Deutsche Forschungsgemeinschaft
Priority Programme 1962

An Optimal Control Problem for Equations with p-Structure and its Finite Element Discretization

Adrian Hirn, Winnifried Wollner

Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization

Preprint Number SPP1962-137r

received on October 20, 2020

Edited by
SPP1962 at Weierstrass Institute for Applied Analysis and Stochastics (WIAS)
Leibniz Institute in the Forschungsverbund Berlin e.V.
Mohrenstraße 39, 10117 Berlin, Germany
E-Mail: spp1962@wias-berlin.de

World Wide Web: http://spp1962.wias-berlin.de/

AN OPTIMAL CONTROL PROBLEM FOR EQUATIONS WITH p-STRUCTURE AND ITS FINITE ELEMENT DISCRETIZATION

ADRIAN HIRN AND WINNIFRIED WOLLNER

Abstract. We analyze a finite element approximation of an optimal control problem that involves an elliptic equation with p-structure (e.g., the p-Laplace) as a constraint. As the nonlinear operator related to the p-Laplace equation mapping the space $W_0^{1,p}(\Omega)$ to its dual $\big(W_0^{1,p}(\Omega)\big)^*$ is not Gâteaux differentiable, first order optimality conditions cannot be formulated in a standard way. Without using adjoint information, we derive novel a priori error estimates for the convergence of the cost functional for both variational discretization and piecewise constant controls.

1. Introduction

In this article, for given $\alpha > 0$ and $u_d \in L^2(\Omega)$ we study the finite element discretization of the following elliptic optimal control problem:

$$\text{(1.1a)} \qquad \text{Minimize} \qquad J(q,u) := \frac{1}{2}\|u-u_d\|_2^2 + \frac{\alpha}{2}\|q\|_2^2$$

subject to the PDE-constraint

$$\text{(1.1b)} \qquad -\operatorname{div} \boldsymbol{S}(\nabla u) = q \qquad \text{in } \Omega,$$

$$\text{(1.1c)} \qquad u = 0 \qquad \text{on } \partial\Omega,$$

and for $q_a, q_b \in \mathbb{R}$, $q_a < q_b$, and w.l.o.g. $0 \in [q_a, q_b]$, the box-constraints

$$\text{(1.1d)} \qquad q_a \le q(x) \le q_b \qquad \text{for a.a. } x \in \Omega,$$

where for given $p > 1$, $\varepsilon \ge 0$ the nonlinear vector field $\boldsymbol{S} : \mathbb{R}^d \to \mathbb{R}^d$ is supposed to have p-structure, or to be more precisely (p,ε)-structure, see Assumption 2.1. Prototypical examples falling into this class are

$$\text{(1.2)} \qquad \boldsymbol{S}(\nabla u) = |\nabla u|^{p-2}\nabla u \qquad \text{or} \qquad \boldsymbol{S}(\nabla u) = \big(\varepsilon^2 + |\nabla u|^2\big)^{\frac{p-2}{2}} \nabla u\,.$$

Throughout the paper, $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, is either a convex polyhedral domain or a bounded convex domain with smooth boundary $\partial\Omega \in C^2$. We include the case of curved boundary in our analysis as we need to assume certain conditions on the regularity of u that are so far only available for domains with smooth boundary. For simplicity of the exposition we will restrict the analysis to $d = 2$ in case of a curved boundary.

(Adrian Hirn) Hochschule Esslingen, Robert-Bosch-Strasse 1, 73037 Göppingen, Germany

(Winnifried Wollner) Technische Universität Darmstadt, Fachbereich Mathematik, Dolivostr. 15, 64293 Darmstadt, Germany.

E-mail addresses: adrian.hirn@hs-esslingen.de, wollner@mathematik.tu-darmstadt.de.

Date: October 20, 2020.

1991 *Mathematics Subject Classification.* 49K20, 49M25, 65N30, 65N12.

Key words and phrases. p-Laplacian, optimization, finite element method, optimality conditions, a priori error estimates.

Equations with p-structure arise in various physical applications, such as in the theory of plasticity, bimaterial problems in elastic-plastic mechanics, non-Newtonian fluid mechanics, blood rheology and glaciology, see, e.g., [25, 31, 34] and the references therein. The first operator in (1.2) corresponds to the p-Laplace equation. For $\varepsilon > 0$ the second operator in (1.2) regularizes the degeneracy of the p-Laplacian as the modulus of the gradient tends to zero for $p > 2$, to infinity for $p < 2$ respectively. For $\varepsilon = 0$ $(1.2)_2$ reduces to the p-Laplace operator $(1.2)_1$. Finite element (FE) approximations of the p-Laplace equation and related equations have been widely investigated, see [1, 21, 23, 29]. Attention to optimal control of quasi-linear PDEs is given, e.g., in [9, 11]. The extension to the parabolic case can be found in [6, 8]. The approximation of optimal control problems in the coefficient for the p-Laplacian, by its ε-regularization, is studied in [10]. For the problem (1.1) some DWR-type a posteriori error estimates have been utilized in [24]. To the best knowledge of the authors, no a priori discretization error results are available for this problem class; by our work we want to fill this gap. We point out that the case $\varepsilon = 0$ is included in our analysis.

A standard procedure for the finite element analysis of an optimal control problem consists in deriving first order optimality conditions and exploiting the properties of the adjoint state. Unfortunately, this analysis requires the existence of a suitable adjoint state. To see that this can not be guaranteed by standard theory, note that the nonlinear operator related to the p-Laplace equation maps the space $W_0^{1,p}(\Omega)$ to its dual $\big(W_0^{1,p}(\Omega)\big)^*$. Hence, its formal derivative would be a linear operator mapping $W_0^{1,p}(\Omega)$ into its dual. As it can be seen in the calculations yielding (5.7), the corresponding linear operator is positive, and thus injective, further it is clearly self-adjoint, cf., (5.2). Hence, unless $p = 2$, the linear operator can not be surjective, see, e.g., the discussion in [26]. Hence, standard KKT-theory is not applicable in the natural setting.

Despite this lack of standard theory, for $\varepsilon > 0$, we are able to show existence of an suitable discrete adjoint state allowing a discrete optimality system suitable for a variational discretization in the spirit of [27]. Due to lack of first order optimality conditions on the continuous level, we cannot attain additional regularity of the adjoint variable. Without additional regularity of these variables we cannot expect more than qualitative convergence for them. Hence, to establish a priori error estimates, we follow techniques established for elliptic optimization problems with state [17, 35] or gradient-state [36] constraints where also no convergence rates of the adjoint variable are available. Although in our analysis we can adopt ideas from [36], we have to cope with several challenges due to the nonlinear degenerate PDE-constraint. For discretization of (1.1) we consider two possible approaches: (a) variational discretization with piecewise linear states and (b) piecewise linear states and piecewise constant controls. In case of (a) the control space is discretized implicitly by the discrete adjoint equation. We show that the sequence of discrete, global, minimizers $(\overline{q}_h, \overline{u}_h)$ for mesh size $h \in (0, 1]$ has a strong accumulation point $(\overline{q}, \overline{u})$ that is a global optimal solution to (1.1). Under a certain realistic regularity assumption for solutions of the state equation (Assumption 2.4), we prove a quantitative convergence estimate for the cost functional value for both variational discretization and piecewise constant controls in Theorems 7.2 and 7.3.

For the proof of these estimates, we combine methods from [36] with quasi-norm techniques from [21] in order to handle the degeneracy of the nonlinear operator. Our method does not require additional regularity of the control variable. The required regularity in Assumption 2.4 is verified for the p-Laplace equation on bounded convex domains with C^2-boundary in [13, 16].

The plan of the paper is as follows: In Section 2, we fix our notation and we clarify the structure of the nonlinear vector field $\boldsymbol{S}$. Further, we state our assumption on the regularity of solutions to (1.1b), (1.1c) (Assumption 2.4). Section 3 is concerned with the precise formulation of the optimal control problem (1.1) and its solvability. In Section 4, we describe its finite element discretization, followed in Section 5 by an analysis of the first order optimality conditions. In Section 6, we collect and extend several results on the finite element approximation of the p-Laplace equation in order to apply them in Section 7 to the convergence analysis of the optimal control problem. There we verify without any regularity assumption that the sequence of discrete minimizers $(\overline{q}_h, \overline{u}_h)$ has a strong accumulation point $(\overline{q}, \overline{u})$ that is an optimal solution to (1.1). Under the regularity Assumption 2.4 we then prove a priori error estimates quantifying the order of convergence in the cost functional.

2. Preliminaries

To begin with, we clarify our notation and we state important properties of the nonlinear operator in (1.1b). Further, we pose our assumption on the regularity of solutions to the state equation (1.1b) & (1.1c) that will be crucial for our analysis.

2.1. **Notation.** The set of all positive real numbers is denoted by $\mathbb{R}^+$. Let $\mathbb{R}_0^+ := \mathbb{R}^+ \cup \{0\}$. The Euclidean scalar product of two vectors $\boldsymbol{\xi}$, $\boldsymbol{\eta} \in \mathbb{R}^d$ is denoted by $\boldsymbol{\xi} \cdot \boldsymbol{\eta}$. We set $|\boldsymbol{\eta}| := (\boldsymbol{\eta} \cdot \boldsymbol{\eta})^{1/2}$. Often we use c as a generic constant whose value may change from line to line but does not depend on important variables. We write $a \sim b$ if there exist constants c, $C > 0$ independent of all relevant quantities such that $cb \le a \le Cb$. Similarly, the notation $a \lesssim b$ stands for $a \le Cb$.

