Hamburger Beiträge
zur Angewandten Mathematik

A priori error estimates for optimal control problems with constraints on the gradient of the state on nonsmooth polygonal domains

Winnifried Wollner

Nr. 2012-02
January 2012

A priori error estimates for optimal control problems with constraints on the gradient of the state on nonsmooth polygonal domains

Winnifried Wollner

Abstract. In this article we are concerned with the finite element discretization of optimal control problems subject to a second order elliptic PDE and additional pointwise constraints on the gradient of the state.

We will derive error estimates for the convergence of the cost functional under mesh refinement. Subsequently error estimates for the control and state variable are obtained.

As an intermediate tool we will also analyze a Moreau-Yosida regularized version of the optimal control problem. In particular we will derive convergence rates for the cost functional and the primal variables. To this end we will employ new techniques in estimating the L^∞ norm of the feasibility error which could also be used to improve existing estimates in the state constrained case.

Mathematics Subject Classification (2010). 49M25, 65N12, 65N30.

Keywords. a priori error estimates, pointwise gradient state constraints, optimization with PDEs, finite elements, nonconvex polygonal domain.

1. Introduction

We are concerned with an analysis of the discretization error for optimal control problems of second order elliptic equations subject to constraints on the gradient of the state. Such problems have some natural application for instance in cooling processes or structural optimization when high stresses have to be avoided.

Despite these interesting applications first order state constraints have hardly been recognized in mathematics. In the works [3,4] the case of optimal control of semilinear elliptic equations with pointwise first order state constraints was studied under the assumption that the domain $\Omega \subset \mathbb{R}^n$ possesses a $C^{1,1}$ boundary. In particular, they studied the adjoint equation and derived first order necessary optimality conditions. It is immediately clear that their results carry over to the case of a polygonally bounded domain, as long as the linearized state equation (with homogeneous Dirichlet boundary values) defines an isomorphism between $W^{2,t}(\Omega) \cap H_0^1(\Omega)$ and $L^t(\Omega)$ for some $t > n$. However, even for $n = 2$ this requires a convex domain which is usually too restrictive for applications. However, in a recent publication [19] it was shown that even on nonconvex domains such problems may remain well posed.

In [10] a Moreau-Yosida based framework for PDE-constrained optimization with constraints on the derivative of the state is developed and used to develop a semismooth Newton algorithm unfortunately their work does not directly carry over to our problem class, because the presence of corner singularities is contradicting the assumptions made in there article. In [16] an investigation of barrier methods for this problem class is conducted.

When concerned with the discretization of the infinite dimensional problem using finite elements, recent results where obtained in [5, 7, 13]. However in all cases the domain was either smooth or polygonally bounded with sufficiently small interior angles. Concerning adaptive discretization methods we refer to [18] and the recent contribution [9].

The rest of this article is structured as follows. In Section 2 we will discuss the problem class under consideration. Then we will consider its discretization in Section 2.1. In Section 3 we will derive an priori error estimate for a certain semi discretization of the problem. The estimates are essentially the same as those obtained in [5, 7, 13]. Unfortunately for this semi discretization the control has to be chosen orthogonal to certain dual singular functions. Since this is not feasible in general we require further analysis. For this purpose we consider a Moreau-Yosida regularization of the state constraint in Section 4. Here we will derive convergence of both the cost functional and the primal variables depending on the penalization parameter. Parts of the analysis will be similar to the work of [8] but with further complications due to the missing regularity of the control-to-state mapping. We will however employ a new L^∞ estimate for the feasibility violation which could also be used to improve the convergence results obtained in [8, 17] for state constrained problems. For the case without corner singularities one can find similar results obtained simultaneously in [11].

With these preparations we can finally derive the main convergence result in Section 5 for a computationally feasible discretization.

2. Problem formulation

In what follows, let $\Omega \subset \mathbb{R}^2$ be a bounded polygonal domain. We are concerned with optimization problems governed by a linear elliptic PDE. For simplicity we consider

$$-\Delta u = q \quad \text{in } \Omega, \tag{2.1a}$$

$$u = 0 \quad \text{on } \partial\Omega. \tag{2.1b}$$

It is then clear, that this operator defines an isomorphism $-\Delta\colon V = H^1_0(\Omega) \to H^{-1}(\Omega)$.

Now we let $r > 2$ be a given number and define $Q = L^r(\Omega)$. We are then particularly interested in an optimal control problem of the form

$$\underset{Q\times V}{\text{Minimize}}\ J(q,u) := \frac{1}{2}\|u-u^d\|^2_{L^2} + \frac{1}{r}\|q\|^r_{L^r}, \tag{2.2a}$$

$$\text{such that } (u,q) \text{ satisfies (2.1)}, \tag{2.2b}$$

$$\text{and such that } |\nabla u| \le 1 \text{ in } \overline{\Omega}. \tag{2.2c}$$

If Ω would be a smooth domain, or a convex polygon, well posedness of (2.2) would follow, e.g, from [4]. However, for a general polygon Ω the results do not carry over easily. This is due to the conflicting nature of the constraint $|\nabla u| \le 1$ and the existence of corner singularities due to the reentrant corners of the domain. This means that for given $q \in Q$ the solution u of (2.1) is neither in $C^1(\overline{\Omega})$ nor in $W^{1,\infty}(\Omega)$. Thus the constraint $|\nabla u| \le 1$ can not be posed easily in this topology. Nonetheless, problem (2.2) is well posed, see [19]. In particular (2.2) admits a unique solution $(\overline{q}, \overline{u}) \in Q \times V$. Moreover there exists a number $t > 2$ depending on the angles in the corners of the domain, such that $\overline{u} \in W^{2,t}(\Omega) \cap V$. Denote the image of $W^{2,t}(\Omega) \cap V$ under $-\Delta$ by I then, again following [19], we have that I is closed in $L^t(\Omega)$. With this preparations we have in addition that $\overline{q} \in I \cap Q$.

For the exposition of this article it is convenient to assume that Ω has only one reentrant corner v with interior angle $\omega > \pi$. Then by inspection of the proof of [19, Lemma 2.2] we obtain the following bounds on t depending on ω

$$\begin{aligned} &t < \frac{\omega}{\pi}\Big(\frac{\omega}{\pi} - 1\Big)^{-1} && \text{if } \omega \in (\pi, 2\pi),\\ &t < 4 && \text{if } \omega = 2\pi. \end{aligned} \tag{2.3}$$

Further restrictions on t are possible due to convex corners of Ω.

This means that we are able to restate problem (2.2) equivalently as follows

$$\underset{Q\cap I\times V}{\text{Minimize}}\ J(q,u) := \frac{1}{2}\|u-u^d\|^2_{L^2} + \frac{1}{r}\|q\|^r_{L^r}, \tag{2.4a}$$

$$\text{such that } (u,q) \text{ satisfies (2.1)}, \tag{2.4b}$$

$$\text{and such that } |\nabla u| \le 1 \text{ in } \overline{\Omega}. \tag{2.4c}$$

2.1. Discretization

In a next step we consider the discretization of these problems. To this end we start by discretizing the state equation (2.1).

Let $(\mathcal{T}_h)_{h\in(0,1]}$ be a given family of triangulations, consisting of triangles or quadrilaterals which are *affine-equivalent* to their respective reference elements, such that $\operatorname{diam}(T) \le h$ for all $T \in \mathcal{T}_h$, $h \in (0,1]$. We assume throughout that the family is quasi-uniform in the sense of [2, Definition 4.4.13], that is, there exists $\rho > 0$ such that, for each $T \in \mathcal{T}_h$ and $h \in (0,1]$ there exists a ball $B_T \subset T$ such that $\operatorname{diam}(B_T) \ge \rho h$.

We define the discrete state space $V_h \subset V$ as the space of continuous piecewise linear (or bi-linear) functions with respect to the mesh $\mathcal{T}_h$.

We remark that the restrictions we imposed on the family $(\mathcal{T}_h)_{h\in(0,1]}$ ensure that the usual interpolation error results, best approximation results, and inverse estimates hold [2, Sec. 4 and 5].

