

Hamburger Beiträge

zur Angewandten Mathematik

**A-priori error bounds for finite element
approximation of elliptic optimal control problems
with gradient constraints**

Klaus Deckelnick and Michael Hinze

Revised Version, 06.01.2014

Nr. 2013-10
September 2013

A-priori error bounds for finite element approximation of elliptic optimal control problems with gradient constraints

Klaus Deckelnick and Michael Hinze

Abstract. The finite element approximation of an elliptic optimal control problem with pointwise bounds on the gradient of the state is considered. We review recent results on the error analysis for various discretization approaches and prove a new bound for the problem without control constraints.

Mathematics Subject Classification (2010). Primary 49J20; Secondary 65N30.

Keywords. Elliptic optimal control problem, state constraints, error estimates.

1. Introduction

The subject of this report is the finite element approximation of optimal control problems with constraints on the gradient of the state. A typical example is the optimization of a cooling process in which the temperature acts as the state variable and large temperature gradients are prohibited in order to avoid possible damage in the material. We shall restrict ourselves to a model problem which involves the optimal control of a linear elliptic partial differential equation in the presence of pointwise bounds on the gradient of the state, while the control variable can be both constrained or unconstrained. In order to discretize this problem it is common to approximate the underlying objective functional by a sequence of functionals which are obtained by discretizing the state equation with the help of a finite element method. Natural choices in this step are continuous, piecewise linear finite elements but also the lowest order Raviart–Thomas mixed finite element. The control variable can be handled in two ways: either by variational discretization (see [11]), which means that the first order optimality conditions give rise to an implicit discretization in terms of the discrete adjoint state; another possibility consists in discretizing the control explicitly, typically by piecewise constant functions.

This report focusses on the a-priori error analysis for the abovementioned approaches. Apart from reviewing results that have been obtained in [5, 8, 12] we prove a new bound in the case in which the control variable is unconstrained and the objective

functional contains an L^r -norm ($r > 2$). In the remaining part of the paper we present a number of test calculations.

Let us close this section with a short survey of related publications. Elliptic optimal control problems with gradient constraints in nonsmooth polygonal domains are considered by Wollner in [16, 17]. While [16] is concerned with the existence of solutions, first order conditions and regularity, [17] derives a-priori error bounds for a finite element discretization. A general Moreau–Yosida framework for the treatment of elliptic optimal control problems with state and gradient constraints is presented by Hintermüller and Kunisch in [9]. Interior point approaches are investigated by Schiela and Wollner in [13]. In [15] Wollner presents an a-posteriori error analysis for an interior point approach to elliptic optimal control problems with general state constraints, including the case of pointwise bounds on the gradient of the state. A residual based adaptive approach to elliptic optimal control problems with pointwise gradient state constraints is presented by Hintermüller et al. in [10].

2. Mathematical setting

Let $\Omega \subset \mathbb{R}^d$ ($d = 2, 3$) be a bounded domain with a $C^{1,1}$ -boundary and consider the differential operator

$$\mathcal{A}y := - \sum_{i,j=1}^d \partial_{x_j}(a_{ij}y_{x_i}) + a_0y,$$

where for simplicity the coefficients a_{ij} and a_0 are assumed to be smooth functions on $\bar{\Omega}$. In what follows we assume that $a_{ij} = a_{ji}$, $a_0 \geq 0$ in Ω and that there exists $c_0 > 0$ such that

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

We consider the elliptic boundary value problem

$$\begin{aligned} \mathcal{A}y &= u & \text{in } \Omega \\ y &= 0 & \text{on } \partial\Omega. \end{aligned} \tag{2.1}$$

It is well-known that for every $1 < p < \infty$ (2.1) has a unique solution $y \in W^{2,p}(\Omega) \cap W_0^{1,p}(\Omega)$ with

$$\|y\|_{W^{2,p}} \leq C\|u\|_{L^p}. \tag{2.2}$$

Here $\|\cdot\|_{L^p}$ and $\|\cdot\|_{W^{k,p}}$ denote the usual Lebesgue and Sobolev norms. If $p = 2$ we simply write $\|\cdot\| = \|\cdot\|_{L^2}$.

We consider the following optimal control problem:

$$\min_{u \in K} J(u) = \frac{1}{2} \int_{\Omega} |y - y_0|^2 + \frac{\alpha}{r} \int_{\Omega} |u|^r \tag{2.3}$$

where y solves (2.1) and $\nabla y \in C$.

Here, $\alpha > 0$ and $y_0 \in L^2(\Omega)$ are given, while

$$C = \{\mathbf{z} \in C^0(\bar{\Omega})^d \mid |\mathbf{z}(x)| \leq \delta, x \in \bar{\Omega}\},$$

for some given $\delta > 0$ and $|\cdot|$ denotes the Euclidian norm in \mathbb{R}^d . Furthermore, we consider the following two possible choices for K and r :

(I) $K = \{u \in L^\infty(\Omega) \mid a \leq u \leq b \text{ a.e. in } \Omega\}$, $r = 2$, where $a < b$ are given constants.

(II) $K = L^r(\Omega)$ for some $r > d$.

Note that in both cases a well-known embedding result implies that $\nabla y \in C^0(\bar{\Omega})^d$ for the solution of (2.1), so that the condition $\nabla y \in C$ in (2.3) makes sense.

Existence of solutions, first order conditions as well as the structure and regularity of multipliers for control problems with pointwise constraints on the gradient of the state were investigated by Casas and Fernández in [4]. The authors allow a semi-linear state equation and rather general constraints on the control and the gradient of the state. The above choices (I) and (II) fit into the framework of [4]. In order to formulate the first order optimality conditions we introduce the space of regular Borel measures $\mathcal{M}(\bar{\Omega})$, which is the dual space of $C^0(\bar{\Omega})$. The norm on $\mathcal{M}(\bar{\Omega})$ is given by

$$\|\mu\|_{\mathcal{M}(\bar{\Omega})} = \sup_{f \in C^0(\bar{\Omega}), |f| \leq 1} \int_{\bar{\Omega}} f d\mu.$$

In case (I) we assume in addition that the following Slater condition holds:

$$\exists \hat{u} \in K \quad |\nabla \hat{y}(x)| < \delta, \quad x \in \bar{\Omega}, \quad \text{where } \hat{y} \text{ solves (2.1) with } u = \hat{u}. \quad (2.4)$$

Note that in case (II) one may simply choose $\hat{u} = 0$ to satisfy this condition.