Let $\omega \subset \Omega$ be a measurable, nonempty set. The d-dimensional Lebesgue measure of ω is denoted by $|\omega|$. The mean value of a Lebesgue integrable function f over ω is denoted by

$$\langle f \rangle_\omega := \fint_\omega f(x)\,\mathrm{d}x := \frac{1}{|\omega|}\int_\omega f(x)\,\mathrm{d}x.$$

For $\nu \in [1,\infty]$, $L^\nu(\Omega)$ stands for the Lebesgue space and $W^{m,\nu}(\Omega)$ for the Sobolev space of order m. For $\nu > 1$ we use the notation $W_0^{1,\nu}(\Omega)$ for the Sobolev space with vanishing traces on $\partial\Omega$. The $L^\nu(\omega)$-norm is denoted by $\|\cdot\|_{\nu;\omega}$ and the $W^{m,\nu}(\omega)$-norm is denoted by $\|\cdot\|_{m,\nu;\omega}$. For $\nu \in (1,\infty)$ and $\frac{1}{\nu} + \frac{1}{\nu'} = 1$, i.e., $\nu' = \frac{\nu}{\nu-1}$, the dual space of $W_0^{1,\nu}(\omega)$ is denoted by $W^{-1,\nu'}(\omega) = \big(W_0^{1,\nu}(\omega)\big)^*$ and for its dual norm we write $\|\cdot\|_{-1,\nu';\omega}$. For the $L^2(\omega)$ inner product we use the notation $(\cdot,\cdot)_\omega := \int_\omega uv\,\mathrm{d}x$. This notation of norms and inner products is also used for vector-valued functions. In case of $\omega = \Omega$, we usually omit the index Ω, e.g., $\|\cdot\|_\nu = \|\cdot\|_{\nu;\Omega}$.

We recall the important Poincaré inequality: For $\nu \in (1,\infty)$ there holds

$$\|u\|_{\nu;\omega} \le c_P \|\nabla u\|_{\nu;\omega} \qquad \forall u \in W_0^{1,\nu}(\omega). \tag{2.1}$$

There exist diverse generalizations of Poincaré's inequality. We will make use of the following version that goes back to [5]: Let ω be a bounded convex open subset of $\mathbb{R}^d$, $d \ge 1$, and $\varphi : [0,\infty) \to [0,\infty)$ be continuous and convex with $\varphi(0) = 0$. Let $u : \omega \to \mathbb{R}^N$, $N \ge 1$, be in $W^{1,1}(\omega)$ such that $\varphi(|\nabla u|) \in L^1(\omega)$. Then there holds

$$\int_\omega \varphi\left(\frac{|u(x) - \langle u\rangle_\omega|}{\delta}\right) \mathrm{d}x \le \left(\frac{V_d \delta^d}{|\omega|}\right)^{1-\frac{1}{d}} \int_\omega \varphi(|\nabla u(x)|)\,\mathrm{d}x \tag{2.2}$$

where δ is the diameter of ω and V_d is the volume of the unit ball in $\mathbb{R}^d$.

2.2. **Properties of the nonlinear operator.** In this section we state our assumptions on the nonlinear operator $\boldsymbol{S}$. Further, we discuss important properties of the nonlinear operator and we indicate how it relates to so-called N-functions.

Assumption 2.1 (Nonlinear operator)**.** *We assume that the nonlinear operator* $\boldsymbol{S} : \mathbb{R}^d \to \mathbb{R}^d$ *belongs to* $C^0(\mathbb{R}^d, \mathbb{R}^d) \cap C^1(\mathbb{R}^d \setminus \{\mathbf{0}\}, \mathbb{R}^d)$ *and satisfies* $\boldsymbol{S}(\mathbf{0}) = \mathbf{0}$. *Furthermore, we assume that the operator* $\boldsymbol{S}$ *possesses* (p, ε)*-structure, i.e., there exist* $p \in (1, \infty)$, $\varepsilon \in [0, \infty)$, *and constants* $C_0, C_1 > 0$ *such that*

$$\sum_{i,j=1}^{d} \partial_i S_j(\boldsymbol{\xi})\eta_i\eta_j \geq C_0(\varepsilon + |\boldsymbol{\xi}|)^{p-2}|\boldsymbol{\eta}|^2, \tag{2.3a}$$

$$|\partial_i S_j(\boldsymbol{\xi})| \leq C_1(\varepsilon + |\boldsymbol{\xi}|)^{p-2} \tag{2.3b}$$

holds for all $\boldsymbol{\xi}$, $\boldsymbol{\eta} \in \mathbb{R}^d$ *with* $\boldsymbol{\xi} \neq \mathbf{0}$ *and all* $i, j \in \{1, \ldots, d\}$.

Important examples of nonlinear operators $\boldsymbol{S}$ satisfying Assumption 2.1 are those derived from a potential with (p, ε)-structure, i.e., there exists a convex function $\Phi : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ belonging to $C^1(\mathbb{R}_0^+) \cap C^2(\mathbb{R}^+)$ and satisfying $\Phi(0) = 0$, $\Phi'(0) = 0$ such that for all $\boldsymbol{\xi} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ and $i = 1, \ldots, d$ there holds

$$S_i(\boldsymbol{\xi}) = \partial_i \left(\Phi(|\boldsymbol{\xi}|)\right) = \Phi'(|\boldsymbol{\xi}|)\frac{\xi_i}{|\boldsymbol{\xi}|}\,. \tag{2.4}$$

If in addition Φ possesses (p, ε)-structure, i.e., if there exist $p \in (1, \infty)$, $\varepsilon \in [0, \infty)$ and constants $C_2, C_3 > 0$ such that for all $t > 0$ there holds

$$C_2(\varepsilon + t)^{p-2} \leq \Phi''(t) \leq C_3(\varepsilon + t)^{p-2}\,, \tag{2.5}$$

then one can show (see [4]) that $\boldsymbol{S}$ satisfies Assumption 2.1. Note that (1.2) falls into this class. We will briefly discuss how the operator $\boldsymbol{S}$ with (p, ε)-structure relates to N-functions that are standard in the theory of Orlicz spaces, cf., [19]. We define a convex function $\varphi : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ by

$$\varphi(t) := \int_0^t (\varepsilon + s)^{p-2} s \,\mathrm{d}s. \tag{2.6}$$

The function φ belongs to $C^1(\mathbb{R}_0^+) \cap C^2(\mathbb{R}^+)$ and satisfies uniformly in $t > 0$

$$\min\{1, p-1\}(\varepsilon + t)^{p-2} \leq \varphi''(t) \leq \max\{1, p-1\}(\varepsilon + t)^{p-2}\,. \tag{2.7}$$

Therefore, the inequalities (2.3a) and (2.3b) defining the (p, ε)-structure of $\boldsymbol{S}$ can be expressed equivalently in terms of the convex function φ:

$$\sum_{i,j=1}^{d} \partial_i S_j(\boldsymbol{\xi})\eta_i\eta_j \geq \widetilde{C}_0\varphi''(|\boldsymbol{\xi}|)|\boldsymbol{\eta}|^2, \qquad |\partial_i S_j(\boldsymbol{\xi})| \leq \widetilde{C}_1\varphi''(|\boldsymbol{\xi}|) \qquad \forall i, j = 1, \ldots, d\,.$$

The function φ is an example of an N-function satisfying the Δ_2-condition, see, e.g., [19,22]. In view of (2.7) the function φ satisfies uniformly in t the equivalence

$$\varphi''(t)t \sim \varphi'(t)\,. \tag{2.8}$$

Several studies on the finite element analysis of the p-Laplace equation indicate that a p-structure-adapted quasi-norm is crucial for error estimation. To this end, for given $\psi \in C^1([0, \infty))$ we introduce the family of shifted functions $\{\psi_a\}_{a \geq 0}$ by

$$\psi_a(t) := \int_0^t \psi_a'(s)\,\mathrm{d}s \qquad \text{with} \qquad \psi_a'(t) := \frac{\psi'(a+t)}{a+t}t\,. \tag{2.9}$$

For $\psi = \varphi$ given by (2.6), there holds $\varphi_a(t) \sim (\varepsilon + a + t)^{p-2} t^2$ uniformly in $t \geq 0$. [19] provides the following inequality of Young-type: For all $\delta > 0$ there exists $c(\delta) > 0$ such that for all s, t, $a \geq 0$ there holds

$$s\varphi_a'(t) + \varphi_a'(s)t \leq \delta\varphi_a(s) + c(\delta)\varphi_a(t)\,, \tag{2.10}$$

where the constant $c(\delta)$ only depends on p and δ (it is independent of ε, a).

We define a function $\boldsymbol{F} : \mathbb{R}^d \to \mathbb{R}^d$ associated to the nonlinear operator $\boldsymbol{S}$ with (p, ε)-structure by

$$\boldsymbol{F}(\boldsymbol{\xi}) := (\varepsilon + |\boldsymbol{\xi}|)^{\frac{p-2}{2}}\,\boldsymbol{\xi} \tag{2.11}$$

where p and ε are the same as in Assumption 2.1. The vector fields $\boldsymbol{S}$ and $\boldsymbol{F}$ are closely related to each other as depicted by the following lemma provided by [18,19].

Lemma 2.2. *For $p \in (1,\infty)$, $\varepsilon \in [0,\infty)$ let $\boldsymbol{S}$ satisfy Assumption 2.1, let $\boldsymbol{F}$, φ, and $\varphi_{|\boldsymbol{\xi}|}$ be defined by* (2.11), (2.6), *and* (2.9) *respectively. Then there holds for all $\boldsymbol{\xi}$, $\boldsymbol{\eta} \in \mathbb{R}^d$*

$$\begin{aligned}\big(\boldsymbol{S}(\boldsymbol{\xi}) - \boldsymbol{S}(\boldsymbol{\eta})\big)\cdot(\boldsymbol{\xi}-\boldsymbol{\eta}) &\sim (\varepsilon + |\boldsymbol{\xi}| + |\boldsymbol{\eta}|)^{p-2}|\boldsymbol{\xi}-\boldsymbol{\eta}|^2 \\ &\sim \varphi_{|\boldsymbol{\xi}|}(|\boldsymbol{\xi}-\boldsymbol{\eta}|) \sim |\boldsymbol{F}(\boldsymbol{\xi}) - \boldsymbol{F}(\boldsymbol{\eta})|^2, \\ |\boldsymbol{S}(\boldsymbol{\xi}) - \boldsymbol{S}(\boldsymbol{\eta})| &\sim \varphi_{|\boldsymbol{\xi}|}'(|\boldsymbol{\xi}-\boldsymbol{\eta}|) \sim (\varepsilon + |\boldsymbol{\xi}| + |\boldsymbol{\eta}|)^{p-2}|\boldsymbol{\xi}-\boldsymbol{\eta}|,\end{aligned}$$

where the constants only depend on p; they are independent of $\varepsilon \geq 0$ in particular.

