Discretization error estimates for an optimal control problem in a nonconvex domain

Th. Apel[1], thomas.apel@unibw.de, A. Rösch[2], arnd.roesch@oeaw.ac.at, and G. Winkler[1], gunter.winkler@unibw.de

[1] Institut für Mathematik und Bauinformatik, Fakultät für Bauingenieur- und Vermessungswesen, Universität der Bundeswehr München, 85577 Neubiberg, Germany

[2] Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenbergerstraße 69, A-4040 Linz, Austria

Summary. An optimal control problem for a 2-d elliptic equation and with pointwise control constraints is investigated. The domain is assumed to be polygonal but non-convex. The corner singularities are treated by a priori mesh grading. A second order approximation of the optimal control is constructed by a projection of the discrete adjoint state. Here we summarize the results from [1] and add further numerical tests.

Key words: Linear-quadratic optimal control problems, error estimates, corner singularities, control constraints, superconvergence.

1 Introduction

This paper is concerned with the a 2-d elliptic optimal control problem with pointwise control constraints. The state and the adjoint state are discretized by continuous, piecewise linear functions on a family of graded finite element meshes. The control is initially discretized with piecewise constants on the same meshes, but this control is used only for solving the system of discretized equations. Finally, an improved control is constructed by postprocessing the adjoint state. This approach was suggested and analysed for sufficiently smooth solutions by Meyer and Rösch [3]. The results of our analysis [1] of the case of non-smooth solutions are summarized in Section 2.

In Section 3, we present some new numerical tests of this method. It can be seen that graded meshes are indeed suited to retain the convergence order of smooth solutions in the non-smooth case. Moreover, we see that the boundary between active and non-active controls is approximated well although the method does not specially target to this aim. The results show that it is not necessary to adapt the mesh to these a priori unknown curves.

2 Theory

In this section, we summarize our results from [1]; and therefore we closely follow that paper. We consider the elliptic optimal control problem

$$J(\bar{u}) = \min_{u \in U_{\mathrm{ad}}} J(u), \quad J(u) := F(Su, u), \tag{1}$$

$$F(y, u) := \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)}, \tag{2}$$

where the associated state $y = Su$ to the control u is the weak solution of the state equation

$$Ly = u \quad \text{in } \Omega, \qquad y = 0 \quad \text{on } \Gamma = \partial\Omega, \tag{3}$$

and the control variable is constrained by

$$a \le u(x) \le b \qquad \text{for a.a. } x \in \Omega. \tag{4}$$

The function $y_d \in L^\infty(\Omega)$ is the desired state, a and b are real numbers, and the regularization parameter $\nu > 0$ is a fixed positive number. Moreover, $\Omega \subset \mathbb{R}^2$ is a bounded polygonal domain with boundary Γ. The set of admissible controls is $U_{\mathrm{ad}} := \{u \in L^2(\Omega) : \ a \le u \le b \ \text{ a.e. in } \Omega\}$. The second order elliptic operator L is defined by

$$Ly := -\nabla \cdot (A\nabla y) + \mathbf{a} \cdot \nabla y + a_0 y, \tag{5}$$

where the coefficients $A = A^T \in C(\bar{\Omega}, \mathbb{R}^{2\times 2})$, $\mathbf{a} \in C(\bar{\Omega}, \mathbb{R}^2)$, $a_0 \in C(\bar{\Omega})$, satisfy the usual ellipticity and coercivity conditions $\xi^T A\xi \ge m_0 \xi^T \xi$ for all $\xi \in \mathbb{R}^2$ and $a_0 - \frac{1}{2}\nabla \cdot \mathbf{a} \ge 0$.