Theorem 2.1. *An element $u \in K$ is a solution of (2.3) if and only if there exist $\mu \in \mathcal{M}(\bar{\Omega})^d$ and $p \in L^t(\Omega)$ ($t < \frac{d}{d-1}$) such that*

$$\int_{\Omega} p \mathcal{A}z - \int_{\Omega} (y - y_0)z - \int_{\bar{\Omega}} \nabla z \cdot d\mu = 0 \quad \forall z \in W^{2,t'}(\Omega) \cap W_0^{1,t'}(\Omega), \quad (2.5)$$

$$\int_{\Omega} (p + \alpha|u|^{r-2}u)(\tilde{u} - u) \geq 0 \quad \forall \tilde{u} \in K, \quad (2.6)$$

$$\int_{\bar{\Omega}} (\mathbf{z} - \nabla y) \cdot d\mu \leq 0 \quad \forall \mathbf{z} \in C. \quad (2.7)$$

Here, y is the solution of (2.1) and $\frac{1}{t} + \frac{1}{t'} = 1$.

Proof. see, [4, Theorem 3] and [4, Corollary 1]. □

Remark 2.2. We may infer from (2.6) that in case

$$(I) \quad u(x) = \text{Proj}_{[a,b]} \left(-\frac{p(x)}{\alpha} \right) \text{ a.a. } x \in \Omega,$$

$$(II) \quad u(x) = -\alpha^{-\frac{1}{r-1}} |p(x)|^{\frac{2-r}{r-1}} p(x) \text{ a.a. } x \in \Omega.$$

In the latter case it is shown in [12, Corollary 5] that this relation together with (2.5) implies that $u \in W^{\frac{1-d}{r-1}, r}(\Omega)$ for any $\epsilon > 0$. An embedding result (see [14, 4.6.1]) then yields $u \in L^{p_\epsilon}(\Omega)$, where $p_\epsilon = \frac{r-1}{1-\frac{1}{r}+\epsilon}$ for any $\epsilon > 0$.

3. Finite element discretization

Let \mathcal{T}_h be a triangulation of Ω with maximum mesh size $h := \max_{T \in \mathcal{T}_h} \text{diam}(T)$. We suppose that $\bar{\Omega}$ is the union of the elements of \mathcal{T}_h ; boundary elements are allowed to have one curved face. 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 κh .

3.1. Piecewise linear approximation of the state

Let us recall the definition of the space of linear finite elements,

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

and let $X_{h0} := X_h \cap H_0^1(\Omega)$. For a given function $u \in L^2(\Omega)$ we denote by $y_h \in X_{h0}$ the solution of

$$\int_{\Omega} A \nabla y_h \cdot \nabla v_h + \int_{\Omega} a_0 y_h v_h = \int_{\Omega} u v_h \quad \text{for all } v_h \in X_{h0}. \quad (3.1)$$

Here, we have abbreviated $A(x) = (a_{ij}(x))_{i,j=1}^d$. Let us define

$$C_h := \{\mathbf{c}_h : \bar{\Omega} \rightarrow \mathbb{R}^d \mid \mathbf{c}_h|_T \text{ is constant and } |\mathbf{c}_h|_T| \leq \delta, T \in \mathcal{T}_h\}. \quad (3.2)$$

We approximate (2.3) by the following control problem depending on the mesh parameter h :

$$\min_{u \in K} J_h(u) := \frac{1}{2} \int_{\Omega} |y_h - y_0|^2 + \frac{\alpha}{r} \int_{\Omega} |u|^r \quad (3.3)$$

subject to y_h solves (3.1) and $\nabla y_h \in C_h$.

Note that the control variable is not discretized. Problem (3.3) represents a convex infinite-dimensional optimization problem of similar structure as problem (2.3), but with only finitely many constraints on the state. The following first order conditions yield an implicit discretization of the control variable in terms of the discrete adjoint state. Using (2.4) it is not difficult to see that a Slater condition holds for (3.3) provided that $0 < h \leq h_0$.

Lemma 3.1. *Problem (3.3) has a unique solution $u_h \in K$. For $0 < h \leq h_0$ there are $\mu_{\mathbf{T}} \in \mathbb{R}^d$, $T \in \mathcal{T}_h$ and $p_h \in X_{h0}$ such that*

$$\int_{\Omega} (A \nabla v_h \cdot \nabla p_h + a_0 v_h p_h) - \int_{\Omega} (y_h - y_0) v_h - \sum_{T \in \mathcal{T}_h} \nabla v_h|_T \cdot \mu_{\mathbf{T}} = 0 \quad \forall v_h \in X_{h0}, \quad (3.4)$$

$$\int_{\Omega} (p_h + \alpha |u_h|^{r-2} u_h) (\tilde{u} - u_h) \geq 0 \quad \forall \tilde{u} \in K, \quad (3.5)$$

$$\sum_{T \in \mathcal{T}_h} (\mathbf{c}_h|_T - \nabla y_h|_T) \cdot \mu_{\mathbf{T}} \leq 0 \quad \forall \mathbf{c}_h \in C_h. \quad (3.6)$$

Here, $y_h \in X_{h0}$ is the solution of (3.1) with right hand side u_h .

Proof. see [4, Theorem 7] with the choices $U = L^r(\Omega)$, $K \subset U$, $C_h \subset Z := \mathbb{R}^{N_h}$, where N_h is the number of triangles in \mathcal{T}_h . \square

Remark 3.2. Similar to Remark 2.2 we deduce from (3.5) that for

$$\text{(I)} \quad u_h(x) = \text{Proj}_{[a,b]} \left(-\frac{p_h(x)}{\alpha} \right) \text{ a.a. } x \in \Omega,$$

$$\text{(II)} \quad u_h(x) = -\alpha^{-\frac{1}{r-1}} |p_h(x)|^{\frac{2-r}{r-1}} p_h(x) \text{ a.a. } x \in \Omega,$$

so that in both cases the discrete control is expressed implicitly in terms of the piecewise linear discrete costate p_h , the relation however being nonlinear.

For the unconstrained case **(II)**, the following error bound has been proved in [8, Theorem 2.5]:

Theorem 3.3. *Let u and u_h be the solutions of (2.3) and (3.3) in case **(II)** respectively. Then there exists $h_0 > 0$ such that*

$$\|u - u_h\|_{L^r} \leq Ch^{\frac{1}{r}(1-\frac{d}{r})}, \quad \|y - y_h\| \leq Ch^{\frac{1}{2}(1-\frac{d}{r})}$$

for all $0 < h \leq h_0$.