Due to Lemma 2.2, for all $u, v \in W^{1,p}(\Omega)$ the equivalence

$$(\boldsymbol{S}(\nabla u) - \boldsymbol{S}(\nabla v), \nabla u - \nabla v)_\Omega \sim \|\boldsymbol{F}(\nabla u) - \boldsymbol{F}(\nabla v)\|_2^2 \sim \int_\Omega \varphi_{|\nabla u|}(|\nabla u - \nabla v|)\,\mathrm{d}x \tag{2.12}$$

holds true with constants only depending on p. We refer to each quantity in (2.12) as the *quasi-norm* or *natural distance* following, e.g., [1–3, 20]. It has been used very successfully in the finite element analysis of equations with p-structure.

The following lemma, from [28], shows the connection between the natural distance and the Sobolev norms:

Lemma 2.3. *For $p \in (1,\infty)$ and $\varepsilon \in [0,\infty)$ let the operator $\boldsymbol{S}$ satisfy Assumption 2.1 and let $\boldsymbol{F}$ be defined by* (2.11). *Then for all $u, v \in W^{1,p}(\Omega)$ there holds*

(i) *in the case $p \in (1,2]$, with constants only depending on p,*

$$\begin{aligned}\|\nabla(u-v)\|_p^2 &\lesssim \|\boldsymbol{F}(\nabla u) - \boldsymbol{F}(\nabla v)\|_2^2\ \ \|\varepsilon + |\nabla u| + |\nabla v|\|_p^{2-p}, \\ \|\boldsymbol{F}(\nabla u) - \boldsymbol{F}(\nabla v)\|_2^2 &\lesssim \|\nabla(u-v)\|_p^p\,.\end{aligned}$$

(ii) *in the case $p \in [2,\infty)$, with constants only depending on p,*

$$\|\nabla(u-v)\|_p^p \lesssim \|\boldsymbol{F}(\nabla u) - \boldsymbol{F}(\nabla v)\|_2^2 \lesssim \|\varepsilon + |\nabla u| + |\nabla v|\|_p^{p-2}\ \ \|\nabla(u-v)\|_p^2\,.$$

In particular, all constants appearing in (i) and (ii) are independent of $\varepsilon \geq 0$.

2.3. **Regularity assumption.** We impose our assumption on the regularity of solutions to the state equation that will later enable us to derive a priori error estimates for the finite element approximation of (1.1).

Assumption 2.4. *We assume that for $q \in L^{\max\{2,p'\}}(\Omega)$, $\frac{1}{p} + \frac{1}{p'} = 1$, the weak solution u to the equation* (1.1b), (1.1c) *with p-structure satisfies the regularity*

$$\boldsymbol{S}(\nabla u) \in W^{1,2}(\Omega) \qquad \textit{and} \qquad u \in W^{2,2}(\Omega)$$

and that there exist positive constants c_1, c_2, γ such that

$$\|\boldsymbol{S}(\nabla u)\|_{1,2} \leq c_1\|q\|_2 \qquad \textit{and} \qquad \|u\|_{2,2} \leq c_2\|q\|_{\max\{2,p'\}}^{\gamma} \tag{2.13}$$

The regularity Assumption 2.4 is satisfied for certain data:

- In [13] it is shown that, if $\Omega \subset \mathbb{R}^d$, $d \geq 2$, is a bounded convex open set, $q \in L^2(\Omega)$ and $p \in (1,\infty)$, the weak solution u to the equation (1.1b), (1.1c) with p-structure satisfies the regularity $\boldsymbol{S}(\nabla u) \in W^{1,2}(\Omega)$ with
$$C_1\|q\|_2 \leq \|\boldsymbol{S}(\nabla u)\|_{1,2} \leq C_2\|q\|_2,$$
where the constants C_1, C_2 only depend on p, d. In particular, the analysis carried out in [13] covers the p-Laplacian with $\boldsymbol{S}(\nabla u) = |\nabla u|^{p-2}\nabla u$.
- In [14], on certain domains Lipschitz-continuous solutions are obtained whenever $q \in L^r(\Omega)$ with $r > d$ has mean-value zero and $\varepsilon > 0$. [13, Remark 2.7] claims that this implies $W^{2,2}$-regularity.
- In [16] it is shown that, if $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$ is a bounded domain with C^2-boundary, $p \in (1,2]$ and $q \in L^{p'}(\Omega)$, the weak solution u of the p-Laplace equation (1.1b), (1.1c) with $\boldsymbol{S}(\nabla u) = |\nabla u|^{p-2}\nabla u$ fulfills
$$u \in W^{2,2}(\Omega) \qquad \text{with} \qquad \|\nabla^2 u\|_2 \leq C\|q\|_{p'}^{\frac{1}{p-1}}.$$
- In [32] it is shown that, if $\Omega \subset \mathbb{R}^2$ is either convex or has C^2-boundary, for $p \in (1,2)$ the weak solution u of the p-Laplace equation (1.1b), (1.1c) with $\boldsymbol{S}(\nabla u) = |\nabla u|^{p-2}\nabla u$ satisfies
$$q \in L^r \text{ with } r > 2 \qquad \Longrightarrow \qquad u \in W^{2,2}(\Omega).$$

As a consequence, Assumption 2.4 is satisfied for the p-Laplace equation in the case $p \in (1,2]$ if, e.g., $\Omega \subset \mathbb{R}^d$, $d \geq 2$, is a bounded convex domain with C^2-boundary and $q \in L^{p'}(\Omega)$. Since later we take $q \in \mathcal{Q}_{\mathrm{ad}} \subset L^\infty(\Omega)$, we can weaken Assumption 2.4: It is sufficient to assume that $q \in L^\infty(\Omega)$ implies $\boldsymbol{S}(\nabla u) \in W^{1,2}(\Omega)$ and $u \in W^{2,2}(\Omega)$ with $\|\boldsymbol{S}(\nabla u)\|_{1,2} \lesssim \|q\|_\infty$ and $\|u\|_{2,2} \lesssim \|q\|_\infty^\gamma$ (replacing (2.13)).

3. Optimal control problem

In this section, we give a precise definition of the optimal control problem (1.1). For $\frac{1}{p} + \frac{1}{p'} = 1$, i.e., $p' = \frac{p}{p-1}$, the natural spaces for the states and controls are
$$\mathcal{V} := W_0^{1,p}(\Omega), \quad Q := L^{\max\{2,p'\}}(\Omega), \quad \mathcal{Q}_{\mathrm{ad}} := \big\{q \in Q \mid q_a \leq q \leq q_b \text{ a.e. in } \Omega\big\}.$$
The weak formulation of the state equation (1.1b), (1.1c) reads:

For a given control $q \in \mathcal{Q}_{\mathrm{ad}}$ find the state $u = u(q) \in \mathcal{V}$ such that
$$(\boldsymbol{S}(\nabla u), \nabla\varphi)_\Omega = (q,\varphi)_\Omega \qquad \forall\varphi \in \mathcal{V}. \tag{3.1}$$

We now investigate stability and continuity properties of the solution $u \in \mathcal{V}$ with respect to the control q. It will be suitable for this to consider variations of q in $W^{-1,p'}(\Omega)$.

Lemma 3.1. *For each $p \in (1,\infty)$ and $q \in \mathcal{Q}_{\mathrm{ad}}$ there exists a unique solution $u = u(q) \in \mathcal{V}$ to (3.1). This solution satisfies the a priori estimate*
$$\|\nabla u\|_p \leq c_1\left(\|q\|_{-1,p'}^{\frac{1}{p-1}} + c_2\varepsilon\right), \tag{3.2}$$
where $c_1 > 0$ only depends on Ω, p and $c_2 = 1$ if $p < 2$ and $c_2 = 0$ otherwise.

Proof. Lemma 2.2 implies that the operator $-\operatorname{div}\boldsymbol{S}(\nabla\cdot) : W_0^{1,p}(\Omega) \to W^{-1,p'}(\Omega)$ is strictly monotone. Using the theory of monotone operators (see [38,41]), we can thus easily conclude that for each $q \in Q$ there exists a unique solution $u = u(q) \in \mathcal{V}$ to (3.1). The proof of (3.2) is standard and can be found, e.g., in [28]. □

The next lemma states that the solution operator $q \mapsto u = u(q)$ is locally Hölder-continuous.

Lemma 3.2. *For $p \in (1,\infty)$ let $u_1 = u(q_1) \in \mathcal{V}$ and $u_2 = u(q_2) \in \mathcal{V}$ be the solutions to the state equation* (3.1) *for the right-hand side $q_1 \in \mathcal{Q}_{\rm ad}$ and $q_2 \in \mathcal{Q}_{\rm ad}$. Then there exist constants only depending on p, Ω such that*

$$\|\boldsymbol{F}(\nabla u_1) - \boldsymbol{F}(\nabla u_2)\|_2 \lesssim \begin{cases} \|\varepsilon + |\nabla u_1| + |\nabla u_2|\|_p^{\frac{2-p}{2}} \|q_1 - q_2\|_{-1,p'} & \text{for } p \le 2, \\ \|q_1 - q_2\|_{-1,p'}^{\frac{p'}{2}} & \text{for } p \ge 2, \end{cases}$$

$$\|\nabla u_1 - \nabla u_2\|_p \lesssim \begin{cases} \|\varepsilon + |\nabla u_1| + |\nabla u_2|\|_p^{2-p} \|q_1 - q_2\|_{-1,p'} & \text{for } p \le 2, \\ \|q_1 - q_2\|_{-1,p'}^{\frac{1}{p-1}} & \text{for } p \ge 2. \end{cases}$$

Proof. As $u_1 \in \mathcal{V}$ and $u_2 \in \mathcal{V}$ solve (3.1) with right-hand side q_1 and q_2, there holds

$$(\boldsymbol{S}(\nabla u_1) - \boldsymbol{S}(\nabla u_2), \nabla \varphi)_\Omega = \langle q_1 - q_2, \varphi \rangle \qquad \forall \varphi \in W_0^{1,p}(\Omega).$$