Finally, we define $\Pi_h : L^1(\Omega) \to V_h$ to be the natural extension of the L^2-projection operator, that is, for $q \in L^1(\Omega)$, we define $\Pi_h u \in V_h$ via

$$(\Pi_h q, \varphi) = (q, \varphi) \qquad \forall \varphi \in Q^h. \tag{2.5}$$

It is shown in [6] that Π_h is stable as an operator from $L^p(\Omega)$ to $L^p(\Omega)$, for any $p \in [1, \infty]$, that is, there exist constants c_p, independent of h, such that

$$\|\Pi_h f\|_{L^p} \le c_p \|f\|_{L^p} \qquad \forall f \in L^p(\Omega). \tag{2.6}$$

Now we can discretize the state equation. For fixed $q \in Q$ we search for a solution of the following

$$(\nabla u_h, \nabla \varphi_h) = (q, \varphi_h) \quad \forall\, \varphi_h \in V_h. \tag{2.7}$$

This is already sufficient to obtain a finite dimensional optimization problem. This is due to the fact, that it is sufficient to consider equivalence classes of functions $q, p \in Q$ given by the identification $\Pi_h q = \Pi_h p$ as controls. Then for the minimization of the cost functional in (2.2) it is sufficient to take the unique element out of these classes with minimal L^r-Norm. Moreover due to the first order optimality conditions these elements can be expressed in an explicit way, see, e.g., [12] where this idea was explored first.

In particular the discretized version of (2.2) becomes

$$\operatorname*{Minimize}_{Q\times V_h} J(q_h, u_h) := \frac{1}{2}\|u_h - u^d\|_{L^2}^2 + \frac{1}{r}\|q_h\|_{L^r}^r, \tag{2.8a}$$

$$\text{such that } (u_h, q_h) \text{ satisfies (2.7)}, \tag{2.8b}$$

$$\text{and such that } |\nabla u_h| \le 1 \text{ a.e. in } \overline{\Omega}. \tag{2.8c}$$

In addition, we can also discretize (2.4) and get

$$\operatorname*{Minimize}_{Q\cap I\times V_h} J(q_h, u_h) := \frac{1}{2}\|u_h - u^d\|_{L^2}^2 + \frac{1}{r}\|q_h\|_{L^r}^r, \tag{2.9a}$$

$$\text{such that } (u_h, q_h) \text{ satisfies (2.7)}, \tag{2.9b}$$

$$\text{and such that } |\nabla u_h| \le 1 \text{ a.e. in } \overline{\Omega}. \tag{2.9c}$$

Unlike the continuous case the optimal control problems (2.8) and (2.9) are not equivalent. We will start with an analysis of (2.9). This analysis will follow the lines of the arguments used in [13] where some of the arguments have to be refined due to the presence of corner singularities. However, (2.9) is not useful in practical computations. This is because the restriction to the controls to lie in I can not be imposed. Hence we will continue our exposition with the analysis of (2.8) based upon the results obtained during the analysis of (2.9).

Remark 2.1. *We remark, that the space I is characterized by a so called dual singular function s_{-1} which is known. However the characterization involves the unknown solution u of* (2.1). *In particular, a function $q \in I$ if and only if with the corresponding solution u to* (2.1) *it holds*

$$(q, s_{-1}) + (u, \Delta s_{-1}) = 0.$$

This representation is still of some use in order to calculate the singular coefficients and thereby accelerating convergence of the finite element method for the forward problem, see, e.g., [1]. In order to keep the presentation simple we will not follow such ideas to improve convergence of the discrete problems.

3. Analysis of the semi discretization

In this section we will analyze the error between (2.9) and (2.2) or (2.4) respectively.

To this end, we denote the unique solution to (2.2) or (2.4) by $(\overline{q}, \overline{u})$. The unique solutions to (2.9) will be denoted by $(\overline{q}_h^{\perp}, \overline{u}_h^{\perp})$.

Then similar to the proof of [13, Theorem 1] we obtain

Theorem 3.1. *Let $(\overline{q}, \overline{u}) \in Q \cap I \times W^{2,t}(\Omega) \cap V$ be the solution to* (2.2) *with $r \geq t > 2$. Further, let $(\overline{q}_h^{\perp}, \overline{u}_h^{\perp}) \in Q \cap I \times V_h$ be the solutions to* (2.9). *Then, for any $\varepsilon > 0$, there exists a constant $C > 0$ independent of $h \in (0, 1]$ such that*

$$|J(\overline{q}, \overline{u}) - J(\overline{q}_h^{\perp}, \overline{u}_h^{\perp})| \leq Ch^{\beta}$$

where $\beta = 1 - 2/t - \varepsilon$.

Proof. We begin our proof by considering the Ritz projection $u_h \in V_h$ of $\overline{u}$ defined by

$$(\nabla u_h, \nabla \varphi_h) = (\overline{q}, \varphi_h) \quad \forall \varphi_h \in V_h.$$

Then, because $\overline{u} \in W^{2,t}(\Omega) \subset W^{1,\infty}$ we have by [14, Theorem 2]

$$\|\nabla \overline{u} - \nabla u_h\|_{\infty} \leq ch^{\beta} \|\overline{u}\|_{C^{1,1-2/t}} \leq ch^{\beta} \|\overline{q}\|_{Q}. \tag{3.1}$$

Apart from this argument the rest of the proof is the same as in the case of a smooth domain, see [13, Theorem 1]. In particular, with $\tilde{c} \geq c\|q\|_Q$, we get that

$$(1 - \tilde{c}h^{\beta})|\nabla u_h| \leq (1 - \tilde{c}h^{\beta})|\nabla \overline{u}| + (1 - \tilde{c}h^{\beta})ch^{\beta}\|\overline{q}\|_Q \leq 1 \quad \text{a.e. in } \Omega.$$

From this we get that

$$(\tilde{q}_h, \tilde{u}_h) = (1 - \tilde{c}h^{\beta})(\overline{q}, u_h) \tag{3.2}$$

defines an element $(\tilde{q}_h, \tilde{u}_h) \in Q \cap I \times V_h$ which is feasible for (2.9). It is then clear, that

$$\|\overline{q} - \tilde{q}_h\|_Q + \|\overline{u} - \tilde{u}_h\|_2 \leq ch^{\beta}\|\overline{q}\|_Q.$$

Hence, we get

$$|J(\overline{q}, \overline{u}) - J(\tilde{q}_h, \tilde{u}_h)| \leq ch^{\beta}\|\overline{q}\|_Q$$

from local Lipschitz continuity of J. Furthermore we have

$$J(\overline{q}_h^\perp, \overline{u}_h^\perp) \le J(\tilde{q}_h, \tilde{u}_h)$$

because $(\tilde{q}_h, \tilde{u}_h)$ is feasible for (2.9). This yields

$$J(\overline{q}_h^\perp, \overline{u}_h^\perp) - J(\overline{q}, \overline{u}) \le J(\tilde{q}_h, \tilde{u}_h) - J(\overline{q}, \overline{u}) \le ch^\beta.$$

In particular the sequence $\|\overline{q}_h^\perp\|_Q^r \le 2J(\overline{q}_h^\perp, \overline{u}_h^\perp)$ is bounded.

In order to show the reverse inequality, i.e.,

$$-ch^\beta \le J(\overline{q}_h^\perp, \overline{u}_h^\perp) - J(\overline{q}, \overline{u})$$

we use the same line of arguments. We define for each given solution $(\overline{q}_h^\perp, \overline{u}_h^\perp) \in Q \cap I \times V_h$ to (2.9) a continuous function $u \in V$ using

$$(\nabla u, \nabla \varphi) = (\overline{q}_h^\perp, \varphi) \quad \forall \varphi \in V.$$

Due to the fact that $\overline{q}_h^\perp \in Q \cap I$ we have $u \in W^{2,t}(\Omega)$ and hence we get from [14, Theorem 2] that

$$\|\nabla \overline{u}_h^\perp - \nabla u\|_\infty \le ch^\beta \|u\|_{C^{1,1-2/t}} \le ch^\beta \|\overline{q}_h^\perp\|_Q$$

as in (3.1). Now one can continue analog by shifting to obtain a pair $(\hat{q}, \hat{u})$ which is feasible for (2.2) such that

$$|J(\overline{q}_h^\perp, \overline{u}_h^\perp) - J(\hat{q}, \hat{u})| \le ch^\beta.$$

Note that the constant c is independent of h because $\overline{q}_h^\perp$ can be bounded independent of h.

