) For distributed control we have $Bu^\mu = u^\mu \in W^{1,q'}(\Omega)$ by (2.3), for all $q' < d/(d-1)$.
(ii) In the case of Neumann boundary control, by definition of $B$, $u^\mu = \alpha^{-1} B^* p^\mu$ is the boundary trace of $p^\mu$, which is contained in $W^{1-1/q',q'}(\partial\Omega)$.
(iii) For finite dimensional control, the regularity of $Bu^\mu$ depends, of course, on the input fields $f_i$.

We start with the case, where $a(\cdot,\cdot)$ is $H^2$-regular. This holds for example if $\Omega$ is a convex domain, and $a_{ij}$ is Lipschitz continuous. Then it is well known (cf. e.g. [4]) that for $f \in L^2(\Omega)$

$$\left\|\mathcal{G}(f) - \mathcal{G}_h(f)\right\|_\infty \le Ch^{2-d/2}\|f\|_{L_2}. \tag{4.1}$$

Thus we conclude from Theorem 3.3

**Corollary 4.2** *If $B : U \to L_2(\Omega)$ is bounded, then for an $H^2$-regular problem we have the estimate*

$$\left\|u^\mu - u_h^\mu\right\|_U + \left\|y^\mu - y_h^\mu\right\|_{H^1} \le Ch^{1-\frac{d}{4}} \quad \textit{for all } 0 < h \le h_0.$$

*Proof* Since $u^\mu$ and $u_h^\mu$ are uniformly bounded in $U$ and $B$ is bounded, (4.1) can be applied to provide for the right hand side of (3.10) the estimate $\|y^\mu - y^h\|_\infty + \|\tilde{y} - y_h^\mu\|_\infty \le Ch^{2-d/2}$. Theorem 3.3 then yields (4.2). □

If we have more smoothness, we obtain better approximation results. From now on, let us—for the sake of simplicity—consider problems on smoothly bounded domains with smooth coefficients. Finite element approximations would then have to be accomplished by curved elements.

If $f \in L^\infty(\Omega)$ and $V = H_0^1(\Omega)$, then, for example, [23] yields for $d = 2, 3$

$$\left\|\mathcal{G}(f) - \mathcal{G}_h(f)\right\|_\infty \le Ch^2|\log h|^2\|f\|_\infty,$$

so that we obtain

**Corollary 4.3** *If $Bu^\mu, Bu_h^\mu \in L_\infty(\Omega)$ are bounded uniformly in $L_\infty(\Omega)$ for $\mu \to 0$ and $h \to 0$, then*

$$\left\|u^\mu - u_h^\mu\right\|_U + \left\|y^\mu - y_h^\mu\right\|_{H^1} \le Ch|\log h| \quad \textit{for all } 0 < h \le h_0.$$

A uniform bound for $Bu^\mu, Bu_h^\mu \in L_\infty(\Omega)$ can be observed a-posteriori in many, but not all optimal control problems with distributed control, and it holds in the case of finite dimensional control, if the input fields $f_i$ are all in $L_\infty(\Omega)$.

For distributed control $Bu^\mu \in W^{1,q'}(\Omega)$ for all $q' < d/(d-1)$ (cf. Example 4.1(i)), because $p^\mu \in W^{1,q'}(\Omega)$. Analogously, Lemma 3.2 implies that all $\tilde{p}$, as defined in (3.8) are uniformly bounded in $W^{1,q'}(\Omega)$ for $\mu, h \to 0$. Using a stability result for the Galerkin projection in $W^{1,q'}(\Omega)$ (cf. e.g. [3, Theorem 8.5.3]), its Galerkin projection $p_h^\mu$ is also uniformly bounded in $W^{1,q'}(\Omega)$. Since $B = Id$, the same holds for $u_h^\mu$, and thus also for $Bu_h^\mu$.