Testing this equation with $\varphi = u_1 - u_2$ and employing Lemma 2.2, we get

$$\|\boldsymbol{F}(\nabla u_1) - \boldsymbol{F}(\nabla u_2)\|_2^2 \sim \langle q_1 - q_2, u_1 - u_2 \rangle \le \|q_1 - q_2\|_{-1,p'} \|u_1 - u_2\|_{1,p} \tag{3.3}$$

for $p \in (1,\infty)$. Poincaré's inequality and Lemma 2.3 imply

$$\begin{aligned} \|\boldsymbol{F}(\nabla u_1) - \boldsymbol{F}(\nabla u_2)\|_2^2 &\lesssim \|q_1 - q_2\|_{-1,p'} \|\nabla u_1 - \nabla u_2\|_p \\ &\lesssim \begin{cases} \|q_1 - q_2\|_{-1,p'} \|\varepsilon + |\nabla u_1| + |\nabla u_2|\|_p^{\frac{2-p}{2}} \|\boldsymbol{F}(\nabla u_1) - \boldsymbol{F}(\nabla u_2)\|_2 & \text{for } p \le 2, \\ \|q_1 - q_2\|_{-1,p'} \|\boldsymbol{F}(\nabla u_1) - \boldsymbol{F}(\nabla u_2)\|_2^{\frac{2}{p}} & \text{for } p \ge 2. \end{cases} \end{aligned}$$

This yields the desired estimate in the natural distance. Using similar arguments,

$$\begin{aligned} \|q_1 - q_2\|_{-1,p'} \|\nabla u_1 - \nabla u_2\|_p &\overset{(3.3)}{\gtrsim} \|\boldsymbol{F}(\nabla u_1) - \boldsymbol{F}(\nabla u_2)\|_2^2 \\ &\gtrsim \begin{cases} \|\varepsilon + |\nabla u_1| + |\nabla u_2|\|_p^{p-2} \|\nabla u_1 - \nabla u_2\|_p^2 & \text{for } p \le 2, \\ \|\nabla u_1 - \nabla u_2\|_p^p & \text{for } p \ge 2. \end{cases} \end{aligned}$$

From this we obtain the desired estimate in the $W_0^{1,p}$-norm. ☐

For given $\alpha > 0$ and $u_d \in L^2(\Omega)$ we define the cost functional $J : Q \times \mathcal{V} \to \mathbb{R}$ as

$$J(q,u) := \frac{1}{2}\|u - u_d\|_2^2 + \frac{\alpha}{2}\|q\|_2^2$$

We aim to solve the following optimal control problem:

$$\text{Minimize } J(q,u) \quad \text{subject to (3.1) and } (q,u) \in \mathcal{Q}_{\rm ad} \times \mathcal{V}. \tag{P}$$

We tacitly let $J(q,u) = \infty$ whenever $u \notin L^2(\Omega)$. For the finite element analysis of (P) we will later utilize the following relation which holds for all $(q_1,u_1),(q_2,u_2) \in Q \times \mathcal{V}$ due to the parallelogram law:

$$\begin{aligned} \tfrac{1}{2}\left\|\tfrac{u_1-u_2}{2}\right\|_2^2 + \tfrac{\alpha}{2}\left\|\tfrac{q_1-q_2}{2}\right\|_2^2 + J\big(\tfrac{1}{2}(q_1+q_2), \tfrac{1}{2}(u_1+u_2)\big) \\ \le \tfrac{1}{2} J(q_1,u_1) + \tfrac{1}{2} J(q_2,u_2). \end{aligned} \tag{3.4}$$

Further, we often make use of the continuous embedding $W^{1,p}(\Omega) \subset L^2(\Omega)$ for $p \ge \frac{2d}{d+2}$. As a start, we deal with the existence of solutions to (P).

Theorem 3.3. *For $p \in (1,\infty)$ and $\varepsilon \ge 0$ the optimal control problem (P) has at least one globally optimal control $\overline{q} \in \mathcal{Q}_{\rm ad}$ with corresponding optimal state $\overline{u} = u(\overline{q}) \in \mathcal{V}$.*

Proof. The proof follows standard arguments, cf., [30,33]. According to Lemma 3.1, for each control $q \in \mathcal{Q}_{\rm ad}$ the state equation (3.1) has a unique solution $u = u(q) \in \mathcal{V}$. The functional J is bounded from below. Thus, there exists

$$j := \inf_{q \in \mathcal{Q}_{\rm ad}} J(q, u(q)).$$

Let $\{(q_n, u_n)\}_{n=1}^{\infty}$ be a minimizing sequence, i.e.,

$$q_n \in \mathcal{Q}_{\rm ad}, \qquad u_n := u(q_n), \qquad J(q_n, u_n) \to j \qquad \text{for } n \to \infty.$$

As $\mathcal{Q}_{\rm ad}$ is non-empty, convex, closed and bounded in $L^{\max\{p',2\}}(\Omega)$, it is weakly sequentially compact. Hence, there exists a subsequence denoted again by $\{q_n\}_{n=1}^{\infty}$, that weakly converges in $L^{\max\{p',2\}}(\Omega)$ to a function $\overline{q} \in \mathcal{Q}_{\rm ad}$,

$$q_n \rightharpoonup \overline{q} \qquad \text{weakly in } L^{\max\{p',2\}}(\Omega),$$

and thus strongly in $W^{-1,p'}(\Omega)$. Then, Poincaré's inequality and Lemma 3.2 imply

$$\|u_n - u(\overline{q})\|_{1,p} \lesssim \|\nabla u_n - \nabla u(\overline{q})\|_p \to 0 \qquad (n \to \infty), \tag{3.5}$$

where in the case $p < 2$ also (3.2) was used. If $p \geq \frac{2d}{d+2}$, we directly obtain from (3.5) that $u_n \to u(\overline{q})$ strongly in $L^2(\Omega)$. If $p < \frac{2d}{d+2}$, we proceed differently. Since

$$\|u_n\|_2^2 \lesssim \|u_n - u_d\|_2^2 + \|u_d\|_2^2 \leq 2J(q_n, u_n) + \|u_d\|_2^2 \leq C,$$

it holds in addition for some function $\overline{u} \in L^2(\Omega)$

$$u_n \rightharpoonup \overline{u} \qquad \text{weakly in } L^2(\Omega),$$

and due to $p < 2$, $u_n \rightharpoonup \overline{u}$ weakly in $L^p(\Omega)$. In view of (3.5) we have $u_n \to u(\overline{q})$ strongly in $L^p(\Omega)$. As the weak limit is unique, this yields $\overline{u} = u(\overline{q})$, hence

$$(q_n, u_n) \rightharpoonup (\overline{q}, \overline{u}) = (\overline{q}, u(\overline{q})) \quad \text{weakly in } L^2(\Omega) \times L^2(\Omega)\,.$$

The functional J is convex and continuous on $L^2(\Omega) \times L^2(\Omega)$, so it is weakly lower-semicontinuous. Thus, from this weak convergence we conclude

$$j = \lim_{n\to\infty} J(q_n, u_n) \geq \liminf_{n\to\infty} J(q_n, u_n) \geq J(\overline{q}, \overline{u}) \geq j\,,$$

i.e., $\overline{q}$ is a optimal control. The proof is completed. □

4. Finite element discretization

In this section, we introduce the discretization of the optimal control problem (1.1). We assume that Ω is either a convex polygonal/polyhedral domain or a bounded convex domain with C^2-boundary where for the latter case $d = 2$ is assumed. Let $\mathbb{T}_h$ be a shape regular decomposition of Ω into d-dimensional simplices such that, if Ω is polyhedral, there holds $\overline{\Omega} = \bigcup_{K\in\mathbb{T}_h} \overline{K}$, or if Ω has a curved boundary, the corner points of

$$\overline{\Omega_h} := \bigcup_{K\in\mathbb{T}_h} \overline{K} \subset \overline{\Omega}$$

belong to $\partial\Omega$. By h_K, we denote the diameter of a cell $K \in \mathbb{T}_h$, and by ρ_K the supremum of diameters of inscribed balls. The mesh parameter h represents the maximum diameter of the cells, i.e., $h := \max\{h_K;\, K \in \mathbb{T}_h\}$. We assume that $\mathbb{T}_h$ is non-degenerate, see [7], i.e.,

$$\max_{K\in\mathbb{T}_h} \frac{h_K}{\rho_K} \leq \kappa_0 \qquad \forall h. \tag{4.1}$$

For $K \in \mathbb{T}_h$ we define the set of neighbors N_K and the neighborhood S_K by

$$\begin{aligned} N_K &:= \{K' \in \mathbb{T}_h : \overline{K'} \cap \overline{K} \neq \emptyset\}, \\ S_K &:= \operatorname{int} \bigcup_{K'\in N_K} \overline{K'}. \end{aligned} \tag{4.2}$$

The sets S_K are open, bounded and connected. The non-degeneracy (4.1) of $\mathbb{T}_h$ implies the following properties of $\mathbb{T}_h$, in which all constants are independent of h:

$$|S_K| \sim |K| \text{ for all } K \in \mathbb{T}_h \text{ and } \#N_K \leq m_0 \text{ for some } m_0 \in \mathbb{N}. \tag{4.3}$$

If $\partial\Omega$ is curved, we need to be able to estimate the integral over the part of the domain not covered by Ω_h. As the treatment of curved boundaries is not the purpose of the paper, in this case we restrict ourselves to space dimension $d = 2$ for ease of presentation.

To this end, we need the following

Lemma 4.1. *Let Ω be a convex domain with C^2-boundary and $d = 2$. Then for $u \in W^{1,2}(\Omega)$ the integral of $|u|^2$ on the stripe $\Sigma_h := \overline{\Omega} \setminus \overline{\Omega_h}$ is bounded by*

$$\int_{\Sigma_h} |u(x)|^2 \, \mathrm{d}x \lesssim h^2 \|u\|^2_{1,2;\Omega} \,. \tag{4.4}$$

Its proof can be found, e.g., in the proof of [37, Satz 3.3].