This yields the desired lower bound, i.e.,

$$-ch^\beta \le J(\overline{q}_h^\perp, \overline{u}_h^\perp) - J(\hat{q}, \hat{u}) \le J(\overline{q}_h^\perp, \overline{u}_h^\perp) - J(\overline{q}, \overline{u}) \le J(\tilde{q}_h, \tilde{u}_h) - J(\overline{q}, \overline{u}) \le ch^\beta$$

and concludes the proof. □

The convergence of the cost functional implies convergence of the primal variables.

Corollary 3.1. *Let $(\overline{q}, \overline{u}) \in Q \cap I \times W^{2,t}(\Omega) \cap V$ be the solution to (2.2) with $r \ge t > 2$. Further, let $(\overline{q}_h^\perp, \overline{u}_h^\perp) \in Q \cap I \times V_h$ be the solutions to (2.9). Then, for any $\varepsilon > 0$, there exists a constant $C > 0$ independent of $h \in (0, 1]$ such that*

$$\|\overline{q} - \overline{q}_h^\perp\|_Q^r + \|\overline{u} - \overline{u}_h^\perp\|^2 \le Ch^{1-2/t-\varepsilon}.$$

Proof. The proof is identical to the one for [13, Corollary 1]. □

4. Regularization

Before we come to the analysis of the error between (2.2) and (2.8), we will need some additional analysis. In particular, we are interested in the following regularized problems for given $\gamma > 0$

$$\underset{Q \times V}{\text{Minimize}}\ J_\gamma(q, u) := J(q, u) + \frac{\gamma}{2}\|(|\nabla u| - 1)^+\|^2,$$
$$\text{such that } (u, q) \text{ satisfies (2.1).}$$

Similar problems have been analyzed in [10]. Unfortunately their analysis was done under the assumption, that the state equation (2.1) defines an isomorphism between $W^{2,t}(\Omega) \cap V$ and $L^t(\Omega)$ which is not the case in our setting. Further we will require bounds on the rate of convergence of the primal variables similar to those obtained in [8]. Again the arguments are complicated by the fact, that the state equation does not yield sufficient regularity.

We note, that even though one can show convergence of the sequence of minimizers to to above problem to those of (2.2). However, in the above setting the convergence speed may be dominated by the existence of the corner singularities. As we know that they do not appear in the solution we will apply an additional filter to remove at least parts of the influence of the reentrant corner.

To do so we need to separate the influence of the corner singularities. Hence we define the set $I^\perp$ as

$$I^\perp = \{p \in Q^* = L^{p'}(\Omega) \,|\, (q,p) = 0 \; \forall \, q \in Q \cap I\}$$

where $p' = \frac{p}{p-1}$. The set $I^\perp$ is a finite dimensional linear space generated by so called dual singular functions. The dimension m of $I^\perp$ is equal to the number of non convex corners of the domain Ω, see [19]. In our case, this means by assumption $m = 1$.

Then we can define the finite dimensional linear space $Q_s \subset Q$ as follows

$$Q_s = \{q \in Q \,|\, \exists p \in I^\perp \colon \; (q,p) \neq 0\,\} \cup \{0\}.$$

This gives the following representation of Q

$$Q = Q \cap I \oplus Q_s.$$

Let $\{q^\perp\}$ be a basis of $I^\perp$. Then we choose $\{q^s\} \subset Q_s$ as dual basis to $\{q^\perp\}$, i.e., $(q^s, q^\perp) = 1$.

In particular, we can write any element $q \in Q$ as

$$q = q^r + \alpha q^s$$

where $q^r \in Q \cap I$ and $\alpha = (q, q^\perp) \in \mathbb{R}$ are uniquely determined. In particular $q \in Q \cap I$ if and only if $\alpha = 0$. Corresponding to this relation we can also rewrite any solution u to (2.1) with right hand side q as

$$u = u^r + \alpha u^s$$

where $u^r \in W^{2,t}(\Omega) \cap V$ and u^s behaves as $r^{\pi/\omega}$ in the vicinity of the reentrant corner.

Then we can state the regularized problem as follows

$$\begin{aligned} &\underset{Q\times V}{\text{Minimize}}\; J_\gamma(q,u) := J(q,u) + \frac{\gamma}{2}\|(|\nabla u| - 1)^+\|^2 + \frac{\gamma}{2}|(q,q^\perp)|^2, \\ &\qquad\qquad \text{such that } (u,q) \text{ satisfies (2.1).} \end{aligned} \tag{4.1}$$

We note, that by standard arguments, there exists unique solutions $(\overline{q}_\gamma, \overline{u}_\gamma) \in Q \times V$ to (4.1). Further, let $(\overline{q}_\gamma, \overline{u}_\gamma) \in Q \times V$ be the solution to (2.2). Due to the fact, that

$$J(\overline{q}_\gamma, \overline{u}_\gamma) \le J_\gamma(\overline{q}_\gamma, \overline{u}_\gamma) \le J_\gamma(\overline{q}, \overline{u}) = J(\overline{q}, \overline{u})$$

we immediately obtain boundedness of $\|\overline{q}_\gamma\|_Q$. Further, this gives the relation

$$\|(\nabla \overline{u}_\gamma| - 1)^+\|^2 \le C\gamma^{-1}, \qquad |\alpha_\gamma|^2 = |(\overline{q}_\gamma, q^\perp)|^2 \le C\gamma_2^{-1}. \tag{4.2}$$

Our analysis starts with an analysis of the feasibility error.

Lemma 4.1. *Let* $(\overline{q}_\gamma, \overline{u}_\gamma) \in Q \times V$ *be the solution to* (4.1)*. Further, denote* $\overline{q}_\gamma = q^r + \alpha_\gamma q^s \in Q \cap I \oplus Q_s$ *and* $\overline{u}_\gamma = u^r_\gamma + \alpha_\gamma u^s$ *the corresponding splitting of the state variable.*

Then there exists a constant c independent of γ such that

$$\|(|\nabla u^r_\gamma| - 1)^+\|_\infty \le c\gamma^{-1/2(\beta/(1+\beta))},$$

where β is the same as in Theorem 3.1*.*

Proof. To obtain the convergence of u^r_γ in the maximum norm, we define

$$f(x) = (|\nabla u^r_\gamma| - 1)^+.$$

We remark, that by embedding theorems, we know that $f \in C^{0,\beta}(\Omega)$. Then define

$$\varepsilon_\gamma = \max_{x \in \overline{\Omega}} f(x).$$

An easy computation shows that (for $\gamma \ge 1$)

$$\|(|\nabla u^r_\gamma| - 1)^+\|^2 \le c(\|(|\nabla \overline{u}_\gamma| - 1)^+\|^2 + |\alpha_\gamma|^2) \le c\gamma^{-1}.$$

We assume w.l.o.g. that $\varepsilon_\gamma > 0$. Then by Hölder continuity of f we get that

$$\begin{aligned} c\gamma^{-1} &\ge \|f\|^2 \\ &\ge \int_{\{f \ge \varepsilon_\gamma/2\}} |f(x)|^2 \, dx \\ &\ge \frac{\varepsilon_\gamma^2}{4} \int_{\{f \ge \varepsilon_\gamma/2\}} dx \\ &\ge \frac{\varepsilon_\gamma^2}{4} c\varepsilon_\gamma^{2/\beta} \\ &\ge c\varepsilon_\gamma^{2+2/\beta}. \end{aligned}$$

Hence by definition

$$\begin{aligned} \|(|\nabla u^r_\gamma| - 1)^+\|_\infty &= \varepsilon_\gamma \\ &\le c\gamma^{-\beta/(2\beta+2)} \end{aligned}$$

which shows the assertion. □

The rate of convergence of $\|(|\nabla u^r_\gamma| - 1)^+\|_\infty$ is the same which could be obtained following the analysis of [8, Lemma 3.1]. Unfortunately this rate also limits our ability to derive convergence estimates for the primal variables. Hence we will spend some effort on improving these results, the techniques employed here have been developed simultaneously in [11]. We will derive them here nonetheless because we will have to face some additional difficulties due to the presence of the corner singularities.