In this setting we have the following estimates:

**Corollary 4.4** *If $Bu^\mu, Bu_h^\mu \in W^{1,q'}(\Omega)$ for all $q' < d/(d-1)$ are uniformly bounded for $\mu \to 0$ and $h \to 0$, then under the above regularity assumptions we have for all $\varepsilon > 0$:*

$$\|u^\mu - u_h^\mu\|_U + \|y^\mu - y_h^\mu\|_{H^1} \le Ch^{2-d/2-\varepsilon} \quad \text{for all } 0 < h \le h_0. \tag{4.2}$$

*Proof* Under these assumptions, the estimate

$$\|\mathcal{G}(f) - \mathcal{G}_h(f)\|_\infty \le Ch^{3-d/q'}\|f\|_{W^{1,q'}(\Omega)}$$

can be derived, see [18, Chap. 3] and [24] for details. Since $q'$ can be chosen arbitrarily close to $d/(d-1)$, Theorem 3.3 yields (4.2). □

Let us consider briefly the case of Neumann boundary control in two dimensions. According to Example 4.1 we can expect $Bu^\mu \in W^{1-1/q',q'}(\partial\Omega)$ for all $q' < 2$, and a uniform bound for $Bu_h^\mu$, similarly as above.

**Corollary 4.5** *If $Bu^\mu, Bu_h^\mu \in W^{1-1/q',q'}(\partial\Omega)$ for all $q' < 2$ are uniformly bounded for $\mu \to 0$ and $h \to 0$, then under the above regularity assumptions we have for all $\varepsilon > 0$*

$$\|u^\mu - u_h^\mu\|_U + \|y^\mu - y_h^\mu\|_{H^1} \le Ch^{1/2-\varepsilon} \quad \text{for all } 0 < h \le h_0.$$

*Proof* If $Bu^\mu \in W^{1-1/q',q'}(\partial\Omega)$, a general regularity result [1, Theorem 9.2] yields for the regularity of the state $y$:

$$\|y^\mu\|_{W^{2,q'}(\Omega)} \le c\|Bu^\mu\|_{W^{1-1/q',q'}(\partial\Omega)},$$

and thus uniform boundedness of $y$ in $\|y\|_{W^{2,q'}}$. Now [3, Corollary 8.6.3] shows boundedness of the Galerkin projection in $W^{1,q}(\Omega)$, if $q$ is sufficiently close to 2. Hence, we obtain

$$\|y^\mu - y^h\|_{W^{1,q}} \le C \inf_{v_h \in V_h} \|y^\mu - v_h\|_{W^{1,q}} \le ch^{1+1/q-1/q'}\|y^\mu\|_{W^{2,q'}}.$$

As already noted, $q'$ and $q$ can be chosen arbitrarily close to 2, and thus we conclude

$$\|y^\mu - y^h\|_{L_\infty} \le c\|y^\mu - y^h\|_{W^{1,q}} \le Ch^{1-\varepsilon} \quad \forall \varepsilon = 1/q' - 1/q > 0.$$

Application of Theorem 3.3 yields the desired result. □

For boundary control in the case $d = 3$, it may happen that $y^\mu \notin L_\infty(\Omega)$, so we cannot expect in general that $L_\infty$-error estimates for $y^\mu$ hold.

## 5 Parameter adjustment

The analysis in the preceding two sections allows us to optimally select $\mu$ for given $h$, or vice versa. Considering the results of the previous sections, and neglecting the possible occurrence of logarithmic terms, we have shown that there exists a constant $C$ independent of $\mu$ such that

$$\|u - u_h^\mu\|_U \le \|u - u^\mu\|_U + \|u^\mu - u_h^\mu\|_U \le C\big(\mu^{1/2} + h^\gamma\big), \tag{5.1}$$

for some $\gamma$, depending on the regularity of the partial differential equation.

For given $\mu > 0$ the mesh size $h$ on the right hand side can be adjusted on the basis of (5.1) as

$$h(\mu) \sim \mu^{1/(2\gamma)} \quad \text{or vice versa} \quad \mu(h) \sim h^{2\gamma}.$$

The optimal error bound obtained by Corollary 4.2 then is

$$\|u - u_h^\mu\|_U \le Ch^{1-\frac{d}{4}}$$

if $\mu$ is chosen proportional to $h^{2-d/2}$. If we have the regularity of Corollary 4.4 the optimal adjustment is given by

$$h(\mu) \sim \begin{cases} \sqrt{\mu} & \text{if } d = 2, \\ \mu & \text{if } d = 3, \end{cases} \tag{5.2}$$

and it is $h(\mu) \sim \sqrt{\mu}$ independent of $d$ if the requirements of Corollary 4.3 are met. For Neumann boundary control, one will chose $h \sim \mu$, on the basis of Corollary 4.5. This provides a qualitative guideline when to stop the interior point method for a fixed discretization, or how to refine the discretization for a given $\mu$.