The proof relies on a careful combination of the information given by the primal and adjoint equations and we present the main ideas in the following section for a mixed finite element discretization of the state equation. The bounds in Theorem 3.3 are still satisfied if one employs a discretization of the control variable by piecewise constant functions on \mathcal{T}_h , see [8, Theorem 2.8]. We also remark that the above results are obtained by Ortner and Wollner in [12] without making use of adjoint information by working directly with the functionals J and J_h . Such a technique was previously used in [6] for the numerical analysis of elliptic optimal control problems with pointwise bounds on the state.

In general, both the control u and the adjoint variable p have low regularity even allowing jumps. For this reason, piecewise linear, continuous finite elements are not ideally suited for the discretization as they tend to develop oscillations near discontinuities. In the next section we present an alternative approach on the basis of a mixed finite element approach of lowest order for the state equation. This approach leads in particular to piecewise constant approximations for the state, costate and control and therefore seems to be better suited to handle discontinuities.

3.2. Mixed finite element approximation of the state

As already mentioned we now use a mixed formulation in order to approximate the solution of (2.1). Let us introduce

$$H(\text{div}, \Omega) := \{\mathbf{w} \in L^2(\Omega)^d \mid \text{div} \mathbf{w} \in L^2(\Omega)\}$$

and write $(y, \mathbf{v}) = \mathcal{G}(u)$, where $\mathbf{v} = A\nabla y$ and y is the solution of (2.1).

We use a mixed finite element method based on the lowest order Raviart–Thomas element. Let

$$\mathbf{V}_h := RT_0(\Omega, \mathcal{T}_h) := \{\mathbf{w}_h \in H(\text{div}, \Omega) \mid \mathbf{w}_h|_T \in RT_0(T) \text{ for all } T \in \mathcal{T}_h\},$$

where $RT_0(T) = \{\mathbf{w} : T \rightarrow \mathbb{R}^d \mid \mathbf{w}(x) = a + \beta x, a \in \mathbb{R}^d, \beta \in \mathbb{R}\}$. Furthermore, let

$$Y_h := \{z_h \in L^2(\Omega) \mid z_h \text{ is constant on each } T \in \mathcal{T}_h\}.$$

For a given function $u \in L^r(\Omega)$ the discrete solution $(y_h, \mathbf{v}_h) \in Y_h \times \mathbf{V}_h$ is given by

$$\int_{\Omega} A^{-1} \mathbf{v}_h \cdot \mathbf{w}_h + \int_{\Omega} y_h \operatorname{div} \mathbf{w}_h = 0 \quad \forall \mathbf{w}_h \in \mathbf{V}_h \quad (3.7)$$

$$\int_{\Omega} z_h \operatorname{div} \mathbf{v}_h - \int_{\Omega} a_0 y_h z_h + \int_{\Omega} u z_h = 0 \quad \forall z_h \in Y_h. \quad (3.8)$$

Introducing $\mathcal{G}_h(u) = (y_h, \mathbf{v}_h) \in Y_h \times \mathbf{V}_h$ as an approximation of \mathcal{G} it is well-known ([3]) that the following error estimate holds:

$$\|y - y_h\| + \|\mathbf{v} - \mathbf{v}_h\| \leq Ch(\|y\|_{H^1} + \|A\nabla y\|_{H^1}) \leq Ch\|y\|_{H^2} \leq Ch\|u\| \quad (3.9)$$

by (2.2). In what follows it will be crucial to control the error between \mathbf{v} and \mathbf{v}_h in $L^\infty(\Omega)$.

Lemma 3.4. *Let $d < p < \infty$, $u \in L^p(\Omega)$ and $(y, \mathbf{v}) = \mathcal{G}(u)$, $(y_h, \mathbf{v}_h) = \mathcal{G}_h(u)$. Then there exists $h_0 > 0$ such that for $0 < h \leq h_0$*

$$\|\mathbf{v} - \mathbf{v}_h\|_{L^\infty} \leq Ch^{1-\frac{d}{p}} |\log h|^{1-\frac{2}{p}} \|u\|_{L^p}.$$

Proof. Let us denote by T the linear operator which assigns to u the error $\mathbf{v} - \mathbf{v}_h$. We deduce from (3.9) that

$$\|T\|_{L^2 \rightarrow L^2} \leq Ch.$$

On the other hand we infer from [7, Corollary 3] that there exists $h_0 > 0$ so that for $0 < h \leq h_0$

$$\|\mathbf{v} - \mathbf{v}_h\|_{L^\infty} \leq Ch |\log h| \|u\|_{L^\infty}$$

for all $u \in L^\infty(\Omega)$, so that

$$\|T\|_{L^\infty \rightarrow L^\infty} \leq Ch |\log h|.$$

The Riesz convexity theorem then implies that

$$\|T\|_{L^p \rightarrow L^p} \leq \|T\|_{L^2 \rightarrow L^2}^{\frac{2}{p}} \|T\|_{L^\infty \rightarrow L^\infty}^{1-\frac{2}{p}}$$

and hence

$$\|\mathbf{v} - \mathbf{v}_h\|_{L^p} \leq Ch^{\frac{2}{p}} (h |\log h|)^{1-\frac{2}{p}} \|u\|_{L^p} = h |\log h|^{1-\frac{2}{p}} \|u\|_{L^p}$$

for all $u \in L^p(\Omega)$. Denoting by I_h the usual Lagrange interpolation operator we deduce with the help of standard interpolation estimates, (2.2) and an inverse estimate that

$$\begin{aligned} \|\mathbf{v} - \mathbf{v}_h\|_{L^\infty} &\leq \|\mathbf{v} - I_h \mathbf{v}\|_{L^\infty} + \|I_h \mathbf{v} - \mathbf{v}_h\|_{L^\infty} \\ &\leq ch^{1-\frac{d}{p}} \|\mathbf{v}\|_{W^{1,p}} + ch^{-\frac{d}{p}} \|I_h \mathbf{v} - \mathbf{v}_h\|_{L^p} \\ &\leq ch^{1-\frac{d}{p}} \|u\|_{L^p} + ch^{-\frac{d}{p}} \|\mathbf{v} - I_h \mathbf{v}\|_{L^p} + ch^{-\frac{d}{p}} \|\mathbf{v} - \mathbf{v}_h\|_{L^p} \\ &\leq ch^{1-\frac{d}{p}} \|u\|_{L^p} + ch^{1-\frac{d}{p}} |\log h|^{1-\frac{2}{p}} \|u\|_{L^p} \end{aligned}$$

which yields the result. \square

Similarly to (3.3) we now consider the following discrete control problem:

$$\begin{aligned} \min_{u \in K} J_h(u) &:= \frac{1}{2} \int_{\Omega} |y_h - y_0|^2 + \frac{\alpha}{r} \int_{\Omega} |u|^r \\ \text{subject to } (y_h, \mathbf{v}_h) &= \mathcal{G}_h(u) \text{ and } \left(\int_T A^{-1} \mathbf{v}_h \right)_{T \in \mathcal{T}_h} \in C_h, \end{aligned} \quad (3.10)$$

where C_h is as in (3.2) and $\int_T \cdot = \frac{1}{|T|} \int_T \cdot$. We note that the control again is not discretized and that the gradient of the state variable is only constrained on average on each element. Similar to Lemma 3.1 and Remark 3.2 one has