We focus on state equations with non-smooth solutions. Let us assume that the domain $\Omega \subset \mathbb{R}^2$ has exactly one reentrant corner with interior angle $\omega > \pi$ located at the origin. Due to the local nature of corner singularities in elliptic problems this not a loss of generality. We denote by $r := r(x) = |x|$ the Euclidean distance to this corner. The solution of the elliptic boundary value problem

$$Ly = g \quad \text{in } \Omega, \qquad y = 0 \quad \text{on } \Gamma,$$

has typically an r^λ-singularity where $\lambda \in (1/2, 1)$ is a real number which is defined by the coefficient matrix A and the angle ω. In the case of the Dirichlet problem for the Laplace operator, the value of λ is explicitly known, $\lambda = \pi/\omega$. In more general cases this can also be computed.

Via (3), the operator S associates a state $y = Su$ to the control u. We denote by S^* the solution operator of the adjoint problem

$$L^* p = y - y_d \quad \text{in } \Omega, \qquad p = 0 \quad \text{on } \Gamma, \tag{6}$$

that means, we have $p = S^*(y - y_d)$. Since we can also write $p = S^*(Su - y_d) = Pu$ with an affine operator P we call the solution $p = Pu$ the associated adjoint state to u.

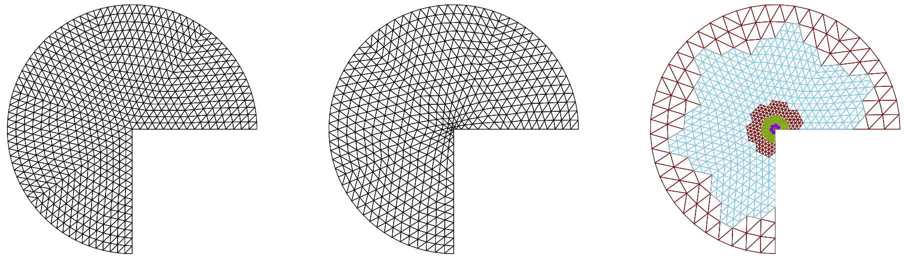

Fig. 1. Ω with a quasi-uniform mesh ($\mu = 1.0$) and with graded meshes ($\mu = 0.6$)

Introducing the projection

$$\Pi_{[a,b]} f(x) := \max(a, \min(b, f(x))),$$

the condition

$$\bar{u} = \Pi_{[a,b]} \left(-\frac{1}{\nu} \bar{p} \right). \tag{7}$$

is necessary and sufficient for the optimality of $\bar{u}$.

The optimal control problem is now discretized by a finite element method. We analyze a family of graded triangulations $(T_h)_{h>0}$ of $\bar{\Omega}$ with the global mesh size h and a grading parameter $\mu < \lambda$. We assume that the individual element diameter $h_T := \operatorname{diam} T$ of any element $T \in T_h$ is related to the distance $r_T := \inf_{x \in T} |x|$ of the triangle to the corner by the relation

$$\begin{aligned} c_1 h^{1/\mu} &\le h_T \le c_2 h^{1/\mu} \quad \text{for } r_T = 0, \\ c_1 h r_T^{1-\mu} &\le h_T \le c_2 h r_T^{1-\mu} \text{ for } r_T > 0. \end{aligned} \tag{8}$$

For a 2-dimensional domain the number of elements of such a triangulation is of order h^{-2}. Figure 1 shows an example domain with a uniform mesh and graded meshes. Implementational aspects are given in Section 3. On these meshes, we define the finite element spaces

$$\begin{aligned} U_h &:= \{u_h \in L^\infty(\Omega) : \ u_h|_T \in \mathcal{P}_0 \text{ for all } T \in T_h\}, \quad U_h^{\text{ad}} := U_h \cap U_{\text{ad}}, \\ V_h &:= \{y_h \in C(\bar{\Omega}) : \ y_h|_T \in \mathcal{P}_1 \text{ for all } T \in T_h \text{ and } y_h = 0 \text{ on } \Gamma\}, \end{aligned}$$

where $\mathcal{P}_k$, $k = 0, 1$, is the space of polynomials of degree less than or equal to k.