## 6 Numerical examples

Finally we illustrate our theoretical findings by a numerical example. We choose $\Omega = [0;1] \times [0;1]$, $A = -\Delta + I$, $B = Id$, $\overline{y} = 0.55$, $y_0 = 2 \cdot x_1 \cdot x_2$, $\alpha = 10^{-3}$. The discretization of the state $y$ and the adjoint state $p$ is performed by a linear finite element method, based on the DUNE library [2]. Computational solutions can be seen in Fig. 1.

For the evaluation of the barrier integrals we use the trapezoidal rule, as analyzed in Sect. 3.2. For the numerical solution we use an interior point Newton path-following method in the variables $y$ and $p$ similar to the one analyzed in [28]. The

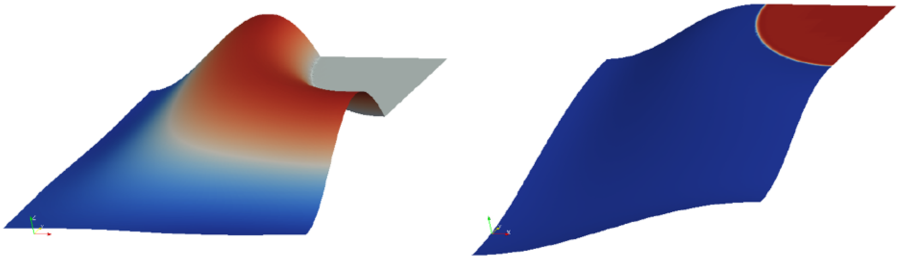

**Fig. 1** *Left*: optimal control. *Right*: optimal state with active set

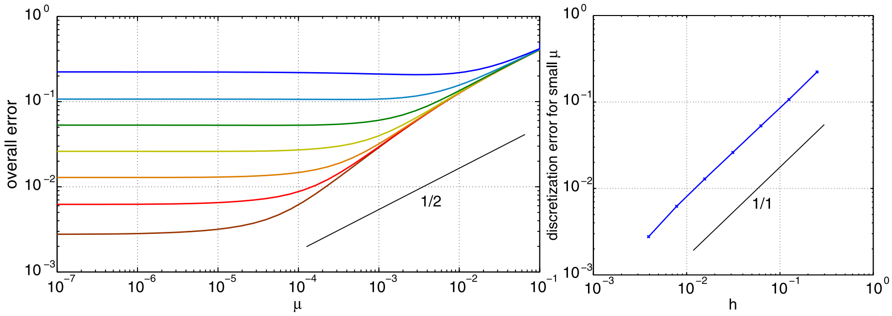


**Fig. 2** *Left*: Overall errors plotted against $\mu$ for $h = 2^{-k}$, $k = 2, \dots, 8$. *Right*: Discretization errors plotted against $h$ for small $\mu$

resulting linear systems of equations are solved by the direct sparse solver *PARDISO* [25].

We compute $y_h^\mu$ and $p_h^\mu$ and estimate their overall errors w.r.t. $y_*$, $p_*$ by comparing them with a discrete solution that is computed on a very fine grid for a very small $\mu$. The choices of $h$ were $h = 2^{-k}$ with $k = 2, \dots, 8$. The plots in Fig. 2 show $\|y_h^\mu - y_*\|_{H^1} + \sqrt{\alpha}\|u_h^\mu - u_*\|_{L_2}$.

As one can see in the left plot, the error introduced by the regularization dominates, until a break even point is reached. Even on the finest level of discretization, this is for relatively large $\mu \approx 10^{-4}$–$10^{-5}$. In this range would be the most efficient point to stop the algorithm. It is interesting to note that for this particular problem the convergence of the path is slightly faster than $O(\sqrt{\mu})$.

In the right plot we observe that the discretization error for a small $\mu$ behaves like $O(h)$ as predicted by our theory.