Lemma 3.5. *Problem (3.10) has a unique solution $u_h \in K$. There exists $0 < h_1 \leq h_0$ such that for $0 < h < h_1$ there are $\mu_T \in \mathbb{R}^d$, $T \in \mathcal{T}_h$ and $(p_h, \chi_h) \in Y_h \times \mathbf{V}_h$ such that with $(y_h, \mathbf{v}_h) = \mathcal{G}_h(u_h)$ we have*

$$\int_{\Omega} A^{-1} \chi_h \cdot \mathbf{w}_h + \int_{\Omega} p_h \operatorname{div} \mathbf{w}_h + \sum_{T \in \mathcal{T}_h} \mu_T \cdot \int_T A^{-1} \mathbf{w}_h = 0 \quad \forall \mathbf{w}_h \in \mathbf{V}_h, \quad (3.11)$$

$$\int_{\Omega} z_h \operatorname{div} \chi_h - \int_{\Omega} a_0 p_h z_h + \int_{\Omega} (y_h - y_0) z_h = 0 \quad \forall z_h \in Y_h, \quad (3.12)$$

$$\int_{\Omega} (p_h + \alpha |u_h|^{r-2} u_h) (\tilde{u} - u_h) \geq 0 \quad \forall \tilde{u} \in K, \quad (3.13)$$

$$\sum_{T \in \mathcal{T}_h} \mu_T \cdot (\mathbf{c}_h|_T - \int_T A^{-1} \mathbf{v}_h) \leq 0 \quad \forall \mathbf{c}_h \in C_h. \quad (3.14)$$

Remark 3.6. The discrete control u_h and the discrete adjoint state p_h are related by

$$\text{(I)} \quad u_h(x) = \operatorname{Proj}_{[a,b]} \left(-\frac{p_h(x)}{\alpha} \right) \text{ a.a. } x \in \Omega,$$

$$\text{(II)} \quad u_h(x) = -\alpha^{-\frac{1}{r-1}} |p_h(x)|^{\frac{2-r}{r-1}} p_h(x) \text{ a.a. } x \in \Omega.$$

In particular, in both cases the discrete solution u_h is piecewise constant on the triangulation \mathcal{T}_h .

The following a-priori estimate is crucial for the convergence analysis.

Lemma 3.7. *Let $u_h \in L^r(\Omega)$ be the optimal solution of (3.10) with corresponding state $(y_h, \mathbf{v}_h) \in Y_h \times \mathbf{V}_h$ and adjoint variables $(p_h, \chi_h) \in Y_h \times \mathbf{V}_h$, μ_T , $T \in \mathcal{T}_h$. Then*

$$\|u_h\|_{L^r} + \|y_h\| + \sum_{T \in \mathcal{T}_h} |\mu_T| \leq C$$

for all $0 < h \leq h_1$.

Proof. The proof is carried out in [5, Lemma 3.6] for case (I), but the analysis can be adapted to case (II) in a straightforward way. \square

The error analysis depends on the choice of the admissible set K and the structure of the objective functional. In case (I) the controls belong to $L^\infty(\Omega)$ leading to better convergence properties in the state equation. We have the following result:

Theorem 3.8. *Let u and u_h be the solutions of (2.3) and (3.10) in case (I) with corresponding states y and y_h respectively. Then*

$$\|u - u_h\| + \|y - y_h\| \leq Ch^{\frac{1}{2}} |\log h|^{\frac{1}{2}}$$

for all $0 < h \leq h_1$.

Proof. see [5, Theorem 4.1]. □

Let us next turn to case (II), for which Theorem 3.3 gives convergence rates of $O(h^{\frac{1}{r}(1-\frac{d}{r})})$ for the control and $O(h^{\frac{1}{2}(1-\frac{d}{r})})$ for the state if a piecewise linear approximation of the state is used. Adapting the corresponding proof to the case of the Raviart–Thomas element it would be possible to derive the same convergence rates. However, as observed in [12, Remark 2], these rates are not optimal since $u \in W^{\frac{1-\frac{d}{r}-\epsilon}{r-1}, r}(\Omega)$ for any $\epsilon > 0$. The following result improves these bounds and appears to be optimal as far as the control variable is concerned.

Theorem 3.9. *Let u and u_h be the solutions of (2.3) and (3.10) in case (II) with corresponding states y and y_h respectively. Then for every $\rho > 0$ there exists C_ρ such that*

$$\|u - u_h\|_{L^r} \leq C_\rho h^{\alpha_1 - \rho}, \quad \|y - y_h\| \leq C_\rho h^{\alpha_2 - \rho},$$

where $\alpha_1 = \frac{1-\frac{d}{r}}{r-1}$, $\alpha_2 = (1 - \frac{d}{r}) \frac{r}{2(r-1)}$.