Before doing so, we recall that for a solution $(\overline{q},\overline{u}) \in Q\times V$ to (2.2) there exist $\overline{z} \in L^{t'}(\Omega)$ and $\overline{\mu} \in C^*(\overline{\Omega})$ such that the following necessary optimality conditions hold

$$
\begin{aligned}
(\nabla\overline{u},\nabla\varphi) &= (\overline{q},\varphi) && \forall\ \varphi\in V,\\
(-\Delta\varphi,\overline{z}) &= (\overline{u}-u^d,\varphi)+\langle\overline{\mu},\nabla\overline{u}\cdot\nabla\varphi\rangle\rangle_{C^*\times C} && \forall\ \varphi\in W^{2,t}(\Omega)\cap V,\\
(|\overline{q}|^{r-2}\overline{q},\delta q) &= -(\delta q,\overline{z}) && \forall\,\delta q\in Q\cap I,\\
\langle\overline{\mu},\varphi\rangle_{C^*\times C} &\le 0 && \forall\ \varphi\in C(\overline{\Omega}),\varphi\le 0,\\
\langle\overline{\mu},|\nabla\overline{u}|-1\rangle_{C^*\times C} &= 0,
\end{aligned}
\tag{4.3}
$$

see, [19, Theorem 3.3]. Further, by standard arguments for a solution $(\overline{q}_\gamma,\overline{u}_\gamma) \in Q\times V$ to (4.1) there exist $\overline{z}_\gamma \in V$ and $\overline{\mu}_\gamma \in L^2(\Omega)$ such that the following holds

$$
\begin{aligned}
(\nabla\overline{u}_\gamma,\nabla\varphi) &= (\overline{q}_\gamma,\varphi) && \forall\ \varphi\in V,\\
(\nabla\varphi,\nabla\overline{z}_\gamma) &= (\overline{u}_\gamma-u^d,\varphi)+(\overline{\mu}_\gamma,\nabla\overline{u}_\gamma\cdot\nabla\varphi) && \forall\ \varphi\in V,\\
(|\overline{q}_\gamma|^{r-2}\overline{q}_\gamma,\delta q) &= -(\delta q,\overline{z}_\gamma)-\gamma\alpha_\gamma(\delta q,q^\perp) && \forall\,\delta q\in Q,\\
\overline{\mu}_\gamma &= \frac{\gamma}{|\nabla\overline{u}_\gamma|}(|\nabla\overline{u}_\gamma|-1)^+
\end{aligned}
\tag{4.4}
$$

compare [10].

Lemma 4.2. *Let $(\overline{q}_\gamma,\overline{u}_\gamma) \in Q\times V$ be the solution to (4.1).*

Then there exists a constant c independent of γ such that it holds

$$\gamma\|(|\nabla u_\gamma|-1)^+\|_1 \le c, \quad \gamma|\alpha_\gamma| \le c.$$

Proof. We obtain for (4.4) that

$$
\begin{aligned}
\|\overline{q}_\gamma\|_Q^r+\gamma|\alpha_\gamma|^2 &= -(\overline{q}_\gamma,\overline{z}_\gamma)\\
&= (-\nabla\overline{u}_\gamma,\nabla\overline{z}_\gamma)\\
&= -(\overline{u}_\gamma-u^d,\overline{u}_\gamma)-\gamma((|\nabla\overline{u}_\gamma|-1)^+,|\nabla\overline{u}_\gamma|).
\end{aligned}
$$

Now we obtain that $\|\overline{q}_\gamma\|_Q$, $\|\overline{u}_\gamma\|$, and $\gamma|\alpha_\gamma|^2$ are bounded independent of γ because

$$0 \le J(\overline{q}_\gamma,\overline{u}_\gamma) \le J_\gamma(\overline{q}_\gamma,\overline{u}_\gamma) \le J(\overline{q},\overline{u})$$

Hence we have

$$
\begin{aligned}
\gamma\|(|\nabla\overline{u}_\gamma|-1)^+\|_1 &\le \gamma((|\nabla\overline{u}_\gamma|-1)^+,|\nabla\overline{u}_\gamma|)\\
&= -\|\overline{q}_\gamma\|_Q^r-(\overline{u}_\gamma-u^d,\overline{u}_\gamma)-\gamma|\alpha_\gamma|^2\\
&\le c.
\end{aligned}
$$

To get the bound on $\gamma|\alpha_\gamma|$ we test (4.4) with q^s and get

$$\begin{aligned}\gamma|\alpha_\gamma| &= \gamma|\alpha_\gamma(q^s, q^\perp)| \\ &= |(|\overline{q}_\gamma|^{r-2}\overline{q}_\gamma, q^s) + (\overline{u}_\gamma - u^d, u^s) + (\overline{\mu}_\gamma, \nabla\overline{u}_\gamma\nabla u^s)| \\ &\le c\end{aligned}$$

by noting that

$$|(\overline{\mu}_\gamma, \nabla\overline{u}_\gamma\nabla u^s)| \le \gamma\|(|\nabla u| - 1)^+\|\|\nabla u^s\| \le c.$$

□

With these preparations we can derive an improved L^∞ estimate.

Lemma 4.3. *Let* $(\overline{q}_\gamma, \overline{u}_\gamma) \in Q \times V$ *be the solution to* (4.1)*. Further, denote* $\overline{q}_\gamma = q^r + \alpha_\gamma q^s \in Q \cap I \oplus Q_s$ *and* $\overline{u}_\gamma = u^r_\gamma + \alpha_\gamma u^s$ *the corresponding splitting of the state variable. Then there exists a constant c independent of* $\gamma \ge 1$ *such that*

$$\|(|\nabla u^r_\gamma| - 1)^+\|_\infty \le c\gamma^{-(\beta/(\beta+2))},$$

where β *is the same as in Theorem* 3.1*.*

Proof. The proof is analog to the one for Lemma 4.1.

To obtain the convergence of u^r_γ in the maximum norm, we define

$$f(x) = (|\nabla u^r_\gamma| - 1)^+.$$

We remark, that by embedding theorems, we know that $f \in C^{0,\beta}(\Omega)$. Then define

$$\varepsilon_\gamma = \max_{x\in\overline{\Omega}} f(x).$$

An easy computation shows that

$$|(|\nabla u^r_\gamma| - 1)^+| \le |(|\nabla\overline{u}_\gamma| - 1)^+| + |\alpha_\gamma||\nabla u^s|$$

and hence by Lemma 4.2

$$\|(|\nabla u^r_\gamma| - 1)^+\|_1 \le \|(|\nabla u_\gamma| - 1)^+\|_1 + |\alpha_\gamma|\|\nabla u^s\|_1 \le c\gamma^{-1}.$$

We assume w.l.o.g. that $\varepsilon_\gamma > 0$. Then by Hölder continuity of f we get that

$$|f(x) - f(y)| \le c\|x - y\|^\beta$$

and hence if for some $x^* \in \overline{\Omega}$ it holds $f(x^*) = \varepsilon_\gamma = \max_{x\in\overline{\Omega}} f(x)$ then we have that $f(y) \ge \varepsilon_\gamma/2$ if $c\|x - y\|^\beta \le \varepsilon_\gamma/2$. This gives

$$\begin{aligned}c\gamma^{-1} &\ge \|f\|_1 \\ &\ge \int_{\{f\ge\varepsilon_\gamma/2\}} |f(x)|\,dx \\ &\ge \frac{\varepsilon_\gamma}{2}\int_{\{f\ge\varepsilon_\gamma/2\}} dx \qquad (4.5) \\ &\ge \frac{\varepsilon_\gamma}{2} c\varepsilon_\gamma^{2/\beta} \\ &\ge c\varepsilon_\gamma^{1+2/\beta}.\end{aligned}$$

Hence by definition

$$\begin{aligned}\|(|\nabla u_\gamma^r| - 1)^+\|_\infty &= \varepsilon_\gamma \\ &\le c\gamma^{-(\beta/(\beta+2))}\end{aligned}$$

which shows the assertion. □

Remark 4.1. *We remark, that usually the estimate in* (4.5) *is too pessimistic. For instance, if* $\overline{\mu}$ *has support on a curve in* $\overline{\Omega}$ *it is reasonable to assume that in fact*

$$\int_{\{f \ge \varepsilon_\gamma/2\}} dx \ge c\varepsilon_\gamma^{1/\beta}$$

yielding the improved rate

$$\|(|\nabla u_\gamma^r| - 1)^+\|_\infty \le c\gamma^{\frac{-\beta}{\beta+1}}.$$

Moreover, if $\overline{\mu}$ *has a volume contribution, then the set on which the maximum is attained may even be independent of the Hölder continuity, i.e.,*

$$\int_{\{f \ge \varepsilon_\gamma/2\}} dx \ge c$$

then yielding the rate

$$\|(|\nabla u_\gamma^r| - 1)^+\|_\infty \le c\gamma^{-1}.$$

For more details we refer to the forthcoming publication [11].