Let $\mathcal{P}_m(K)$ be the set of polynomials on K of degree less than or equal to m. For the discretization of the state equation, we employ the space $\mathcal{V}_h \subset W^{1,\infty}(\Omega)$ of linear finite elements on the triangulation $\mathbb{T}_h$

$$\hat{\mathcal{V}}_h = \big\{ u_h \in C(\overline{\Omega_h}) \mid u_h|_K \in \mathcal{P}_1(K) \text{ for all } K \in \mathbb{T}_h \text{ and } u_h|_{\partial\Omega_h} = 0 \big\}.$$

If Ω is polyhedral, then there holds $\Omega_h = \Omega$ and we set $\mathcal{V}_h = \hat{\mathcal{V}}_h$. Otherwise we define $\mathcal{V}_h$ as the space of functions $u_h \in \hat{\mathcal{V}}_h$ extended to $\overline{\Omega}$ by setting $u_h = 0$ on the stripe $\Sigma_h := \overline{\Omega} \setminus \overline{\Omega_h}$. For the discretization of the controls, we study two approaches:

(a) Variational discretization: The control variable is not explicitly discretized.

(b) Piecewise constant controls on the family of triangulations $\{\mathbb{T}_h\}$ introduced for the discretization of the state variables:

$$\hat{Q}^0_h = \big\{ q_h : \overline{\Omega_h} \to \mathbb{R} \mid q_h|_K \in \mathcal{P}_0(K) \text{ for all } K \in \mathbb{T}_h \big\}.$$

If Ω is polyhedral, we define $Q^0_h = \hat{Q}^0_h$, otherwise

$$Q^0_h = \big\{ q_h \in Q \mid q_h|_{\Omega_h} \in \hat{Q}^0_h \big\}.$$

The discrete admissible set is $\mathcal{Q}^0_{h,\mathrm{ad}} := Q^0_h \cap \mathcal{Q}_{\mathrm{ad}}$.

Note that for the case of curved boundary the functions in Q^0_h are not restricted to a finite dimensional set on the stripe $\Sigma_h = \overline{\Omega} \setminus \overline{\Omega_h}$. Hence, this space is formally infinite dimensional. However, due to the control cost $\frac{\alpha}{2}\|q\|_2^2$ and the fact that $q_h|_{\Sigma_h}$ has no influence on the discrete solution of (4.8), the optimal discrete control will satisfy $q_h|_{\Sigma_h} = 0$ due to our choice $0 \in [q_a, q_b]$. This means that an implementation can work on the finite dimensional set $\hat{Q}^0_h$. In the subsequent analysis $\mathcal{Q}_{h,\mathrm{ad}}$ equals either $\mathcal{Q}_{\mathrm{ad}}$ in case of variational discretization or $\mathcal{Q}^0_{h,\mathrm{ad}}$ in case of cell-wise constant discretization.

Moreover, we introduce $\hat{\Pi}_h : L^1(\Omega_h) \to \hat{Q}^0_h$ as the natural extension of the L^2-projection, i.e., for $q \in L^1(\Omega_h)$ we define $\hat{\Pi}_h q \in \hat{Q}^0_h$ by

$$(\hat{\Pi}_h q, \varphi_h)_{\Omega_h} = (q, \varphi_h)_{\Omega_h} \qquad \forall \varphi_h \in \hat{Q}^0_h \,. \tag{4.5}$$

If Ω is polyhedral, we set $\Pi_h = \hat{\Pi}_h$. Otherwise we define $\Pi_h : Q \to Q^0_h$ by $(\Pi_h q)\big|_{\Omega_h} = \hat{\Pi}_h q$ and $(\Pi_h q)\big|_{\Sigma_h} = q|_{\Sigma_h}$ for all $q \in Q$. It is well-known that the operator Π_h satisfies for any $\nu \in [1,\infty]$ the stability estimate

$$\|\Pi_h q\|_{\nu;\Omega_h} \leq \|q\|_{\nu;\Omega_h} \qquad \forall q \in L^\nu(\Omega_h). \tag{4.6}$$

From the stability property (4.6) one can derive an interpolation estimate for Π_h.

Lemma 4.2. *There exists a constant $c > 0$ independent of h such that for all $q \in W^{m,\nu}(\Omega)$ with $m \in \{0,1\}$ and $\nu \in (1,\infty)$ there holds*

$$\|q - \Pi_h q\|_{-1,\nu} + h\|q - \Pi_h q\|_\nu \leq c h^{m+1} \|q\|_{m,\nu} \,. \tag{4.7}$$

We omit the proof, as it is a standard consequence of orthogonality, the definition of the norms, and standard error estimates for quasi interpolation operators noting that by definition of Π_h the boundary stripe Σ_h induces no error.

The Galerkin approximation of (3.1) consists in replacing the Banach space $\mathcal{V}$ by the finite element space $\mathcal{V}_h$:

For a given control $q \in \mathcal{Q}_{\mathrm{ad}}$ find the discrete state $u_h = u_h(q) \in \mathcal{V}_h$ with

$$(\boldsymbol{S}(\nabla u_h), \nabla \varphi_h)_\Omega = (q, \varphi_h)_\Omega \qquad \forall \varphi_h \in \mathcal{V}_h. \tag{4.8}$$

Existence of a unique solution u_h to (4.8) as well as an a priori estimate for u_h in $W^{1,p}(\Omega)$ follow by using similar arguments as in the continuous case.

Lemma 4.3. *For each $p \in (1, \infty)$ there exists a unique solution $u_h \in \mathcal{V}_h$ to (4.8). This discrete solution satisfies the a priori estimate*

$$\|\nabla u_h\|_p \le c_1 \left(\|q\|_{-1,p'}^{\frac{1}{p-1}} + c_2 \varepsilon \right), \tag{4.9}$$

where $c_1 > 0$ only depends on Ω, p and $c_2 = 1$ if $p < 2$ and $c_2 = 0$ otherwise.

The following lemma is a discrete version of Lemma 3.2.

Lemma 4.4. *For $p \in (1, \infty)$ let $u_{h1} = u_h(q_1) \in \mathcal{V}_h$ and $u_{h2} = u_h(q_2) \in \mathcal{V}_h$ be the solutions to the discrete equation (4.8) for the right-hand side $q_1 \in \mathcal{Q}_{\mathrm{ad}}$ and $q_2 \in \mathcal{Q}_{\mathrm{ad}}$. Then there exist constants only depending on p, Ω such that*

$$\|\boldsymbol{F}(\nabla u_{h1}) - \boldsymbol{F}(\nabla u_{h2})\|_2 \lesssim \begin{cases} \|\varepsilon + |\nabla u_{h1}| + |\nabla u_{h2}|\|_p^{\frac{2-p}{2}} \|q_1 - q_2\|_{-1,p'} & \text{for } p \le 2, \\ \|q_1 - q_2\|_{-1,p'}^{\frac{p'}{2}} & \text{for } p \ge 2, \end{cases}$$

$$\|\nabla u_{h1} - \nabla u_{h2}\|_p \lesssim \begin{cases} \|\varepsilon + |\nabla u_{h1}| + |\nabla u_{h2}|\|_p^{2-p} \|q_1 - q_2\|_{-1,p'} & \text{for } p \le 2, \\ \|q_1 - q_2\|_{-1,p'}^{\frac{1}{p-1}} & \text{for } p \ge 2. \end{cases}$$

Proof. The proof follows along the same lines as the proof of Lemma 3.2 if the space $\mathcal{V}$ is replaced by $\mathcal{V}_h$. □

Now let us consider the discrete optimal control problem. The discrete analog to (P) reads:

(P_h) Minimize $J(q_h, u_h)$ subject to (4.8) and $(q_h, u_h) \in \mathcal{Q}_{h,\mathrm{ad}} \times \mathcal{V}_h$

Following the same arguments used for the proof of Theorem 3.3, we can conclude the existence of a solution to (P_h).

Lemma 4.5. *For each $h > 0$ there exists an optimal control $\overline{q}_h$ with corresponding optimal state $\overline{u}_h$ of the minimization problem (P_h).*

5. Discrete optimality system

In this section we are concerned with an optimality system for (P_h) that can be utilized for practical computation of the discrete optimal solution. We will close this section with a discussion on the continuous optimality system. For ease of exposition, we restrict ourselves to the particular nonlinear operator $(1.2)_2$, i.e., for

$$a(u)(\varphi) := (\boldsymbol{S}(\nabla u), \nabla \varphi)_\Omega, \qquad \boldsymbol{S}(\nabla u) = \left(\varepsilon^2 + |\nabla u|^2\right)^{\frac{p-2}{2}} \nabla u,$$

we consider the discrete variational formulation of the state equation (4.8)

$$a(u_h)(\varphi_h) = (q, \varphi_h)_\Omega \qquad \forall \varphi_h \in \mathcal{V}_h. \tag{5.1}$$

On the discrete level it can be shown that the semi-linear form a is Gâteaux differentiable for each $\varepsilon > 0$ with Gâteaux derivative

$$
\begin{aligned}
a'(v_h)(w_h,\varphi_h) = &\int_\Omega (\varepsilon^2+|\nabla v_h|^2)^{\frac{p-2}{2}}\nabla w_h\cdot\nabla\varphi_h \,\mathrm{d}x \\
&+(p-2)\int_\Omega (\varepsilon^2+|\nabla v_h|^2)^{\frac{p-4}{2}}(\nabla v_h\cdot\nabla w_h)(\nabla v_h\cdot\nabla\varphi_h)\,\mathrm{d}x.
\end{aligned}
\tag{5.2}
$$

Since this will be crucial for this section, we limit ourselves to $\varepsilon>0$ here.

Now we can define an adjoint problem associated to (4.8): Find $z_h\in\mathcal{V}_h$ such that

$$
a'(u_h)(\varphi_h,z_h) = (u_h-u_d,\varphi_h)_\Omega \qquad \forall \varphi_h\in\mathcal{V}_h. \tag{5.3}
$$

The next lemma concerns the unique solvability of the discrete adjoint problem.