Proof. To begin, we note that for $r \geq 2$

$$(|a|^{r-2}a - |b|^{r-2}b)(a - b) \geq 2^{2-r}|a - b|^r \quad \forall a, b \in \mathbb{R}.$$

Hence, using (2.6) and (3.13),

$$\begin{aligned} \alpha 2^{2-r} \int_{\Omega} |u - u_h|^r &\leq \alpha \int_{\Omega} (|u|^{r-2}u - |u_h|^{r-2}u_h)(u - u_h) \\ &= \int_{\Omega} p_h(u - u_h) + \int_{\Omega} p(u_h - u) \equiv I + II. \end{aligned} \quad (3.15)$$

Let us introduce $(\tilde{y}_h, \tilde{\mathbf{v}}_h) = \mathcal{G}_h(u) \in Y_h \times \mathbf{V}_h$. Using (3.8) and (3.11) we infer for the first term

$$\begin{aligned} I &= - \int_{\Omega} p_h \operatorname{div}(\tilde{\mathbf{v}}_h - \mathbf{v}_h) + \int_{\Omega} a_0 p_h(\tilde{y}_h - y_h) \\ &= \int_{\Omega} A^{-1} \chi_h \cdot (\tilde{\mathbf{v}}_h - \mathbf{v}_h) + \sum_{T \in \mathcal{T}_h} \mu_T \cdot \int_T A^{-1}(\tilde{\mathbf{v}}_h - \mathbf{v}_h) + \int_{\Omega} a_0 p_h(\tilde{y}_h - y_h) \\ &= \int_{\Omega} A^{-1} \chi_h \cdot (\tilde{\mathbf{v}}_h - \mathbf{v}_h) + \sum_{T \in \mathcal{T}_h} \mu_T \cdot \left(P_\delta \left(\int_T A^{-1} \tilde{\mathbf{v}}_h \right) - \int_T A^{-1} \mathbf{v}_h \right) \\ &\quad + \int_{\Omega} a_0 p_h(\tilde{y}_h - y_h) + \sum_{T \in \mathcal{T}_h} \mu_T \cdot \left(\int_T A^{-1} \tilde{\mathbf{v}}_h - P_\delta \left(\int_T A^{-1} \tilde{\mathbf{v}}_h \right) \right), \end{aligned}$$

where P_δ denotes the orthogonal projection onto $\bar{B}_\delta(0) = \{x \in \mathbb{R}^d \mid |x| \leq \delta\}$. Note that

$$|P_\delta(x) - P_\delta(\tilde{x})| \leq |x - \tilde{x}| \quad \forall x, \tilde{x} \in \mathbb{R}^d. \quad (3.16)$$

Since by definition

$$\left(P_\delta\left(\int_T A^{-1}\tilde{\mathbf{v}}_{\mathbf{h}}\right)\right)_{T \in \mathcal{T}_h} \in C_h$$

we deduce from (3.14) that

$$\begin{aligned} I &\leq \int_\Omega A^{-1}\chi_{\mathbf{h}} \cdot (\tilde{\mathbf{v}}_{\mathbf{h}} - \mathbf{v}_{\mathbf{h}}) + \int_\Omega a_0 p_h(\tilde{y}_h - y_h) \\ &\quad + \max_{T \in \mathcal{T}_h} \left| \int_T A^{-1}\tilde{\mathbf{v}}_{\mathbf{h}} - P_\delta\left(\int_T A^{-1}\tilde{\mathbf{v}}_{\mathbf{h}}\right) \right| \sum_{T \in \mathcal{T}_h} |\mu_T|. \end{aligned}$$

In order to estimate the last term we note that $\nabla y \in C$ implies that $(\int_T \nabla y)_{T \in \mathcal{T}_h} = (\int_T A^{-1}\mathbf{v})_{T \in \mathcal{T}_h} \in C_h$. Using Lemma 3.4 with $(y, \mathbf{v}) = \mathcal{G}(u)$, $(\tilde{y}_h, \tilde{\mathbf{v}}_{\mathbf{h}}) = \mathcal{G}_h(u)$ we infer

$$\|\tilde{\mathbf{v}}_{\mathbf{h}} - \mathbf{v}\|_{L^\infty} \leq Ch^{1-\frac{d}{p_\epsilon}} |\log h|^{1-\frac{2}{p_\epsilon}} \|u\|_{L^{p_\epsilon}} = C_\epsilon h^{1-\frac{d}{p_\epsilon}} |\log h|^{1-\frac{2}{p_\epsilon}}, \quad (3.17)$$

since $u \in L^{p_\epsilon}(\Omega)$ with $p_\epsilon = \frac{r-1}{1-\frac{1}{d}+\epsilon}$ ($\epsilon > 0$) in view of Remark 2.2. As a consequence,

$$\begin{aligned} \left| \int_T A^{-1}\tilde{\mathbf{v}}_{\mathbf{h}} - P_\delta\left(\int_T A^{-1}\tilde{\mathbf{v}}_{\mathbf{h}}\right) \right| &\leq \left| \int_T A^{-1}(\tilde{\mathbf{v}}_{\mathbf{h}} - \mathbf{v}) \right| + \left| P_\delta\left(\int_T A^{-1}(\tilde{\mathbf{v}}_{\mathbf{h}} - \mathbf{v})\right) \right| \\ &\leq C \|\tilde{\mathbf{v}}_{\mathbf{h}} - \mathbf{v}\|_{L^\infty} \leq C_\epsilon h^{1-\frac{d}{p_\epsilon}} |\log h|^{1-\frac{2}{p_\epsilon}} \end{aligned}$$

in view of (3.17). Combining this estimate with Lemma 3.7 we deduce

$$I \leq \int_\Omega A^{-1}\chi_{\mathbf{h}} \cdot (\tilde{\mathbf{v}}_{\mathbf{h}} - \mathbf{v}_{\mathbf{h}}) + \int_\Omega a_0 p_h(\tilde{y}_h - y_h) + C_\epsilon h^{1-\frac{d}{p_\epsilon}} |\log h|^{1-\frac{2}{p_\epsilon}}.$$

The symmetry of A , (3.7) and (3.12) finally give

$$\begin{aligned} I &\leq - \int_\Omega (\tilde{y}_h - y_h) \operatorname{div} \chi_{\mathbf{h}} + \int_\Omega a_0 p_h(\tilde{y}_h - y_h) + C_\epsilon h^{1-\frac{d}{p_\epsilon}} |\log h|^{1-\frac{2}{p_\epsilon}} \\ &= \int_\Omega (y_h - y_0)(\tilde{y}_h - y_h) + C_\epsilon h^{1-\frac{d}{p_\epsilon}} |\log h|^{1-\frac{2}{p_\epsilon}}. \end{aligned} \quad (3.18)$$

In order to analyze the second term in (3.15) we define $(y^h, \mathbf{v}^h) = \mathcal{G}(u_h)$. Recalling (2.5) we have

$$\begin{aligned} II &= \int_\Omega p(\mathcal{A}y^h - \mathcal{A}y) \\ &= \int_\Omega (y - y_0)(y^h - y) + \int_{\bar{\Omega}} (\nabla y^h - \nabla y) \cdot d\mu \\ &= \int_\Omega (y - y_0)(y^h - y) + \int_{\bar{\Omega}} (P_\delta(\nabla y^h) - \nabla y) \cdot d\mu + \int_{\bar{\Omega}} (\nabla y^h - P_\delta(\nabla y^h)) \cdot d\mu. \end{aligned}$$