We remark that based upon these preparations one can derive estimates for the primal variables following the ideas of [8, Theorem 2.1] with some modifications due to the presence of corner singularities.

Lemma 4.4. *Let* $(\overline{q}_\gamma, \overline{u}_\gamma) \in Q \times V$ *be the solution to* (4.1) *and* $(\overline{q}, \overline{u}) \in Q \times V$ *be the solution to* (2.2)*. Further, denote* $\overline{q}_\gamma = q^r + \alpha_\gamma q^s \in Q \cap I \oplus Q_s$ *and* $\overline{u}_\gamma = u_\gamma^r + \alpha_\gamma u^s$ *the corresponding splitting of the state variable.*

Then, the following estimate holds:

$$\begin{aligned}\|\overline{q}_\gamma - \overline{q}\|_Q^r + \|u_\gamma^r - \overline{u}\|^2 + \frac{\gamma}{2}\|(|\nabla \overline{u}_\gamma|^2 - 1)^+\|^2 \\ &\le \langle (|\nabla u_\gamma^r|^2 - 1)^+, \overline{\mu}\rangle_{C,C^*} + |\alpha_\gamma|\Big(|(\overline{u} - u^d, u^s)| + |(|\overline{q}|^{r-2}\overline{q}, q^s)|\Big) \\ &\le C(\|(|\nabla u_\gamma^r| - 1)^+\|_\infty + |\alpha_\gamma|).\end{aligned}$$

Proof. First we remark, that for any r there exist a constant $c > 0$ such that

$$c\|f - g\|_{L^r}^r \le (|f|^{r-2}f - |g|^{r-2}g, f - g)$$

holds for any $f, g \in L^r(\Omega)$.

This gives in combination with the necessary optimality conditions (4.3) and (4.4)

$$\begin{aligned}c\|\overline{q}_\gamma - \overline{q}\|_Q^r &\le (|\overline{q}_\gamma|^{r-2}\overline{q}_\gamma - |\overline{q}|^{r-2}\overline{q}, \overline{q}_\gamma - \overline{q}) \\ &= -(\overline{z}_\gamma, \overline{q}_\gamma - \overline{q}) + (\overline{z}, q_\gamma^r - q) - \alpha_\gamma(|\overline{q}|^{r-2}\overline{q}, q^s) - \gamma^2\alpha_\gamma(\overline{q}_\gamma - \overline{q}, q^\perp) \\ &\le -(\overline{z}_\gamma, \overline{q}_\gamma - \overline{q}) + (\overline{z}, q_\gamma^r - q) - \alpha_\gamma(|\overline{q}|^{r-2}\overline{q}, q^s) + \gamma^2\alpha_\gamma(\overline{q}, q^\perp).\end{aligned}$$

Now noting that $(\overline{q}, q^{\perp}) = 0$ we conclude with the necessary optimality conditions (4.3) and (4.4) that

$$\begin{aligned}
c\|\overline{q}_\gamma - \overline{q}\|_Q^r &\le -(\overline{z}_\gamma, \overline{q}_\gamma - \overline{q}) + (\overline{z}, q_\gamma^r - q) - \alpha_\gamma(|\overline{q}|^{r-2}\overline{q}, q^s)\\
&= (\overline{z}_\gamma, \Delta\overline{u}_\gamma - \Delta\overline{u}) - (\overline{z}, \Delta u_\gamma^r - \Delta\overline{u}) - \alpha_\gamma(|\overline{q}|^{r-2}\overline{q}, q^s)\\
&= -(\overline{u}_\gamma - u^d, \overline{u}_\gamma - \overline{u}) - \gamma(\nabla\overline{u}_\gamma/|\nabla\overline{u}_\gamma|(|\nabla\overline{u}_\gamma| - 1)^+, \nabla\overline{u}_\gamma - \nabla\overline{u})\\
&\quad + (\overline{u} - u^d, u_\gamma^r - \overline{u}) + \langle\overline{\mu}, \nabla\overline{u}(\nabla u_\gamma^r - \nabla\overline{u})\rangle - \alpha_\gamma(|\overline{q}|^{r-2}\overline{q}, q^s)\\
&= -\|\overline{u}_\gamma - \overline{u}\|^2 - \alpha_\gamma(\overline{u} - u^d, u^s) - \alpha_\gamma(|\overline{q}|^{r-2}\overline{q}, q^s)\\
&\quad - \gamma(\nabla\overline{u}_\gamma/|\nabla\overline{u}_\gamma|(|\nabla\overline{u}_\gamma| - 1)^+, \nabla\overline{u}_\gamma - \nabla\overline{u}) + \langle\overline{\mu}, \nabla\overline{u}(\nabla u_\gamma^r - \nabla\overline{u})\rangle.
\end{aligned}$$

To proceed we need to rewrite the last two summands on the right hand side. To do so, we note that both $\gamma(|\nabla\overline{u}_\gamma| - 1)^+$ and μ are positive, and hence it is sufficient to estimate the arguments. This yields

$$\begin{aligned}
-\nabla\overline{u}_\gamma(\nabla\overline{u}_\gamma - \nabla\overline{u})/|\nabla\overline{u}_\gamma| &= (-|\nabla\overline{u}_\gamma|^2 + \nabla\overline{u}_\gamma\nabla\overline{u})/|\nabla\overline{u}_\gamma|\\
&\le (-|\nabla\overline{u}_\gamma|^2 + \frac{1}{2}(|\nabla\overline{u}_\gamma|^2 + |\nabla\overline{u}|^2))/|\nabla\overline{u}_\gamma|\\
&= \frac{1}{2}(|\nabla\overline{u}|^2 - |\nabla\overline{u}_\gamma|^2)/|\nabla\overline{u}_\gamma|\\
&\le \frac{1}{2}(1 - |\nabla\overline{u}_\gamma|^2)/|\nabla\overline{u}_\gamma|\\
&\le \frac{1}{2}(1/|\nabla\overline{u}_\gamma| - |\nabla\overline{u}_\gamma|).
\end{aligned}$$

Now, noting that $|\nabla\overline{u}_\gamma|^{-1} < 1$ on the set $\{|\nabla\overline{u}_\gamma| - 1 > 0\}$ we conclude that on this set

$$-\nabla\overline{u}_\gamma(\nabla\overline{u}_\gamma - \nabla\overline{u})/|\nabla\overline{u}_\gamma| \le \frac{1}{2}(1 - |\nabla\overline{u}_\gamma|).$$

Similarly one gets

$$\nabla\overline{u}(\nabla u_\gamma^r - \nabla\overline{u}) \le \frac{1}{2}(|\nabla u_\gamma^r|^2 - 1)^+.$$

Hence the first of the inequalities follows.

The second of the inequalities follows immediately by noting, that $\overline{u}$, $\overline{q}$, u^s, and q^s are independent of γ and that

$$|\nabla u_\gamma^r|^2 - 1 = (|\nabla u_\gamma^r| + 1)(|\nabla u_\gamma^r| - 1) \le c(|\nabla u_\gamma^r| - 1).$$

□

We note that Lemma 4.4 combined with Lemma 4.3 immediately gives a bound on the convergence of the primal variables. Additionally one could use the estimate of Lemma 4.1 in a bootstrapping argument to obtain better convergence orders than those derived there. However the results obtained when following this argument are not better than what we obtained in Lemma 4.3. We obtain the following convergence result.

Corollary 4.1. *Let* $(\overline{q}_\gamma, \overline{u}_\gamma) \in Q \times V$ *be the solution to* (4.1) *and* $(\overline{q}, \overline{u}) \in Q \times V$ *be the solution to* (2.2).

Then the following estimate holds

$$\|\overline{q}_\gamma - \overline{q}\|_Q^r + \|\overline{u}_\gamma - \overline{u}\|^2 \leq c\gamma^{\frac{-\beta}{\beta+2}}.$$

4.1. Convergence rates for the cost functional

Unfortunately for our later analysis we will require rates of convergence for the cost functionals. Clearly we get from local Lipschitz continuity of J in combination with Corollary 4.1 that

$$|J(\overline{q}_\gamma, \overline{u}_\gamma) - J(\overline{q}, \overline{u})| \leq c\gamma^{\frac{-\beta}{r(\beta+2)}}.$$

However, as we want to use the difference of the cost functionals to bound the error in the primal variables this is not sufficient. Therefore we will spend some additional effort on the derivation of convergence rates of the cost functional.