Lemma 5.1. *Let $p\in(1,\infty)$, $\varepsilon>0$ and $b\in W^{-1,\max\{p',2\}}(\Omega)$ with $p'=\frac{p}{p-1}$. For each $h>0$ there exists a unique solution $z_h\in\mathcal{V}_h$ to*

$$
a'(u_h)(\varphi_h,z_h) = \langle b,\varphi_h\rangle \qquad \forall\varphi_h\in\mathcal{V}_h, \tag{5.4}
$$

where u_h solves (4.8)*. The adjoint solution z_h satisfies the a priori estimate*

$$
\|\nabla z_h\|_{\min\{p,2\}} \le c\|b\|_{-1,\max\{p',2\}}, \tag{5.5}
$$

where the constant c only depends on p, ε, Ω, q_a, q_b.

Proof. First of all we prove that if there exists a solution z_h to Problem (5.4), then z_h is uniquely determined. To this end, we assume that z_h^1 and z_h^2 are two functions satisfying (5.4). Setting $\xi_h := z_h^1 - z_h^2$, we observe that

$$
a'(u_h)(\varphi_h,\xi_h) = 0 \qquad \forall\varphi_h\in\mathcal{V}_h. \tag{5.6}
$$

We recall that u_h is uniformly bounded in $W^{1,p}(\Omega)$ by the data, see (4.9). In the case $p\le 2$ we may estimate the quantity $a'(u_h)(\xi_h,\xi_h)$ as follows:

$$
\begin{aligned}
a'(u_h)(\xi_h,\xi_h) &\overset{(5.2)}{\ge} \int_\Omega \left(\varepsilon^2+|\nabla u_h|^2\right)^{\frac{p-2}{2}}|\nabla\xi_h|^2\,\mathrm{d}x \\
&\qquad +(p-2)\int_\Omega \left(\varepsilon^2+|\nabla u_h|^2\right)^{\frac{p-4}{2}}|\nabla u_h|^2|\nabla\xi_h|^2\,\mathrm{d}x \\
&\ge (p-1)\int_\Omega \left(\varepsilon^2+|\nabla u_h|^2\right)^{\frac{p-2}{2}}|\nabla\xi_h|^2\,\mathrm{d}x \\
&\ge (p-1)\int_\Omega (\varepsilon+|\nabla u_h|)^{p-2}|\nabla\xi_h|^2\,\mathrm{d}x.
\end{aligned}
$$

Using Hölder's inequality, $q(x)\in[q_a,q_b]$ a.e., and (4.9), for $p\le 2$ and $\varepsilon>0$ we arrive at

$$
a'(u_h)(\xi_h,\xi_h) \ge (p-1)\|\varepsilon+|\nabla u_h|\|_p^{p-2}\|\nabla\xi_h\|_p^2 \ge c\|\nabla\xi_h\|_p^2.
$$

In the case $p>2$ we can bound the quantity $a'(u_h)(\xi_h,\xi_h)$ from below as follows:

$$
\begin{aligned}
a'(u_h)(\xi_h,\xi_h) &\overset{(5.2)}{=} \int_\Omega (\varepsilon^2+|\nabla u_h|^2)^{\frac{p-2}{2}}\nabla\xi_h\cdot\nabla\xi_h\,\mathrm{d}x \\
&\qquad +(p-2)\int_\Omega (\varepsilon^2+|\nabla u_h|^2)^{\frac{p-4}{2}}(\nabla u_h\cdot\nabla\xi_h)(\nabla u_h\cdot\nabla\xi_h)\,\mathrm{d}x \\
&\ge \int_\Omega (\varepsilon^2+|\nabla u_h|^2)^{\frac{p-2}{2}}|\nabla\xi_h|^2\,\mathrm{d}x \ge \int_\Omega \varepsilon^{p-2}|\nabla\xi_h|^2\,\mathrm{d}x = \varepsilon^{p-2}\|\nabla\xi_h\|_2^2.
\end{aligned}
$$

To sum up, we may deduce that there exists a constant $c=c(p,\varepsilon,\Omega,q_a,q_b)$ with

$$
a'(u_h)(\xi_h,\xi_h) \ge c\|\nabla\xi_h\|_{\min\{p,2\}}^2. \tag{5.7}
$$

From (5.6), (5.7) and Poincaré's inequality we infer $\xi_h\equiv 0$ and, hence, $z_h^1 = z_h^2$. Since the system (5.4) is linear, and the space $\mathcal{V}_h$ is finite dimensional, we can conclude from uniqueness that there exists a solution z_h. For the proof of (5.5), we test (5.4) with $\varphi_h := z_h$. Then we can apply the same arguments that led to (5.7) in order to obtain

$$\|b\|_{-1,\max\{p',2\}}\|z_h\|_{1,\min\{p,2\}} \geq \langle b, z_h\rangle = a'(u_h)(z_h, z_h) \gtrsim \|\nabla z_h\|^2_{\min\{p,2\}}.$$

Together with Poincaré's inequality, this yields the assertion. □

With the help of the discrete adjoint state we can now formulate an optimality system for (P_h):

Lemma 5.2. *Let $\varepsilon > 0$ be given. If a control $\overline{q}_h \in \mathcal{Q}_{h,\mathrm{ad}}$ with state $\overline{u}_h = u_h(\overline{q}_h) \in \mathcal{V}_h$ is an optimal solution to problem (P_h), then there exists an adjoint state $\overline{z}_h \in \mathcal{V}_h$ so that*

$$a(\overline{u}_h)(\varphi_h) = (\overline{q}_h, \varphi_h)_\Omega \qquad \forall \varphi_h \in \mathcal{V}_h \tag{5.8a}$$

$$a'(\overline{u}_h)(\overline{z}_h, \varphi_h) = (\overline{u}_h - u_d, \varphi_h)_\Omega \qquad \forall \varphi_h \in \mathcal{V}_h \tag{5.8b}$$

$$(\alpha\overline{q}_h + \overline{z}_h, \delta q_h - \overline{q}_h)_\Omega \geq 0 \qquad \forall \delta q_h \in \mathcal{Q}_{h,\mathrm{ad}}. \tag{5.8c}$$

Remark 5.3. *It is well-known that the variational inequality* (5.8c) *has a pointwise almost everywhere representation, see, e.g., [40]. Indeed,* (5.8c) *can be rewritten using the projection $P_{[q_a,q_b]}$ onto the interval $[q_a, q_b]$ defined by*

$$P_{[q_a,q_b]}\big(f(x)\big) = \min\Big(q_b, \max\big(q_a, f(x)\big)\Big). \tag{5.9}$$

In the case $\mathcal{Q}_{h,\mathrm{ad}} = \mathcal{Q}_{\mathrm{ad}}$ a control $\overline{q}_h \in \mathcal{Q}_{h,\mathrm{ad}}$ solving (P_h) necessarily satisfies (5.8) *and thus the control $\overline{q}_h$ and the solution $\overline{z}_h$ of* (5.8b) *satisfy the projection formula*

$$\overline{q}_h = P_{[q_a,q_b]}\left(-\frac{1}{\alpha}\overline{z}_h\right).$$

In the case $\mathcal{Q}_{h,\mathrm{ad}} = \mathcal{Q}^0_{h,\mathrm{ad}}$ there holds

$$\overline{q}_h = P_{[q_a,q_b]}\left(-\frac{1}{\alpha}\Pi_h\overline{z}_h\right),$$

where Π_h denotes the L^2-projection on $\mathcal{Q}^0_{h,\mathrm{ad}}$.

The question arises whether an analogous optimality system for (P) can be formulated. A closer look at (5.2) however reveals that this is not an easy task: The natural regularity for the state $u \in W^{1,p}(\Omega)$ is not sufficient to formulate a well-defined, invertible, Gâteaux derivative as already discussed in the introduction.

Still, we can pass to the limit in the discrete adjoint. According to the a priori estimate (5.5) and Lemma 4.3, the discrete adjoint solution $\overline{z}_h$ is uniformly bounded in $W^{1,\min\{p,2\}}(\Omega)$ for $p \geq \frac{2d}{d+2}$ as it is shown by the following calculation:

$$\begin{aligned}\|\nabla\overline{z}_h\|_{\min\{p,2\}} \leq c\|u_h - u_d\|_{-1,\max\{p',2\}} &= \sup_{\varphi\in W_0^{1,\min\{p,2\}}(\Omega)} \frac{(u_h - u_d, \varphi)_\Omega}{\|\varphi\|_{1,\min\{p,2\}}}\\ &\leq \sup_{\varphi\in W_0^{1,\min\{p,2\}}(\Omega)} \frac{\|u_h - u_d\|_2\|\varphi\|_2}{\|\varphi\|_{1,\min\{p,2\}}} \lesssim \|u_h\|_{1,p} + \|u_d\|_2 \lesssim C.\end{aligned}$$

Hence, there exists a function $\overline{z} \in W_0^{1,\min\{p,2\}}(\Omega)$ such that up to a subsequence

$$\overline{z}_h \rightharpoonup \overline{z} \qquad \text{weakly in } W_0^{1,\min\{p,2\}}(\Omega) \qquad (h \to 0)\,. \tag{5.10}$$

Due to the compact embedding $W^{1,\min\{p,2\}}(\Omega) \subset L^2(\Omega)$ for $p > \frac{2d}{d+2}$ we get

$$\overline{z}_h \to \overline{z} \qquad \text{strongly in } L^2(\Omega) \qquad (h \to 0). \tag{5.11}$$

Consequently, the projection formula yields strong convergence of the controls

$$\overline{q}_h = P_{[q_a,q_b]}\left(-\frac{1}{\alpha}\overline{z}_h\right) \to \overline{q} = P_{[q_a,q_b]}\left(-\frac{1}{\alpha}\overline{z}\right) \qquad \text{in } L^2(\Omega) \qquad (h \to 0)$$

for $p > \frac{2d}{d+2}$ in the case of variational discretization, see Remark 5.3. This also shows the additional regularity $q \in W^{1,\min\{p,2\}}(\Omega)$ for any such limit point.