Since $x \mapsto P_\delta(\nabla y^h(x)) \in C$ we infer from (2.7)

$$II \leq \int_\Omega (y - y_0)(y^h - y) + \max_{x \in \bar{\Omega}} |\nabla y^h(x) - P_\delta(\nabla y^h(x))| \|\mu\|_{\mathcal{M}(\bar{\Omega})^d}. \quad (3.19)$$

Let $x \in \bar{\Omega}$, say $x \in T$ for some $T \in \mathcal{T}_h$. Since u_h is feasible for (3.10) we have that $\int_T A^{-1} \mathbf{v}_h \in \bar{B}_\delta(0)$ so that (3.16) implies

$$\begin{aligned} & |\nabla y^h(x) - P_\delta(\nabla y^h(x))| \\ & \leq |\nabla y^h(x) - \int_T A^{-1} \mathbf{v}_h| + |P_\delta(\nabla y^h(x)) - P_\delta(\int_T A^{-1} \mathbf{v}_h)| \\ & \leq 2 |\nabla y^h(x) - \int_T A^{-1} \mathbf{v}_h|. \end{aligned} \quad (3.20)$$

Using a well-known interpolation estimate (cf. [2], Corollary (4.4.7)) and (2.2) we obtain

$$\begin{aligned} |\nabla y^h(x) - \int_T A^{-1} \mathbf{v}_h| &= |A^{-1}(x) \mathbf{v}^h(x) - \int_T A^{-1} \mathbf{v}_h| \\ &\leq |A^{-1}(x)(\mathbf{v}^h - \mathbf{v})(x) - \int_T A^{-1}(\mathbf{v}^h(x) - \mathbf{v})| \\ &\quad + |A^{-1}(x) \mathbf{v}(x) - \int_T A^{-1} \mathbf{v}| + |\int_T A^{-1}(\mathbf{v}^h - \mathbf{v}_h)| \\ &\leq Ch^{1-\frac{d}{r}} \|\mathbf{v}^h - \mathbf{v}\|_{W^{1,r}} + Ch^{1-\frac{d}{p_\epsilon}} \|\mathbf{v}\|_{W^{1,p_\epsilon}} + C \|\mathbf{v}^h - \mathbf{v}_h\|_{L^\infty} \\ &\leq Ch^{1-\frac{d}{r}} \|u_h - u\|_{L^r} + Ch^{1-\frac{d}{p_\epsilon}} \|u\|_{L^{p_\epsilon}} + C \|\mathbf{v}^h - \mathbf{v}_h\|_{L^\infty}. \end{aligned}$$

Applying Lemma 3.4 with $u - u_h$ as well as (3.17) we infer

$$\begin{aligned} \|\mathbf{v}^h - \mathbf{v}_h\|_{L^\infty} &\leq \|(\mathbf{v}^h - \mathbf{v}) - (\mathbf{v}_h - \tilde{\mathbf{v}}_h)\|_{L^\infty} + \|\mathbf{v} - \tilde{\mathbf{v}}_h\|_{L^\infty} \\ &\leq Ch^{1-\frac{d}{r}} |\log h|^{1-\frac{2}{r}} \|u - u_h\|_{L^r} + C_\epsilon h^{1-\frac{d}{p_\epsilon}} |\log h|^{1-\frac{2}{p_\epsilon}}, \end{aligned}$$

which combined with (3.20) yields

$$\max_{x \in \bar{\Omega}} |\nabla y^h(x) - P_\delta(\nabla y^h(x))| \leq Ch^{1-\frac{d}{r}} |\log h|^{1-\frac{2}{r}} \|u - u_h\|_{L^r} + C_\epsilon h^{1-\frac{d}{p_\epsilon}} |\log h|^{1-\frac{2}{p_\epsilon}}.$$

Returning to (3.19) we have

$$II \leq \int_\Omega (y - y_0)(y^h - y) + Ch^{1-\frac{d}{r}} |\log h|^{1-\frac{2}{r}} \|u - u_h\|_{L^r} + C_\epsilon h^{1-\frac{d}{p_\epsilon}} |\log h|^{1-\frac{2}{p_\epsilon}}. \quad (3.21)$$

If we insert (3.21) and (3.18) into (3.15) we finally obtain

$$\begin{aligned} \alpha 2^{2-r} \|u - u_h\|_{L^r}^r &\leq \int_\Omega (y_h - y_0)(\tilde{y}_h - y_h) + \int_\Omega (y - y_0)(y^h - y) \\ &\quad + Ch^{1-\frac{d}{r}} |\log h|^{1-\frac{2}{r}} \|u_h - u\|_{L^r} + C_\epsilon h^{1-\frac{d}{p_\epsilon}} |\log h|^{1-\frac{2}{p_\epsilon}} \\ &= - \int_\Omega |y - y_h|^2 + \int_\Omega ((y_0 - y_h)(y - \tilde{y}_h) + (y - y_0)(y^h - y_h)) \\ &\quad + Ch^{1-\frac{d}{r}} |\log h|^{1-\frac{2}{r}} \|u_h - u\|_{L^r} + C_\epsilon h^{1-\frac{d}{p_\epsilon}} |\log h|^{1-\frac{2}{p_\epsilon}} \\ &\leq -\|y - y_h\|^2 + C(\|y - \tilde{y}_h\| + \|y^h - y_h\|) \\ &\quad + \frac{\alpha 2^{2-r}}{2} \|u - u_h\|_{L^r}^r + Ch^{(1-\frac{d}{r})\frac{r}{r-1}} |\log h|^{(1-\frac{2}{r})\frac{r}{r-1}} + C_\epsilon h^{1-\frac{d}{p_\epsilon}} |\log h|^{1-\frac{2}{p_\epsilon}} \end{aligned}$$

by Young's inequality. A simple calculation shows that

$$1 - \frac{d}{p_\epsilon} = \left(1 - \frac{d}{r}\right) \frac{r}{r-1} - \frac{\epsilon d}{r-1} < \left(1 - \frac{d}{r}\right) \frac{r}{r-1},$$

while $1 - \frac{2}{p_\epsilon} < (1 - \frac{2}{r})\frac{r}{r-1}$. In conclusion we obtain after another application of (3.9)