For each $u \in L^2(\Omega)$, we denote by $S_h u$ the unique element of V_h that satisfies $a(S_h u, v_h) = (u, v_h)_{L^2(\Omega)}$ for all $v_h \in V_h$, where $a : \ H^1(\Omega) \times H^1(\Omega) \to \mathbb{R}$ is the bilinear form defined by $a(y, v) := \int_\Omega \left(\nabla y \cdot (A \nabla v) + \mathbf{b} \nabla v + a_0 y v \right) \, dx$. In other words, $S_h u$ is the approximated state associated with a control u.

The finite dimensional approximation of the optimal control problem is defined by

$$J_h(\bar{u}_h) = \min_{u_h \in U_h^{\text{ad}}} J_h(u_h) \tag{9}$$

with $J_h(u_h) := \frac{1}{2}\|S_h u_h - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u_h\|^2_{L^2(\Omega)}$. The adjoint equation is discretized in the same way: We search $p_h = S_h^*(S_h u_h - y_d) = P_h u_h \in V_h$ such that $a(v_h, p_h) = (S_h u_h - y_d, v_h)_{L^2(\Omega)}$ for all $v_h \in V_h$. The optimal control problem (9) admits a unique solution $\bar{u}_h$, and we denote by $\bar{y}_h = S_h \bar{u}_h$ the optimal discrete state and by $\bar{p}_h = P_h \bar{u}_h$ the optimal discrete adjoint state. In analogy to (7) we define a postprocessed approximate control $\tilde{u}_h$ by a simple projection of the piecewise linear adjoint state $\bar{p}_h$ onto the admissible set U_{ad},

$$\tilde{u}_h := \Pi_{[a,b]}\left(-\frac{1}{\nu}\bar{p}_h\right).$$

Let us now summarize discretization error estimates. Under the assumption that the mesh grading parameter μ satisfies the condition

$$\mu < \lambda, \tag{10}$$

the optimal, piecewise constant approximate control $\bar{u}_h$ satisfies

$$\|\bar{u} - \bar{u}_h\|_{L^2(\Omega)} \le ch\left(\|\bar{u}\|_{L^\infty(\Omega)} + \|y_d\|_{L^\infty(\Omega)}\right) \tag{11}$$

The first order convergence is also observed in numerical tests. Although the difference $\bar{u} - \bar{u}_h$ is of first order, the associated states and adjoint states differ by second order,

$$\|\bar{y} - \bar{y}_h\|_{L^2(\Omega)} \le ch^2\left(\|\bar{u}\|_{L^\infty(\Omega)} + \|y_d\|_{L^\infty(\Omega)}\right), \tag{12}$$
$$\|\bar{p} - \bar{p}_h\|_{L^2(\Omega)} \le ch^2\left(\|\bar{u}\|_{L^\infty(\Omega)} + \|y_d\|_{L^\infty(\Omega)}\right), \tag{13}$$

from which one can conclude that the error of the postprocessed control is also of second order,

$$\|\bar{u} - \tilde{u}_h\|_{L^2(\Omega)} \le ch^2\left(\|\bar{u}\|_{L^\infty(\Omega)} + \|y_d\|_{L^\infty(\Omega)}\right). \tag{14}$$

These results were first proved by Meyer and Rösch [3] for uniform meshes in the smooth case, where the solution of $Ly = f$ is contained in $W^{2,2}(\Omega) \cap W^{1,\infty}(\Omega)$. The main result of our paper [1] is that the error estimates (11)–(14) are also valid in the case of non-convex domains and appropriately graded meshes, (10). Without local mesh grading ($\mu = 1$), only a reduced convergence order is observed.