Theorem 4.1. *Let* $(\overline{q}_\gamma, \overline{u}_\gamma) \in Q \times V$ *be the solution to* (4.1) *and* $(\overline{q}, \overline{u}) \in Q \times V$ *be the solution to* (2.2)*. Further, denote* $\overline{q}_\gamma = q^r + \alpha_\gamma q^s \in Q \cap I \oplus Q_s$ *and* $\overline{u}_\gamma = u_\gamma^r + \alpha_\gamma u^s$ *the corresponding splitting of the state variable.*

Assume that

$$\|(|\nabla u_\gamma^r| - 1)^+\|_\infty \leq c\gamma^{-\theta}$$

then

$$0 \leq J(\overline{q}, \overline{u}) - J(\overline{q}_\gamma, \overline{u}_\gamma) \leq c\gamma^{-\theta}.$$

Proof. Assume that $(1 - c\gamma^{-\theta}) > 0$. Define $\tilde{q}_\gamma = (1 - c\gamma^{-\theta})q_\gamma^r$. Now, denote the corresponding solution to (2.1) by $\tilde{u}_\gamma$. Then it holds by assumption that

$$|\nabla \tilde{u}_\gamma| = (1 - c\gamma^{-\theta})|\nabla u_\gamma^r| \leq (1 - c\gamma^{-\theta})(1 + c\gamma^{-\theta}) < 1.$$

In particular $(\tilde{q}_\gamma, \tilde{u}_\gamma)$ is feasible for (2.2) and hence by local Lipschitz-continuity of J it follows

$$J(\overline{q}, \overline{u}) \leq J(\tilde{q}_\gamma, \tilde{u}_\gamma) \leq J(\overline{q}_\gamma, \overline{u}_\gamma) + c\gamma^{-\theta}.$$

□

Finally, we remark that a uniform convexity property holds for the function J_γ.

Lemma 4.5. *The functional* J_γ *is uniformly convex in the sense that*

$$\tfrac{1}{2}\|u_1 - u_2\|^2 + \tfrac{1}{r}\|q_1 - q_2\|_{L^r}^r + J_\gamma(\tfrac{1}{2}(w_1 + w_3)) \leq \tfrac{1}{2}J_\gamma(w_1) + \tfrac{1}{2}J_\gamma(w_2)$$

holds for all $w_1 = (q_1, u_1) \in Q \times V$ *and* $w_2 = (q_2, u_2) \in Q \times V$.

Proof. We note that the stated uniform convexity holds for the cost functional J by application of Clarkson's inequality to $u_1 - u^d$ and $u_2 - u^d$ as well as q_1 and q_2. Hence it remains to show that

$$\|(\tfrac{1}{2}|\nabla u_1 + \nabla u_2| - 1)^+\|^2 \leq \tfrac{1}{2}\|(|\nabla u_1| - 1)^+\|^2 + \tfrac{1}{2}\|(|\nabla u_2| - 1)^+\|^2.$$

This is clear, as the integral is monotone, the map $x \mapsto \max(0,x)^2 \colon \mathbb{R} \to \mathbb{R}$ is monotone increasing and $x \to |x| - 1 \colon \mathbb{R}^2 \to \mathbb{R}$ is convex. Similarly we have

$$\tfrac{1}{4}|(q_1 + q_2, q^\perp)|^2 \le \tfrac{1}{2}\Big(|(q_1, q^\perp)|^2 + |(q_2, q^\perp)|^2\Big).$$

□

Then combination of Theorem 4.1 and Lemma 4.5 yield the same rates of convergence that we obtained in Corollary 4.1.

Corollary 4.2. *Let* $(\overline{q}_\gamma, \overline{u}_\gamma) \in Q \times V$ *be the solution to* (4.1) *and* $(\overline{q}, \overline{u}) \in Q \times V$ *be the solution to* (2.2).

Then the following estimate holds

$$\|\overline{q}_\gamma - \overline{q}\|_Q^r + \|\overline{u}_\gamma - \overline{u}\|^2 \le c\gamma^{\frac{-\beta}{\beta+2}}.$$

Proof. By Lemma 4.5 we obtain

$$\begin{aligned}
\tfrac{1}{2}\|\overline{u}_\gamma - \overline{u}\|^2 + \tfrac{1}{r}\|\overline{q}_\gamma - \overline{q}\|_{L^r}^r &\le -J_\gamma(\tfrac{1}{2}(\overline{q}_\gamma + \overline{q}, \overline{u}_\gamma + \overline{u})) + \tfrac{1}{2}J_\gamma(\overline{q}_\gamma, \overline{u}_\gamma) + \tfrac{1}{2}J_\gamma(\overline{q}, \overline{u})\\
&\le \tfrac{1}{2}J_\gamma(\overline{q}, \overline{u}) - \tfrac{1}{2}J_\gamma(\overline{q}_\gamma, \overline{u}_\gamma)\\
&\le \tfrac{1}{2}J(\overline{q}, \overline{u}) - \tfrac{1}{2}J(\overline{q}_\gamma, \overline{u}_\gamma).
\end{aligned}$$

This shows the assertion using of Theorem 4.1 and Lemma 4.3. □

Remark 4.2. *We comment shortly on the influence of Remark* 4.1*. Given the comment there the speed of convergence in both Theorem* 4.1 *as well as in Corollary* 4.2 *will enhance to*

$$J(\overline{q}, \overline{u}) - J(\overline{q}_\gamma, \overline{u}_\gamma) \le c\gamma^{\frac{-\beta}{\beta+1}}$$

in the presence of a line measure in $\overline{\mu}$ *and*

$$J(\overline{q}, \overline{u}) - J(\overline{q}_\gamma, \overline{u}_\gamma) \le c\gamma^{-1}$$

in the presence of a volume measure in $\overline{\mu}$*.*

5. Analysis of the full discretization

In this section we will analyze the error between (2.8) and (2.2) or (2.4) respectively.

To this end, we denote the unique solution to (2.2) or (2.4) by $(\overline{q}, \overline{u})$. The unique solutions to (2.8) will be denoted by $(\overline{q}_h, \overline{u}_h)$.

In contrast to the previous section we can no longer consider the solution $u \in V$ of

$$(\nabla u, \nabla \varphi) = (\overline{q}_h, \varphi) \quad \forall \varphi \in V$$

in order to show that the lower bound

$$-ch^\beta \le J(\overline{q}_h, \overline{u}_h) - J(\overline{q}, \overline{u})$$

holds true. This is because the solution u defined above is no longer an element of $W^{2,t}(\Omega)$. However, from Theorem 3.1 we immediately get

$$J(\overline{q}_h, \overline{u}_h) - J(\overline{q}, \overline{u}) \le J(\overline{q}_h^\perp, \overline{u}_h^\perp) - J(\overline{q}, \overline{u}) \le ch^\beta$$

because $Q \subset Q \cap I$. In particular, the solutions $\overline{q}_h$ are uniformly bounded.

Before we come to the analysis of the convergence speed, we will start with some preliminary results. First, we will show convergence $\overline{q}_h \to \overline{q}$ and $\overline{u}_h \to \overline{u}$. With these preparations we will compute the distance between $\overline{q}_h$ and I. Then finally, we can obtain the desired convergence rates.