Finally, we illustrate the effect of coupled gridsize-parameter adjustment and compare its numerical performance with that obtained for fixed, small $\mu$. Since the discretization error is approximately halved if the grid is refined once, (5.2) suggests to perform one refinement of the grid after $\mu$ has been reduced approximately by a factor of 0.25. Figure 3 presents the numerical findings for the initial choices $h = 1/2$, $\mu = 0.1$ (lower diagonal graph), and $h = 1/2$, $\mu \approx 0.01$ (upper diagonal graph). Compared with the pure discretization error (vertical light grey graph), the achieved accuracy is similar for all approaches, for $h \le 2^{-5}$.

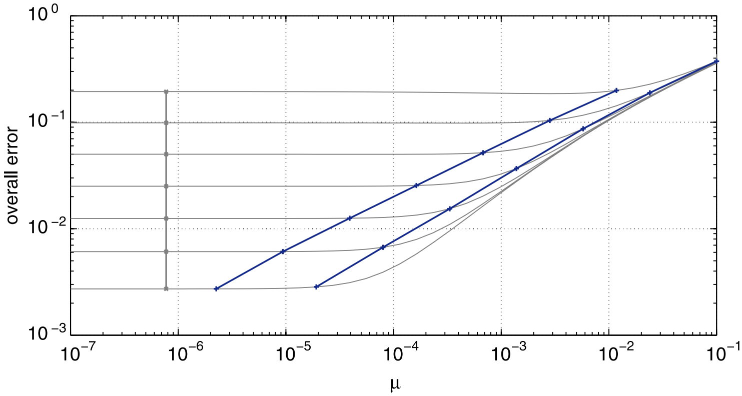


**Fig. 3** Comparison of overall errors for fixed, small $\mu$ (*vertical light grey graph*) and simultaneous grid refinement-$\mu$ reduction (*diagonal blue graphs*). The *marks* are set for $h = 2^{-k}$ with $k = 2, \ldots, 8$

In practical path-following methods an a-posteriori estimate for the remaining length of the central path as described, for example, in [26, Sect. 8.2] may be used, combined with an a-posteriori error estimate for the discretization error, to achieve a proper balancing of the parameters. In this respect we refer to the work [31], where parameter balancing is considered in the context of goal-oriented adaptivity, and to [29], where parameter adjustment and adaptivity are driven by the requirement of keeping Newton iterates in the region of fast convergence during path-following.

**Acknowledgements** The first author gratefully acknowledges the support of the DFG Priority Program 1253 entitled Optimization With Partial Differential Equations. The work of the second author was supported by the DFG Research Center MATHEON "Mathematics for key technologies".

## References


1. Amann, H.: Nonhomogeneous linear and quasilinear elliptic and parabolic boundary value problems. In: Schmeisser, H.J., Triebel, H. (eds.) Function Spaces, Differential Operators and Nonlinear Analysis, pp. 9–126. Teubner, Stuttgart (1993)
2. Bastian, P., Blatt, M., Engwer, C., Dedner, A., Klöfkorn, R., Kuttanikkad, S., Ohlberger, M., Sander, O.: The distributed and unified numerics environment (DUNE). In: Proc. of the 19th Symposium on Simulation Technique in Hanover, September 12–14, 2006
3. Brenner, S., Scott, R.: The Mathematical Theory of Finite Elements, 2nd edn. Springer, New York (2002)
4. Casas, E.: $L^2$ estimates for the finite element method for the Dirichlet problem with singular data. Numer. Math. **47**, 627–632 (1985)
5. Casas, E.: Control of an elliptic problem with pointwise state constraints. SIAM J. Control Optim. **24**, 1309–1322 (1986)
6. Casas, E.: Boundary control of semilinear elliptic equations with pointwise state constraints. SIAM J. Control Optim. **31**, 993–1006 (1993)
7. Deckelnick, K., Günther, A., Hinze, M.: A priori and a posteriori error control for elliptic control problems with pointwise constraints. Oberwolfach Report 13/2008, volume 5, issue 1 (2008)
8. Deckelnick, K., Hinze, M.: Convergence of a finite element approximation to a state constrained elliptic control problem. SIAM J. Numer. Anal. **45**, 1937–1953 (2007)
9. Deckelnick, K., Hinze, M.: A finite element approximation to elliptic control problems in the presence of control and state constraints. Preprint HBAM2007-01, Hamburger Beiträge zur Angewandten Mathematik, Universität Hamburg (2007)