For the proof of the superconvergence results, we needed the following assumption. The formula (7) computes the optimal control $\bar{u}$ by a projection of the adjoint state $\bar{p}$. This reduces the smoothness. While $|r^{1-\mu}\bar{p}|_{W^{2,2}(\Omega)} \le c|r^{1-\mu}\bar{p}|_{W^{2,2}(\Omega)} < \infty$ for $\mu < \lambda < 1$, this is not true for $\bar{u}$ due to kinks at the boundary of the active set. We assume that

$$\sum_{T \in T_h:\, r^{1-\mu}\bar{u} \notin W^{2,2}(T)} \operatorname{meas} T \le ch.$$

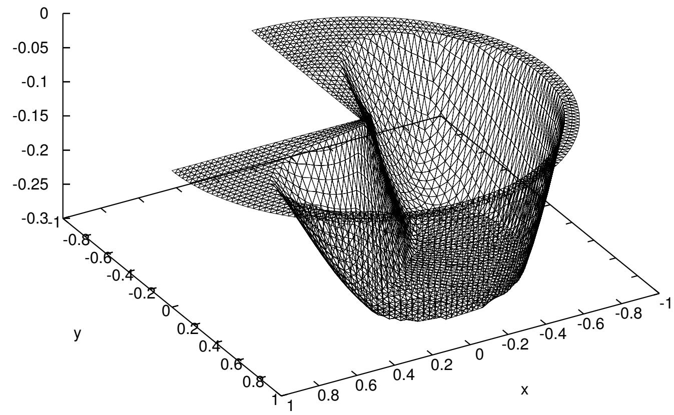

Fig. 2. Example 1. Optimal control function $-0.3 \le u(x) \le -0.1$

3 Numerical Results

Let Ω be a circular sector as shown in Figure 1. In order to construct meshes that fulfil the conditions (8) we transformed the mesh using the mapping $T(x) = x\|x\|^{\frac{1}{\mu}-1}$ near the corner, see figure 1, middle image. An alternative is to use a partitinong strategy, see figure 1, right image.

We choose the example such that the state and dual state have a singularity near the corner. Consider

$$\begin{aligned} -\Delta y + y &= u + f && \text{in } \Omega, \\ -\Delta p + p &= y - y_d && \text{in } \Omega, \\ u &= \Pi_{[a,b]}\left(-\tfrac{1}{\nu}\,p\right) \end{aligned}$$

with homogeneous Dirichlet boundary conditions for y and p.

First Example. In order to have an exact solution we choose the data $f = Ly - u = Ly - \Pi\left(-\frac{1}{\nu}p\right)$ and $y_d = y - L^*p$ such that

$$\begin{aligned} y(r,\varphi) &= (r^\lambda - r^\alpha)\sin\lambda\varphi \\ p(r,\varphi) &= \nu(r^\lambda - r^\beta)\sin\lambda\varphi \end{aligned}$$

are the exact solutions of the optimal control problem. We set $\lambda = \frac{2}{3}$, $\alpha = \beta = \frac{5}{2}$, $\nu = 10^{-4}$, $a = -0.3$ and $b = -0.1$. Figure 2 displays a piecewise linear approximation of the corresponding control function $\bar{u}$. Table 1 shows the reduced convergence rate 2λ on a quasi-uniform mesh ($\mu = 1$) and the optimal rate of convergence of the control on a graded mesh ($\mu = 0.6$).

Figure 3 shows that the error near the corner dominates the global error. The picture visualizes the contribution of each triangle to the global L^2-error. Using graded meshes this error diminishs at least as fast as the global error.

	$\mu = 0.6$		$\mu = 1.0$	
ndof	$\lVert u - \tilde{u} \rVert_{L^2}$	rate	$\lVert u - \tilde{u} \rVert_{L^2}$	rate
18	1.95e-01	0.00	1.95e-01	0.00
55	1.92e-01	0.02	1.92e-01	0.02
189	1.24e-01	0.63	1.31e-01	0.56
697	4.44e-02	1.48	5.87e-02	1.16
2673	1.38e-02	1.69	2.42e-02	1.28
10465	3.79e-03	1.86	9.84e-03	1.30
41409	9.58e-04	1.98	3.93e-03	1.32
164737	2.17e-04	2.14	1.57e-03	1.33

Table 1. Example 1. L^2-error of the computed control $\tilde{u}_h$, $-0.3 \leq u(x) \leq -0.1$