Theorem 5.1. *Let $(\overline{q}_h, \overline{u}_h)$ be the unique solution to* (2.8) *and denote $\overline{q}_h = q_h^r + \alpha_h q_i^s$ with $\alpha_h = (\overline{q}_h, q^\perp)$. Then $\overline{q}_h \rightharpoonup \overline{q}$ in Q and $\overline{u}_h \to \overline{u}$ in $H_0^1(\Omega)$ where $(\overline{q}, \overline{u})$ are the unique solution to* (2.2). *In particular, it holds $\alpha_h \to 0$.*

Proof. As already remarked, we have from Theorem 3.1 and $Q \subset Q \cap I$ that

$$J(\overline{q}_h, \overline{u}_h) \le J(\overline{q}_h^\perp, \overline{u}_h^\perp) \le J(\overline{q}, \overline{u}) + ch^\beta.$$

This shows that $\|\overline{q}_h\|_Q$ is bounded. Hence there exists a weakly convergent subsequence, denoted again by $\overline{q}_h$, with limit q_0. Due to the compact embedding $L^2(\Omega) \subset H^{-1}(\Omega)$ a subsequence $\overline{q}_h$ converges strongly in $H^{-1}(\Omega)$ and hence $\overline{u}_h$ converges strongly in $H_0^1(\Omega)$ to a limit u_0. Now, for any $\varphi \in H_0^1(\Omega)$ there exists a sequence $\varphi_h \in V_h$ with $\varphi_h \to \varphi$ because $\bigcup_{h>0} V_h$ is dense in $H_0^1(\Omega)$. Thus we have

$$(\nabla u_0, \nabla \varphi) \leftarrow (\nabla \overline{u}_h, \nabla \varphi_h) = (\overline{q}_h, \varphi_h) \to (q_0, \varphi_h).$$

To proceed, we note that the sequence $|\nabla \overline{u}_h|$ converges strongly in L^2 and hence, again selecting a subsequence, pointwise almost everywhere. Now $\|\nabla \overline{u}_h\|_\infty \le 1$ which shows $\|\nabla u_0\|_\infty \le 1$.

In particular, (q_0, u_0) are feasible for (2.2). From weak lower semicontinuity of J we deduce

$$J(q_0, u_0) \le \liminf_{h\to 0} J(\overline{q}_h, \overline{u}_h) \le J(\overline{q}, \overline{u}).$$

This shows $q_0 = \overline{q}$ and $u_0 = \overline{u}$.

Finally, as $\overline{q}_h \rightharpoonup \overline{q} \in Q \cap I$ we obtain $\alpha_h = (\overline{q}_h, q^\perp) \to (\overline{q}, q^\perp) = 0$. □

In a next step we try to obtain a convergence rate for the singular coefficient α_h. To do so, we consider the following problems where q_h^r is given as in Theorem 5.1. We search $u^r, u^s \in V$ and $u_h^r, u_h^s \in V_h$ which solve

$$(\nabla u^r, \nabla \varphi) = (q_h^r, \varphi) \qquad \forall\, \varphi \in V \tag{5.1}$$

$$(\nabla u_h^r, \nabla \varphi_h) = (q_h^r, \varphi_h) \qquad \forall\, \varphi_h \in V_h \tag{5.2}$$

$$(\nabla u^s, \nabla \varphi) = (q^s, \varphi) \qquad \forall\, \varphi \in V \tag{5.3}$$

$$(\nabla u_h^s, \nabla \varphi_h) = (q^s, \varphi_h) \qquad \forall\, \varphi_h \in V_h. \tag{5.4}$$

Lemma 5.1. *Let $(\overline{q}_h, \overline{u}_h)$ be the unique solution to* (2.8) *and denote $\overline{q}_h = q_h^r + \alpha_h q_i^s$. Then there exists a constant C independent of h such that*

$$|\alpha_h| \|\nabla u_h^s\|_\infty \le C.$$

Proof. We begin by noting, that $u^r \in W^{2,t}(\Omega)$ by definition of q_h^r. In particular, due to [14, Theorem 2]

$$\|\nabla u^r - \nabla u_h^r\|_\infty \le ch^\beta \|q_h^r\|_Q \le ch^\beta \|\overline{q}_h\|_Q \le ch^\beta.$$

Hence $\|\nabla u_h^r\|_\infty \le C$ independent of $h \in (0,1]$.

This yields

$$\|\alpha_h \nabla u_h^s\|_\infty = \|\nabla \overline{u}_h - \nabla u_h^r\|_\infty \le 1 + C$$

and thus the assertion. $\square$

In a next step, we need to show that $\|\nabla u_h^s\|_\infty$ blows up with a certain rate. This appears to be clear, unfortunately the author could not find a citable source. This is why we need the following lemma.

Lemma 5.2. *Let $\omega > \pi$ be the angle of the non convex corner of Ω. Further, let $u_h^s \in V_h$ be given by (5.4). Then for any $\varepsilon > 0$ there exists a constant c such that for $h > 0$ sufficiently small it holds*

$$\|\nabla u_h^s\|_\infty \ge c h^{-1+\pi/\omega+\varepsilon}.$$

Proof. Denote the non convex corner by v. Let $s > 1$ be given. Then it is well known, that the solution $u^s \in V$ of (5.3) satisfies

$$\max_{\mathrm{dist}(v,x)=h^{1/s}} u^s(x) \ge c_1 h^{\frac{\pi}{s\omega}}$$

for some given constant $c_1 > 0$ and h sufficiently small.

By [15, Theorem 4.1] we now that for any $\varepsilon' > 0$ there exists some $c_2 > 0$ such that

$$\|u^s - u_h^s\|_\infty \le c_2 h^{\frac{\pi}{\omega}-\varepsilon'}.$$

Then for given s it holds

$$c_2 h^{\frac{s-1}{s}\frac{\pi}{\omega}-\varepsilon'} < c_1/2$$

provided that h is sufficiently small.

In particular it holds for any $x \in \Omega$

$$u_h^s(x) \ge u^s(x) - c_2 h^{\frac{\pi}{\omega}-\varepsilon'}.$$

Hence we have for h sufficiently small that

$$\begin{aligned}
\max_{\mathrm{dist}(v,x)=h^{1/s}} u_h^s(x) &\ge c_1 h^{\frac{\pi}{s\omega}} - c_2 h^{\frac{\pi}{\omega}-\varepsilon'} \\
&= h^{\frac{\pi}{s\omega}}\left(c_1 - c_2 h^{\frac{s-1}{s}\frac{\pi}{\omega}-\varepsilon'}\right) \\
&\ge \frac{c_1}{2} h^{\frac{\pi}{s\omega}}.
\end{aligned}$$

On the other hand $h(v) = 0$. And thus we have

$$\begin{aligned}
\max_\Omega |\nabla u_h^s(x)| &\ge \frac{\max_{\mathrm{dist}(v,x)=h^{1/s}} u_h^s(x) - u_h^s(v)}{h^{1/s}} \\
&\ge \frac{c_1}{2} h^{\frac{\pi}{s\omega}} h^{\frac{-1}{s}} \\
&= \frac{c_1}{2} h^{\frac{1}{s}(-1+\frac{\pi}{\omega})}.
\end{aligned}$$

Now, for given $\varepsilon > 0$ such that $-1 + \frac{\pi}{\omega} + \varepsilon < 0$ there exists some $s > 1$ such that

$$-1 + \frac{\pi}{\omega} + \varepsilon = \frac{1}{s}(-1 + \frac{\pi}{\omega}).$$

This proofs the assertion. $\square$

Corollary 5.1. *For any $\varepsilon > 0$ there exists a constant c such that for $h > 0$ sufficiently small the singular coefficients α_h satisfy*

$$|\alpha_h| \leq ch^{1-\pi/\omega-\varepsilon}$$

Proof. The assertion follows immediately from Lemma 5.2 and Lemma 5.1. $\square$

Lemma 5.3. *Let $(\overline{q}_h, \overline{u}_h)$ be the unique solution to (2.8). Define $u^h \in V$ as the solution to*

$$(\nabla u^h, \nabla\varphi) = (\overline{q}_h, \varphi) \quad \forall \varphi \in V.$$

Then it holds

$$\|(|\nabla u^h| - 1)^+\|^2 \leq ch^{2\pi/\omega}.$$

Proof. By definition and the fact, that $|\nabla\overline{u}_h| \leq 1$, we have for almost all $x \in \Omega$

$$\begin{aligned}
(|\nabla u^h(x)| - 1)^+ &= \max(0, |\nabla u^h(x)| - 1) \\
&\leq \max(0, |\nabla u^h(x) - \nabla\overline{u}_h| + |\nabla\overline{u}_h| - 1) \\
&\leq \max(0, |\nabla u^h(x) - \nabla\overline{u}_h|) \\
&= |\nabla u^h(x) - \nabla\overline{u}_h|.
\end{aligned}$$

Hence we get by standard finite error estimates

$$\|(|\nabla u^h(x)| - 1)^+\|^2 \leq \|\nabla u^h - \nabla\overline{u}_h\|^2 \leq ch^{2\pi/\omega}.$$

$\square$

Theorem 5.2. *Let $(\overline{q}_h, \overline{u}_h)$ be the unique solution to (2.8) and $(\overline{q}, \overline{u})$ be the unique solution to (2.2). Then for h sufficiently small, there exists a constant $c > 0$ such that*

$$|J(\overline{q}, \overline{u}) - J(\overline{q}_h, \overline{u}_h)| \leq ch^{\beta_2}$$

where

$$\beta_2 = \frac{\beta}{1+\beta}(1 - \pi/\omega - \varepsilon)$$

for any $\varepsilon > 0$.