$$\|u - u_h\|_{L^r}^r + \|y - y_h\|^2 \leq C_\epsilon h^{(1-\frac{d}{r})\frac{r}{r-1} - \frac{\epsilon d}{r-1}} |\log h|^{(1-\frac{2}{r})\frac{r}{r-1}},$$

from which we deduce the result of the theorem. \square

4. Numerical examples

We consider (2.3) with the choices $\Omega = B_2(0) \subset \mathbb{R}^2$, $\alpha = 1$,

$$C = \{\mathbf{z} \in C^0(\bar{\Omega})^2 \mid |\mathbf{z}(x)| \leq \frac{1}{2}, x \in \bar{\Omega}\}$$

as well as

$$y_0(x) := \begin{cases} \frac{1}{4} + \frac{1}{2} \ln 2 - \frac{1}{4}|x|^2, & 0 \leq |x| \leq 1, \\ \frac{1}{2} \ln 2 - \frac{1}{2} \ln |x|, & 1 < |x| \leq 2. \end{cases}$$

In order to construct a test example we allow an additional right hand side f in the state equation and replace (2.1) by

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

where

$$f(x) := \begin{cases} 2, & 0 \leq |x| \leq 1, \\ 0, & 1 < |x| \leq 2. \end{cases}$$

In case **(I)** we consider $K = \{u \in L^\infty(\Omega) \mid -2 \leq u \leq 2 \text{ a.e. in } \Omega\}$, while in case **(II)** we choose $r = 4$. The optimization problem then has in both cases the unique solution

$$u(x) = \begin{cases} -1 & , 0 \leq |x| \leq 1 \\ 0 & , 1 < |x| \leq 2 \end{cases}$$

with corresponding state $y \equiv y_0$. We note that in case **(I)** the bounds on the control are not active, so that we obtain equality in (2.6), i.e. $p = -u$. Furthermore, the action of the measure μ applied to a vectorfield $\phi \in C^0(\bar{\Omega})^2$ is given by

$$\int_{\bar{\Omega}} \phi \cdot d\mu = - \int_{\partial B_1(0)} x \cdot \phi dS.$$

In what follows we frequently use the experimental order of convergence, which is defined for an error functional $E(h)$ by

$$\text{EOC} = \frac{\ln E(h_1) - \ln E(h_2)}{\ln h_1 - \ln h_2}.$$

For the numerical solution we use the routine `fmincon` contained in the Matlab optimization toolbox. The actual calculations were carried out on a polygonal approximation of $B_2(0)$. Note that while our analysis did not take into account the approximation of the domain, the observed rates show that this error doesn't dominate.



FIGURE 1. Control (left), and adjoint state (right) (variational discretization)

nt	$\ u - u_h\ _{L^4(\Omega)}$	$\ u - u_h\ $	$\ y - y_h\ $
32	$8.34633 \cdot 10^{-1}$	1.36003	$2.20346 \cdot 10^{-1}$
128	$5.88566 \cdot 10^{-1}$	$9.04770 \cdot 10^{-1}$	$7.97200 \cdot 10^{-2}$
512	$4.84191 \cdot 10^{-1}$	$5.82014 \cdot 10^{-1}$	$3.52102 \cdot 10^{-2}$
	0.54884	0.64041	1.59745
	0.29263	0.66136	1.22499

TABLE 1. Errors (top) and EOCs for piecewise linears

4.1. Piecewise linears for the state with variational discretization

Many existing finite element codes employ continuous, piecewise linear finite elements, so that it is natural to use this element in order to discretize the state equation in optimization problems for elliptic pdes. Numerical results for case **(II)** are reported in [8] to which we refer for details. Table 1 shows the experimental order of convergence for the error functionals

$$\|u - u_h\|_{L^4(\Omega)}, \quad \|u - u_h\|, \quad \text{and} \quad \|y - y_h\|.$$

Fig. 1 illustrates the optimal solution u_h and the corresponding adjoint state p_h on a mesh consisting of $nt = 512$ triangles. Note that in view of the relation $u_h(x) = -|p_h(x)|^{-\frac{2}{3}} p_h(x)$ the variational control u_h necessarily is a continuous function, while the exact control u has a jump. This inconsistency is reflected in the appearance of oscillations near the set $\partial B_1(0)$ in Fig. 1, and also affects the performance of the optimization solvers implemented within the `fmincon` package. We conclude that variational discretization combined with continuous, piecewise linear finite elements for the state approximation is not ideally suited to control problems with gradient constraints on the state.

4.2. Mixed finite element approach with variational discretization.

The state equation is now approximated with the help of the lowest order Raviart–Thomas element for which we used the implementation provided by [1]. Numerical results for case **(I)** can be found in [5].

Let us report on the corresponding results for case **(II)**. In Table 2 we display the

NT	$\ u - u_h\ _{L^4(\Omega)}$	$\ u - u_h\ $	$\ y - y_h\ $
32	$6.85 \cdot 10^{-1}$	1.10	$3.00 \cdot 10^{-1}$
128	$6.77 \cdot 10^{-1}$	$8.70 \cdot 10^{-1}$	$1.51 \cdot 10^{-1}$
512	$6.05 \cdot 10^{-1}$	$6.04 \cdot 10^{-1}$	$7.25 \cdot 10^{-2}$
2048	$5.22 \cdot 10^{-1}$	$4.21 \cdot 10^{-1}$	$3.61 \cdot 10^{-2}$
8192	$4.44 \cdot 10^{-1}$	$2.96 \cdot 10^{-1}$	$1.80 \cdot 10^{-2}$
	0.01881	0.36245	1.08340
	0.16899	0.54697	1.09552
	0.21730	0.53219	1.02287
	0.23488	0.51182	1.01139

TABLE 2. Errors and EOCs for the controls and the state with Raviart–Thomas approximation of the state

NT	$\sum_{i=1}^{NT} \mu_T $
32	2.32
128	4.32
512	5.29
2048	5.79
8192	6.04

TABLE 3. Behaviour of the discrete multipliers

experimental order of convergence for the error functionals

$$\|u - u_h\|_{L^4(\Omega)}, \quad \|u - u_h\| \quad \text{and} \quad \|y - y_h\|.$$

The errors show a similar behaviour as in the case of piecewise linear finite elements and are slightly better than predicted by Theorem 3.9. Fig. 2 shows the optimal state and the optimal control on a grid containing $m = 1089$ gridpoints. In Table 3 we display the values of $\sum_{T \in \mathcal{T}_h} |\mu_T|$ which appear to converge to 2π , the total variation of the measure μ . The modulus of $\mu_T, T \in \mathcal{T}_h$ as well as the set of elements T on which $\mu_T \neq 0$ is shown in Fig. 3. It can be seen that these elements concentrate around $|x| = 1$.