6. FE approximation of the p-Laplace equation

Before analyzing the convergence of the discretized optimal control problem, we collect and extend several results regarding the FE approximation of the p-Laplace equation. The first lemma states that the Galerkin approximation is a quasi best-approximation with respect to the natural distance.

Lemma 6.1 (Best-approximation in quasinorms)**.** *For $\varepsilon \geq 0$ and $p \in (1,\infty)$ let $u \in \mathcal{V}$ be the unique solution of* (3.1)*, and $u_h \in \mathcal{V}_h$ its finite element approximation, i.e., $u_h \in \mathcal{V}_h$ is the unique solution of* (4.8)*. Then, there hold for $\Sigma_h := \overline{\Omega} \setminus \overline{\Omega_h}$*

$$\|\boldsymbol{F}(\nabla u) - \boldsymbol{F}(\nabla u_h)\|_{2;\Omega} \lesssim \inf_{\varphi_h \in \mathcal{V}_h} \|\boldsymbol{F}(\nabla u) - \boldsymbol{F}(\nabla \varphi_h)\|_{2;\Omega}, \tag{6.1a}$$

$$\|\boldsymbol{F}(\nabla u) - \boldsymbol{F}(\nabla u_h)\|_{2;\Omega} \lesssim \inf_{\varphi_h \in \mathcal{V}_h} \|\boldsymbol{F}(\nabla u) - \boldsymbol{F}(\nabla \varphi_h)\|_{2;\Omega_h} + \|\boldsymbol{F}(\nabla u)\|_{2;\Sigma_h}, \tag{6.1b}$$

where the constants only depend on p (they are independent of h and ε).

Proof. For polyhedral Ω the lemma is proven in [21]. Using Lemma 2.2, the Galerkin orthogonality ($\mathcal{V}_h \subset \mathcal{V}$), we can deduce for arbitrary $\varphi_h \in \mathcal{V}_h$

$$\begin{aligned}\int_\Omega \varphi_{|\nabla u|}\left(|\nabla u - \nabla u_h|\right)\,\mathrm{d}x &\sim \int_\Omega (\boldsymbol{S}(\nabla u) - \boldsymbol{S}(\nabla u_h)) \cdot (\nabla u - \nabla u_h)\,\mathrm{d}x \\ &\sim \int_\Omega (\boldsymbol{S}(\nabla u) - \boldsymbol{S}(\nabla u_h)) \cdot (\nabla u - \nabla \varphi_h)\,\mathrm{d}x \\ &\lesssim \int_\Omega \varphi'_{|\nabla u|}\left(|\nabla u - \nabla u_h|\right)|\nabla u - \nabla \varphi_h|\,\mathrm{d}x\,.\end{aligned}$$

We apply Young's inequality (2.10) to the shifted function $\varphi_{|\nabla u|}$ in order to obtain

$$\int_\Omega \varphi_{|\nabla u|}\left(|\nabla u - \nabla u_h|\right)\,\mathrm{d}x \lesssim \int_\Omega \varphi_{|\nabla u|}\left(|\nabla u - \nabla \varphi_h|\right)\,\mathrm{d}x\,.$$

Using Lemma 2.2 and taking the infimum over all $\varphi_h \in \mathcal{V}_h$, we arrive at the assertion (6.1a). There holds $\varphi_h|_{\Sigma_h} = 0$ for all $\varphi_h \in \mathcal{V}_h$ and thus (6.1b) easily follows from inequality (6.1a). □

The Scott-Zhang interpolation operator $j_h : W_0^{1,1}(\Omega) \to \mathcal{V}_h$, see [39], is defined in such a way that it fulfills $j_h v = v$ for all $v \in \mathcal{V}_h$ and preserves homogeneous boundary conditions. It is also suitable for interpolation in quasi-norms as it satisfies the following property, see [21]: For all $v \in W^{1,p}(\Omega)$ and $K \in \mathbb{T}_h$ there holds

$$\fint_K |\boldsymbol{F}(\nabla v) - \boldsymbol{F}(\nabla j_h v)|^2\,\mathrm{d}x \lesssim \inf_{\boldsymbol{\eta} \in \mathbb{R}^d} \fint_{S_K} |\boldsymbol{F}(\nabla v) - \boldsymbol{F}(\boldsymbol{\eta})|^2\,\mathrm{d}x, \tag{6.2}$$

where the constant only depends on p. In particular, it is independent of h and ε. On the basis of (6.2) it is a simple matter to derive an interpolation estimate in quasi-norms: As the function $\boldsymbol{F}$ is surjective, (6.2) implies

$$\fint_K |\boldsymbol{F}(\nabla v) - \boldsymbol{F}(\nabla j_h v)|^2\,\mathrm{d}x \lesssim \inf_{\boldsymbol{\xi} \in \mathbb{R}^d} \fint_{S_K} |\boldsymbol{F}(\nabla v) - \boldsymbol{\xi}|^2\,\mathrm{d}x. \tag{6.3}$$

Now let us assume that $v \in W^{1,p}(\Omega)$ satisfies the regularity $\boldsymbol{F}(\nabla v) \in W^{1,2}(\Omega)^d$. If we choose $\boldsymbol{\xi} = \langle \boldsymbol{F}(\nabla v)\rangle_{S_K}$ in (6.3), we can apply Poincaré's inequality (2.2):

$$\fint_K |\boldsymbol{F}(\nabla v) - \boldsymbol{F}(\nabla j_h v)|^2 \, \mathrm{d}x \lesssim \fint_{S_K} h_K^2 |\nabla \boldsymbol{F}(\nabla v)|^2 \, \mathrm{d}x. \tag{6.4}$$

In order to obtain a global version of (6.4), we sum the inequality (6.4) over all elements $K \in \mathbb{T}_h$ and use the mesh properties (4.3):

$$\|\boldsymbol{F}(\nabla v) - \boldsymbol{F}(\nabla j_h v)\|_{2;\Omega_h} \leq ch \|\nabla \boldsymbol{F}(\nabla v)\|_{2;\Omega_h}. \tag{6.5}$$

Combining Lemma 6.1 and (6.5), we obtain an error estimate in quasi-norms.

Lemma 6.2. *For $\varepsilon \geq 0$ and $p \in (1,\infty)$ let $u \in \mathcal{V}$ be the unique solution of* (3.1)*, and $u_h \in \mathcal{V}_h$ its finite element approximation, i.e., $u_h \in \mathcal{V}_h$ is the unique solution of* (4.8)*. In case of curved $\partial\Omega$, we require that $d = 2$.*

(i) *If u satisfies the regularity assumption $\boldsymbol{F}(\nabla u) \in W^{1,2}(\Omega)^d$, then there holds*

$$\|\boldsymbol{F}(\nabla u) - \boldsymbol{F}(\nabla u_h)\|_2 \leq ch \|\boldsymbol{F}(\nabla u)\|_{1,2}.$$

(ii) *If u satisfies $\boldsymbol{S}(\nabla u) \in W^{1,2}(\Omega)^d$ and $u \in W^{2,2}(\Omega)$, then there holds*

$$\|\boldsymbol{F}(\nabla u) - \boldsymbol{F}(\nabla u_h)\|_2 \leq ch \|\boldsymbol{S}(\nabla u)\|_{1,2}^{\frac{1}{2}} \|u\|_{2,2}^{\frac{1}{2}}.$$

All constants c only depend on p (they are independent of h and ε).

Proof. (i) For polyhedral Ω the assertion directly follows by combining (6.1a) and (6.5). If $\partial\Omega$ is curved, we use (6.1b), (6.5) and take into account Lemma 4.1:

$$\begin{aligned}\|\boldsymbol{F}(\nabla u) - \boldsymbol{F}(\nabla u_h)\|_{2;\Omega} &\overset{(6.1\mathrm{b})}{\lesssim} \inf_{\varphi_h \in \mathcal{V}_h} \|\boldsymbol{F}(\nabla u) - \boldsymbol{F}(\nabla \varphi_h)\|_{2;\Omega_h} + \|\boldsymbol{F}(\nabla u)\|_{2;\Sigma_h} \\ &\lesssim \|\boldsymbol{F}(\nabla u) - \boldsymbol{F}(\nabla j_h u)\|_{2;\Omega_h} + \|\boldsymbol{F}(\nabla u)\|_{2;\Sigma_h} \overset{(6.5),(4.4)}{\leq} ch\|\boldsymbol{F}(\nabla u)\|_{1,2;\Omega}.\end{aligned}$$

(ii) From (6.5), the pointwise estimate with constants independent of ε (cf. [4])

$$|\nabla \boldsymbol{F}(\nabla v)|^2 \sim |\nabla \boldsymbol{S}(\nabla v)||\nabla^2 v|$$

and Hölder inequality, we get for all $v \in W^{2,2}(\Omega)$ with $\boldsymbol{S}(\nabla v) \in W^{1,2}(\Omega)^d$

$$\|\boldsymbol{F}(\nabla v) - \boldsymbol{F}(\nabla j_h v)\|_{2;\Omega_h} \lesssim h \|\nabla \boldsymbol{S}(\nabla v)\|_{2;\Omega_h}^{\frac{1}{2}} \|\nabla^2 v\|_{2;\Omega_h}^{\frac{1}{2}}. \tag{6.6}$$

If Ω is polyhedral, the assertion directly follows by combining (6.1a) and (6.6). If $\partial\Omega$ is curved, we use (6.1b), (6.6), Lemma 2.2 and Lemma 4.1:

$$\begin{aligned}\|\boldsymbol{F}(\nabla u) - \boldsymbol{F}(\nabla u_h)\|_{2;\Omega}^2 &\lesssim \inf_{\varphi_h \in \mathcal{V}_h} \|\boldsymbol{F}(\nabla u) - \boldsymbol{F}(\nabla \varphi_h)\|_{2;\Omega_h}^2 + \|\boldsymbol{F}(\nabla u)\|_{2;\Sigma_h}^2 \\ &\lesssim \|\boldsymbol{F}(\nabla u) - \boldsymbol{F}(\nabla j_h u)\|_{2;\Omega_h}^2 + (\boldsymbol{S}(\nabla u), \nabla u)_{\Sigma_h} \\ &\lesssim h^2 \|\nabla \boldsymbol{S}(\nabla u)\|_{2;\Omega_h} \|\nabla^2 u\|_{2;\Omega_h} + \|\boldsymbol{S}(\nabla u)\|_{2;\Sigma_h} \|\nabla u\|_{2;\Sigma_h} \\ &\lesssim h^2 \|\nabla \boldsymbol{S}(\nabla u)\|_{2;\Omega_h} \|\nabla^2 u\|_{2;\Omega_h} + h\|\boldsymbol{S}(\nabla u)\|_{1,2;\Omega} \cdot h\|u\|_{2,2;\Omega}\end{aligned}$$

Taking the square root, we obtain the stated a priori estimate. □

If the solution of the state equation u satisfies the regularity assumption (2.13), according to Lemma 6.2 (ii) the error measured in the natural distance can be bounded in terms of the control q. From this we get error bounds in the $W^{1,p}$-norm.