10. Ekeland, I., Témam, R.: Convex Analysis and Variational Problems. Classics in Applied Mathematics, vol. 28. SIAM, Philadelphia (1999)
11. Gröger, K.: A $W^{1,p}$-estimate for solutions to mixed boundary value problems for second order elliptic differential equations. Math. Ann. **283**(4), 679–687 (1989)
12. Haller-Dintelmann, R., Kaiser, H.-C., Rehberg, J.: Elliptic model problems including mixed boundary conditions and material heterogeneities. J. Math. Pures Appl. **89**, 25–48 (2008)
13. Hintermüller, M., Hinze, M.: A note on an optimal parameter adjustment in a Moreau-Yosida-based approach to state constrained elliptic control problems. SIAM J. Numer. Anal. **47**, 1666–1683 (2009)
14. Hintermüller, M., Kunisch, K.: Path following methods for a class of constrained minimization methods in function spaces. SIAM J. Optim. **17**, 159–187 (2006)
15. Hintermüller, M., Kunisch, K.: Feasible and non-interior path following in constrained minimization with low multiplier regularity. SIAM J. Control Optim. **45**, 1198–1221 (2006)
16. Hinze, M.: A variational discretization concept in control constrained optimization: the linear-quadratic case. J. Comput. Optim. Appl. **30**, 45–63 (2005)
17. Hinze, M., Meyer, C.: Variational discretization of Lavrentiev-regularized state constrained elliptic optimal control problems. Comput. Optim. Appl. (2008). doi:10.1007/s10589-008-9198-1
18. Hinze, M., Pinnau, R., Ulbrich, M., Ulbrich, S.: Optimization with PDE Constraints. Mathematical Modelling: Theory and Applications, vol. 23. Springer, Berlin (2009)
19. Jerison, D., Kenig, C.: The inhomogeneous Dirichlet problem in Lipschitz domains. J. Funct. Anal. **130**, 161–219 (1995)
20. Meyer, C.: Error estimates for the finite-element approximation of an elliptic control problem with pointwise state and control constraints. Control Cybern. **37**, 51–85 (2008)
21. Meyer, C., Prüfert, U., Tröltzsch, F.: On two numerical methods for state-constrained elliptic control problems. Optim. Methods Softw. **22**, 871–899 (2007)
22. Meyer, C., Rösch, A., Tröltzsch, F.: Optimal control of PDEs with regularized pointwise state constraints. Comput. Optim. Appl. **33**, 209–228 (2006)
23. Rannacher, R.: Zur $L^\infty$-Konvergenz linearer finiter Elemente beim Dirichlet-Problem. Math. Z. **149**(1), 69–77 (1975)
24. Schatz, A.H.: Pointwise error estimates and asymptotic error expansion inequalities for the finite element method on irregular grids, I: global estimates. Math. Comput. **67**(223), 877–899 (1998)
25. Schenk, O., Gärtner, K.: On fast factorization pivoting methods for sparse symmetric indefinite systems. Electron. Trans. Numer. Anal. **23**, 158–179 (2006)
26. Schiela, A.: The control reduced interior point method—a function space oriented algorithmic approach. PhD thesis, Free University of Berlin. Dept. Math. and Comp. Sci. (2006)
27. Schiela, A.: Barrier methods for optimal control problems with state constraints. ZIB Report, 07-07 (2007). SIAM J. Optim. (accepted for publication)
28. Schiela, A.: An interior point method in function space for the efficient solution of state constrained optimal control problems. ZIB Report, 07-44 (2008)
29. Schiela, A., Günther, A.: Interior point methods in function space for state constraints—inexact Newton and adaptivity. ZIB Report, 09-01 (2009)
30. Tröltzsch, F.: Optimale Steuerung mit partiellen Differentialgleichungen. Vieweg, Wiesbaden (2005)
31. Wollner, W.: A posteriori error estimates for a finite element discretization of interior point methods for an elliptic optimization problem with state constraints. Comput. Optim. Appl. (2008)