References

- [1] Bahriawati, C., Carstensen, C., *Three Matlab Implementations Of The Lowest-Order Raviart-Thomas MFEM With A Posteriori Error Control*. Computational Methods in Applied Mathematics **5** (2005), 333–361. Software download at www.math.hu-berlin.de/cc/download/public/software/code/Software-4.tar.gz.
- [2] Brenner, S., Scott, R., *The mathematical theory of finite elements*. 2nd edition, Springer New York, 2002.
- [3] Brezzi, F., Fortin, M., *Mixed and hybrid finite element methods*. Springer Series in Computational Mathematics **15**, Springer–Verlag, New York, 1991.

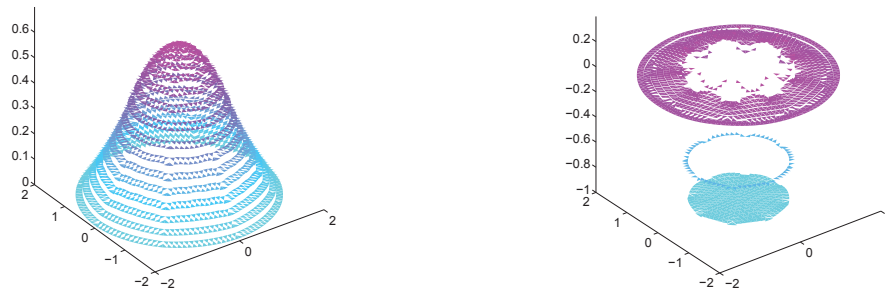
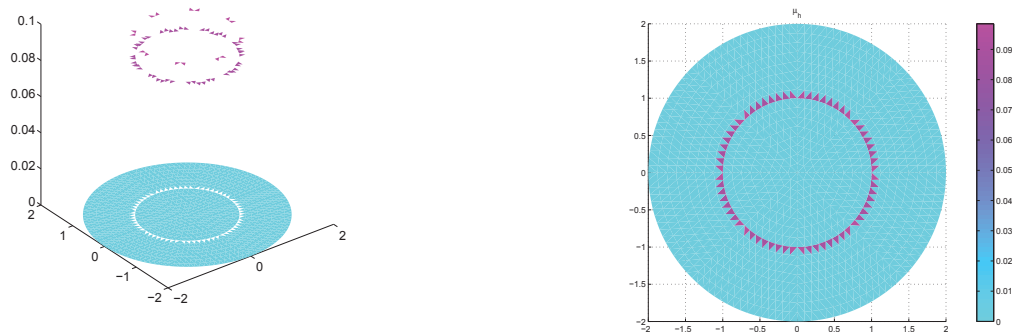


FIGURE 2. Optimal state (left), and optimal control

FIGURE 3. $|\mu_T|$ (left), and support of μ_T

- [4] Casas, E., Fernández, L., *Optimal control of semilinear elliptic equations with pointwise constraints on the gradient of the state*. Appl. Math. Optimization **27** (1993), 35–56.
- [5] Deckelnick, K., Günther, A., and Hinze, M., *Finite element approximation of elliptic control problems with constraints on the gradient*. Numer. Math. **111** (2009), 335-350.
- [6] Deckelnick, K., Hinze, M., *Convergence of a finite element approximation to a state constrained elliptic control problem*. SIAM J. Numer. Anal. **45**, No.5 (2007), 1937–1953.
- [7] Gastaldi, L., Nochetto, R.H., *On L^∞ -accuracy of mixed finite element methods for second order elliptic problems*. Mat. Apl. Comput. **7** (1988), 13–39.
- [8] Günther, A., Hinze, M., *Elliptic Control Problems with Gradient Constraints - Variational Discrete Versus Piecewise Constant Controls*. Comput. Optim. Appl. **49** (2011), 549-566.
- [9] Hintermüller, M., Kunisch, K., *PDE-constrained optimization subject to pointwise constraints on the control, the state, and its derivative*. SIAM J. Optim. **20**(3) (2009), 1133-1156.
- [10] Hintermüller, M., Hinze, M. and Hoppe, R.W.H., *Weak-Duality based Adaptive Finite Element Methods for PDE-Constrained Optimization with Pointwise Gradient State-Constraints*. J. Comp. Math. **30** (2012), 101-123.

- [11] Hinze, M., *A variational discretization concept in control constrained optimization: the linear-quadratic case*. *Comput. Optim. Appl.* **30** (2005), 45–63.
- [12] Ortner, C., Wollner, W., *A Priori Error Estimates for Optimal Control Problems with Pointwise Constraints on the Gradient of the State*. *Numer. Math.* **118**(3) (2011), 587–600.
- [13] Schiela, A., Wollner, W., *Barrier methods for optimal control problems with convex nonlinear gradient state constraints*. *SIAM J. Optim.* **21**(1) (2011), 269–286.
- [14] Triebel, H., *Interpolation theory, function spaces, differential operators..* 2nd edition, Johann Ambrosius Barth, Heidelberg, 1995.
- [15] 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.* **47**(1) (2010), 133–159.
- [16] Wollner, W., *Optimal Control of Elliptic Equations with Pointwise Constraints on the Gradient of the State in Nonsmooth Polygonal Domains*. *SIAM J. Control Optim.* **50**(4) (2012), 2117–2129.
- [17] Wollner, W., *A Priori Error Estimates for Optimal Control Problems with Constraints on the Gradient of the State on Nonsmooth Polygonal Domains*. In: *Control and Optimization with PDE Constraints*, Springer ISNM Series, Vol. **164**, 193–215, Springer Basel (2013).

Acknowledgment

The authors acknowledge support of the DFG Priority Program 1253 through grants DFG HI689/5-1 and DFG DE611/4-2.

Klaus Deckelnick
Institut für Analysis und Numerik
Otto–von–Guericke–Universität Magdeburg
Universitätsplatz 2
39106 Magdeburg
Germany
e-mail: klaus.deckelnick@ovgu.de

Michael Hinze
Schwerpunkt Optimierung und Approximation
Universität Hamburg
Bundesstraße 55
20146 Hamburg
Germany
e-mail: michael.hinze@uni-hamburg.de