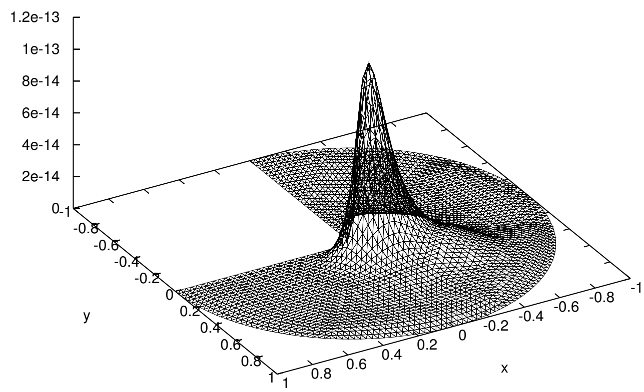

Fig. 3. Example 1. Visualization of the L^2-error of p_h, $-0.3 \leq u(x) \leq 1$, $\mu = 1$

Second Example. We choose now the data f and y_d such that

$$y(r, \varphi) = (r^\lambda - r^\alpha) \sin 3\lambda\varphi$$
$$p(r, \varphi) = \nu(r^\lambda - r^\beta) \sin 3\lambda\varphi$$

with $\lambda = \frac{2}{3}$ and $\alpha = \beta = \frac{5}{2}$. Further we set $a = -0.2$, $b = 0$ and $\nu = 10^{-4}$. We used a mesh that did not even coincide with the boundary of the upper active set $\{x : u(x) = b\}$ in order to show that the method does not need any apriori information about the active set. Figure 4 shows the piecewise constant approximation of the optimal control $\bar{u}$.

Table 2 shows that the convergence rate of the control $\tilde{u}$ is about 2 which was proven in [1]. Table 3 contains the absolute errors and error reduction rates of the approximated state y_h in both the L^2-norm and the H^1-seminorm.

Active Sets. The approximation of the boundary of the active sets is very important for the quality of the computed control, see e.g. [2]. The method presented here approximates the active set by a union of triangles. However, after postprocessing the piecewise linear function $\tilde{u}_h$ gives a much better representation of the active sets. Figure 5 shows the active set of Example 1 on

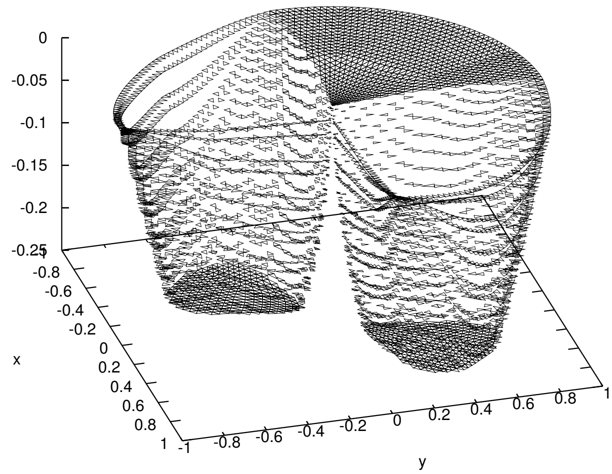

Fig. 4. Example 2. Piecewise constant approximation of optimal control function $\bar{u}$, $-0.2 \le u(x) \le 0$. One can see a singularity near the corner.

	$\mu = 0.6$	
ndof	$\|u - \tilde{u}\|_{L^2}$	rate
18	2.63e-01	0.00
55	2.59e-01	0.02
189	2.33e-01	0.15
697	8.44e-02	1.47
2673	2.36e-02	1.84
10465	6.04e-03	1.96
41409	1.57e-03	1.95
164737	4.31e-04	1.86

Table 2. Example 2. L^2-error of the computed control $\tilde{u}_h$, $-0.2 \le u(x) \le 0$