Proof. In view of Theorem 3.1 we already know that

$$J(\overline{q}_h, \overline{u}_h) \leq J(\overline{q}, \overline{u}) + ch^{\beta}.$$

Hence it remains to derive a lower bound on $J(\overline{q}_h, \overline{u}_h)$. To this end, we define $u^h \in V$ by

$$(\nabla u^h, \nabla\varphi) = (\overline{q}_h, \varphi) \quad \forall \varphi \in V.$$

Now, by standard L^2-error estimates we have

$$\|u - u^h\| \leq ch^{2\pi/\omega}$$

and because $2\pi/\omega \geq 1 > \beta$ we get

$$J(\overline{q}_h, u^h) \leq J(\overline{q}, \overline{u}) + ch^\beta,$$

$$|J(\overline{q}_h, u^h) - J(\overline{q}_h, \overline{u}_h)| \leq ch^\beta. \tag{5.5}$$

From this we immediately see, that if

$$J(\overline{q}, \overline{u}) \leq J(\overline{q}_h, u^h)$$

we would be done.

Hence, we will now assume that

$$J(\overline{q}_h, u^h) \leq J(\overline{q}, \overline{u}).$$

Then we proceed by considering a regularized version of (2.2) namely (4.1). Now we have that

$$|J_\gamma(\overline{q}_\gamma, \overline{u}_\gamma) - J(\overline{q}, \overline{u})| \leq c\gamma^{\frac{-\beta}{\beta+2}}$$

following the result of Theorem 4.1 and Lemma 4.3.

Further, with respect to Lemma 5.3 and Corollary 5.1 we have that

$$|J(\overline{q}_h, u^h) - J_\gamma(\overline{q}_h, u^h)| \leq c\gamma h^{2\pi/\omega} + c\gamma h^{2(1-\pi/\omega-\varepsilon)}$$

and thus

$$\begin{aligned} J(\overline{q}, \overline{u}) - c\gamma^{\frac{-\beta}{2+\beta}} &\leq J_\gamma(\overline{q}_\gamma, \overline{u}_\gamma) \\ &\leq J_\gamma(\overline{q}_h, u^h) \leq J(\overline{q}_h, u^h) + c\gamma h^{2\pi/\omega} + c\gamma h^{2(1-\pi/\omega-\varepsilon)} \\ &\leq J(\overline{q}, \overline{u}) + c\gamma h^{2\pi/\omega} + c\gamma h^{2(1-\pi/\omega-\varepsilon)} + ch^\beta. \end{aligned}$$

Now, in order to obtain the best possible rate of convergence, we choose $\gamma = h^{-x}$ where $x \geq 0$ solves

$$\max_{x\geq 0} \min\Big(x\frac{\beta}{2+\beta}, 2\frac{\pi}{\omega} - x, 2(1-\pi/\omega-\varepsilon) - x\Big) = f^*. \tag{5.6}$$

To do so, we note that $2\pi/\omega > 2-2\pi/\omega-\varepsilon$ since $\omega \leq 2\pi$. Hence the minimizer is obtained when the two terms $x\frac{\beta}{2+\beta} =$ and $2(1-\frac{\pi}{\omega}-\varepsilon)-x$ are equilibrated. This happens at

$$\overline{x} = (1 - \frac{\pi}{\omega} - \varepsilon)\frac{2+\beta}{1+\beta}$$

with the value

$$f^* = \overline{x}\frac{\beta}{2+\beta} = \frac{\beta}{1+\beta}(1-\pi/\omega-\varepsilon) = \beta_2 < \beta.$$

Thus we obtain

$$J(\overline{q}, \overline{u}) - ch^{\beta_2} \leq J(\overline{q}_h, u^h) \leq J(\overline{q}, \overline{u}) + ch^{\beta_2}$$

which shows the assertion. □

Convergence of the primal variables follows analog to Corollary 4.2 using the uniform convexity of J_γ to get that

$$\|\overline{q}_\gamma - \overline{q}_h\|_Q^r + \|\overline{u}_\gamma - \overline{u}_h\|^2 \leq ch^{\beta_2}.$$

References

[1] H. Blum and M. Dobrowolski, *On finite element methods for elliptic equations on domains with corners*, Computing, 28 (1982), pp. 53–63.

[2] S. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer Verlag, New York, 3. ed., 2008.

[3] E. Casas and J. F. Bonnans, *Contrôle de systèmes elliptiques semilinéares comportant des contraintes sur l'état*, in Nonlinear Partial Differential Equations and their Applications 8, H. Brezzis and J. Lions, eds., Longman, New York, 1988, pp. 69–86.

[4] E. Casas and L. A. Fernández, *Optimal control of semilinear elliptic equations with pointwise constraints on the gradient of the state*, Appl. Math. Optim., 27 (1993), pp. 35–56.

[5] K. Deckelnick, A. Günther, and M. Hinze, *Finite element approximation of elliptic control problems with constraints on the gradient*, Numer. Math., 111 (2008), pp. 335–350.

[6] J. Douglas, Jr., T. Dupont, and L. Wahlbin, *The stability in L^q of the L^2-projection into finite element function spaces*, Numer. Math., 23 (1974/75), pp. 193–197.

[7] A. Günther and M. Hinze, *Elliptic control problems with gradient constraints - variational discrete versus piecewise constant controls*, Comput. Optim. Appl., 49 (2009), pp. 549–566.

[8] M. Hintermüller and M. Hinze, *Moreau-Yosida regularization in state constrained elliptic control problems: Error estimates and parameter adjustment*, SIAM J. Numer. Anal., 47 (2009), pp. 1666–1683.

[9] M. Hintermüller, M. Hinze, and R. H. W. Hoppe, *Weak-duality based adaptive finite element methods for PDE-constrained optimization with pointwise gradient state-constraints*, Preprint 2010-08, Hamburger Beiträge zur Angewandten Mathematik, 2010.

[10] M. Hintermüller and K. Kunisch, *PDE-constrained optimization subject to pointwise constraints on the control, the state, and its derivative*, SIAM J. Optim., 20 (2009), pp. 1133–1156.

[11] M. Hintermüller, A. Schiela, and W. Wollner, *The length of the primal-dual path in moreau-yosida-based path-following for state constrained optimal control*, Preprint To appear, Hamburger Beiträge zur Angewandten Mathematik, 2012.

[12] M. Hinze, *A variational discretization concept in control constrained optimization: The linear-quadratic case*, Comp. Optim. Appl., 30 (2005), pp. 45–61.

[13] C. Ortner and W. Wollner, *A priori error estimates for optimal control problems with pointwise constraints on the gradient of the state*, Numer. Math., 118 (2011), pp. 587–600.

[14] A. Schatz, *A weak discrete maximum principle and stability of the finite element method in L_∞ on plane polygonal domains. I*, Math. Comp., 33 (1980), pp. 77–91.

[15] A. Schatz and L. B. Wahlbin, *Maximum norm estimates in the finite element method on plane polygonal domains. Part 1*, Math. Comp., 32 (1978), pp. 73–109.

[16] A. Schiela and W. Wollner, *Barrier methods for optimal control problems with convex nonlinear gradient state constraints*, SIAM J. Optim., 21 (2011), pp. 269–286.

[17] M. Ulbrich, *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems*, MOS-SIAM Series on Optimization, Society for Industrial and Applied Mathematics (SIAM), 2011.

[18] W. Wollner, *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 (2010), pp. 133–159.

[19] W. Wollner, *Optimal control of elliptic equations with pointwise constraints on the gradient of the state in nonsmooth polygonal domains*, Preprint 2011-11, Hamburger Beiträge zur Angewandten Mathematik, 2011.

Winnifried Wollner
Department of Mathematics
University of Hamburg
Bundesstrasse 55
D-20146 Hamburg
Germany
e-mail: `winnifried.wollner@math.uni-hamburg.de`