	$\mu = 0.6$				$\mu = 1$			
ndof	$\|y - y_h\|_{L^2}$	rate	$\|y - y_h\|_{H^1}$	rate	$\|y - y_h\|_{L^2}$	rate	$\|y - y_h\|_{H^1}$	rate
18	1.55e-01	0.00	1.78e+00	0.00	1.55e-01	0.00	1.78e+00	0.00
55	3.92e-02	1.98	1.04e+00	0.77	4.35e-02	1.83	1.10e+00	0.69
189	7.68e-03	2.35	5.74e-01	0.86	1.10e-02	1.98	6.84e-01	0.69
697	1.99e-03	1.94	3.06e-01	0.91	3.55e-03	1.63	4.24e-01	0.69
2673	6.18e-04	1.69	1.61e-01	0.93	1.23e-03	1.53	2.64e-01	0.68
10465	1.58e-04	1.97	8.38e-02	0.94	3.91e-04	1.66	1.65e-01	0.68
41409	3.97e-05	1.99	4.33e-02	0.95	1.23e-04	1.67	1.04e-01	0.67
164737	1.00e-05	1.99	2.22e-02	0.96	3.87e-05	1.67	6.52e-02	0.67

Table 3. Example 2. L^2- and H^1-errors of the computed state y_h, $-0.2 \le u(x) \le 0$

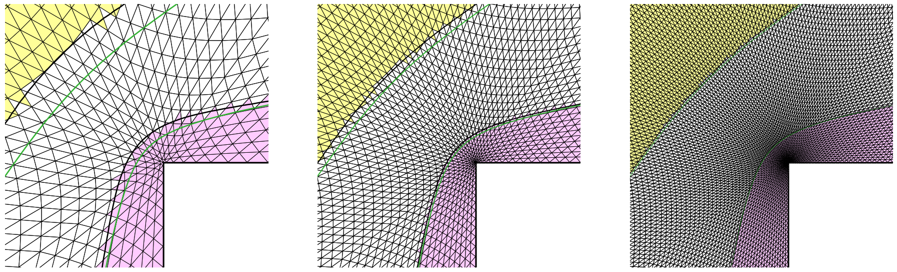

Fig. 5. Example 1: Active triangles and boundary of active sets, (zoom of region near singularity), left: ndof=2673, middle: ndof=10465, right: ndof=41409

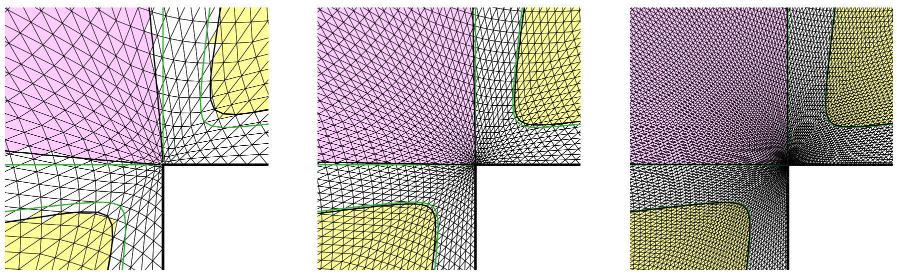

Fig. 6. Example 2: Active triangles and boundary of active sets, $-0.2 \le u(x) \le 0$, $\mu = 1.0$, $\alpha = \beta = \frac{5}{2}$, (zoom of region near singularity), left: ndof=2673, middle: ndof=10465, right: ndof=41409

different meshes. The active triangles are shaded. The black curve shows the computed boundary of the active set as represented by $\tilde{u}_h$. The second curve displays the exact boundary. Clearly, the approximation improves with decreasing mesh size. Figure 6 shows the same behavior for the second example.

References

1. T. Apel, A. Rösch, and G. Winkler, *Optimal control in nonconvex domains*, RICAM Report, 2005-17. `http://www.ricam.oeaw.ac.at/publications/reports/`.
2. M. Hinze, *A variational discretization concept in control constrained optimization: The linear-quadratic case*, Computational Optimization and Applications, 30 (2005), pp. 45–63.
3. C. Meyer and A. Rösch, *Superconvergence properties of optimal control problems*, SIAM J. Control and Optimization, 43 (2004), pp. 970–985.