Corollary 6.3. *For $p \in (1,\infty)$, $\varepsilon \geq 0$ and any $q \in \mathcal{Q}_{\mathrm{ad}}$ let $u = u(q) \in \mathcal{V}$ be the solution of* (3.1) *and let $u_h = u_h(q) \in \mathcal{V}_h$ be its discrete approximation, i.e.,* *the solution of* (4.8)*. In case of curved* $\partial\Omega$ *we require that* $d = 2$*. If* u *satisfies Assumption 2.4, there exist constants only depending on* p *with*

$$\|\boldsymbol{F}(\nabla u) - \boldsymbol{F}(\nabla u_h)\|_2 \le ch, \qquad \|\nabla u - \nabla u_h\|_p \le \begin{cases} ch & \text{for } p \le 2, \\ ch^{\frac{2}{p}} & \text{for } p \ge 2. \end{cases} \tag{6.7}$$

In particular, the constants do not depend on the mesh size h *and* ε*.*

Proof. The error estimate in the natural distance directly follows from Lemma 6.2 (ii), Assumption 2.4 and $\|q\|_{\max\{2,p'\}} \lesssim \max\{|q_a|, |q_b|\}$:

$$\|\boldsymbol{F}(\nabla u) - \boldsymbol{F}(\nabla u_h)\|_2 \le Ch\|q\|_2^{\frac{1}{2}}\|q\|_{\max\{2,p'\}}^{\frac{\gamma}{2}} \le Ch.$$

In order to derive the error estimates in the $W^{1,p}$-norm, we apply Lemma 2.3

$$\|\boldsymbol{F}(\nabla u) - \boldsymbol{F}(\nabla u_h)\|_2^2 \gtrsim \begin{cases} \|\varepsilon + |\nabla u| + |\nabla u_h|\|_p^{p-2}\|\nabla u - \nabla u_h\|_p^2 & \text{for } p \le 2, \\ \|\nabla u - \nabla u_h\|_p^p & \text{for } p \ge 2, \end{cases}$$

and, if $p \le 2$, use the stability Lemmas 3.1, 4.3 and $\|q\|_{p'} \lesssim \max\{|q_a|, |q_b|\}$. □

If higher regularity is not available for the solution of state equation (3.1), we still have strong convergence of its finite element approximation in $W^{1,p}(\Omega)$.

Lemma 6.4. *For* $p \in (1, \infty)$*,* $\varepsilon \ge 0$ *let* $u \in \mathcal{V}$ *be the unique solution of the state equation* (3.1) *and let* $u_h \in \mathcal{V}_h$ *be the unique solution of its discrete approximation* (4.8)*, each for the right-hand side* $q \in W^{-1,p'}(\Omega)$*. Then* u_h *converges strongly in* $\mathcal{V}$ *to* u *for* $h \to 0$*, i.e.,*

$$\lim_{h\to 0}\|u - u_h\|_{1,p} = 0. \tag{6.8}$$

Proof. The lemma is proven in [12] for the case that Ω is polyhedral. Let Ω be a bounded convex domain with $\partial\Omega \in C^2$. Since $C_0^\infty(\Omega)$ is dense in $\mathcal{V}$, there exists a sequence $(\Phi_n) \subset C_0^\infty(\Omega)$ with

$$\|u - \Phi_n\|_{1,p} \to 0 \qquad (n \to \infty). \tag{6.9}$$

Let $i_h : C(\overline{\Omega}) \to \mathcal{V}_h$ denote the Lagrange interpolation operator, see [15]. On the stripe $\Sigma_h = \overline{\Omega} \setminus \overline{\Omega_h}$ we can set $(i_h\Phi)\big|_{\Sigma_h} = 0$ for $\Phi \in C(\overline{\Omega})$. Applied to Φ_n, for all $K \in \mathbb{T}_h$ it satisfies

$$\|\Phi_n - i_h\Phi_n\|_{1,p;K} \le c|K|^{1/p}h_K\|\Phi_n\|_{2,\infty;K}\,. \tag{6.10}$$

Because of Lemma 6.1 the finite element solution u_h fulfills

$$\begin{aligned}\|\boldsymbol{F}(\nabla u) - \boldsymbol{F}(\nabla u_h)\|_2 &\lesssim \inf_{\varphi_h\in\mathcal{V}_h}\|\boldsymbol{F}(\nabla u) - \boldsymbol{F}(\nabla\varphi_h)\|_2 \lesssim \|\boldsymbol{F}(\nabla u) - \boldsymbol{F}(\nabla i_h\Phi_n)\|_2 \\ &\lesssim \|\boldsymbol{F}(\nabla u) - \boldsymbol{F}(\nabla\Phi_n)\|_2 + \|\boldsymbol{F}(\nabla\Phi_n) - \boldsymbol{F}(\nabla i_h\Phi_n)\|_2\,. \end{aligned} \tag{6.11}$$

As the support of Φ_n, $\operatorname{supp}(\Phi_n)$, is compact and $\operatorname{supp}(\Phi_n) \subset \Omega$, there exists $h_0 = h_0(n) > 0$ such that

$$h < h_0(n) \qquad \Rightarrow \qquad \operatorname{supp}(\Phi_n) \subset \overline{\Omega_h} \qquad \text{with} \qquad \overline{\Omega_h} = \bigcup_{K\in\mathbb{T}_h}\overline{K}.$$

We can then infer from (6.10) that for $h < h_0(n)$ the estimate

$$\|\Phi_n - i_h\Phi_n\|_{1,p;\Omega} \le c|\Omega|^{1/p}h\|\Phi_n\|_{2,\infty;\Omega} \tag{6.12}$$

holds true. As the natural distance relates to the $W_0^{1,p}$-norm, see Lemma 2.3, we can combine (6.11) and (6.12) in order to obtain for each $n \in \mathbb{N}$

$$\lim_{h\to 0}\|\boldsymbol{F}(\nabla u) - \boldsymbol{F}(\nabla u_h)\|_2 \lesssim \|\boldsymbol{F}(\nabla u) - \boldsymbol{F}(\nabla\Phi_n)\|_2\,.$$

By employing Lemma 2.3 and recalling (6.9), from this we infer the assertion. □

7. Convergence of the approximation of the optimal control problem

This section contains the main results of the paper: Without assuming any regularity we show for the case of piecewise constant controls that the sequence of discrete global optimal solutions $(\overline{q}_h, \overline{u}_h)$ has a strong accumulation point $(\overline{q}, \overline{u}) \in \mathcal{Q}_{\mathrm{ad}} \times \mathcal{V}$ that is a global optimal solution of the original optimal control problem. Under the regularity Assumption 2.4 we then prove a priori error estimates quantifying the order of convergence for both variational discretization and piecewise constant controls.

Theorem 7.1 (Convergence of global minimizers)**.** *For $\varepsilon \in [0, \infty)$ and $p \in [\frac{2d}{d+2}, \infty)$ let $\boldsymbol{S}$ satisfy Assumption 2.1. For each $h > 0$ let $\overline{q}_h \in \mathcal{Q}_{h,\mathrm{ad}}$ be a discrete global optimal control and $\overline{u}_h = u_h(\overline{q}_h) \in \mathcal{V}_h$ the corresponding discrete optimal state, i.e., $(\overline{q}_h, \overline{u}_h)$ solves (P_h). Then the sequence $(\overline{q}_h, \overline{u}_h)$ has a weak accumulation point $(\overline{q}, \overline{u}) \in \mathcal{Q}_{\mathrm{ad}} \times \mathcal{V}$. Further, any weak accumulation point is also a strong accumulation point, i.e., up to a subsequence*

$$\overline{q}_h \to \overline{q} \text{ in } L^2(\Omega), \qquad \overline{u}_h \to \overline{u} \text{ in } W^{1,p}(\Omega) \qquad (h \to 0).$$

Moreover, any such point $(\overline{q}, \overline{u})$ is a global optimal solution of (P).

Proof. For each $h > 0$ let $(\overline{q}_h, \overline{u}_h)$ be a global solution of (P_h). Weak accumulation points of $(\overline{q}_h, \overline{u}_h)$ exist in $L^{\max\{p',2\}}(\Omega) \times W_0^{1,p}(\Omega)$ due to the uniform a priori bounds

$$\|\overline{q}_h\|_{\max\{p',2\}} \lesssim \max\{|q_a|, |q_b|\}, \qquad \|\overline{u}_h\|_{1,p} \overset{(4.9)}{\lesssim} C \qquad \text{uniformly in } h\,. \tag{7.1}$$

Now, let $(\bar{q}, \bar{u})$ be a global minimizer of (P), whose existence is ensured by Theorem 3.3.

For piecewise constant controls, i.e., $\mathcal{Q}_{h,\mathrm{ad}} = \mathcal{Q}^0_{h,\mathrm{ad}}$ we define Π_h as L^2-projection given in Section 4. For variational controls, i.e., $\mathcal{Q}_{h,\mathrm{ad}} = \mathcal{Q}_{\mathrm{ad}}$, we set $\Pi_h = \mathrm{Id}$. Then sequence of minimizers $(\overline{q}_h, \overline{u}_h)$ satisfies

$$J(\overline{q}_h, \overline{u}_h) \le J(\Pi_h \overline{q}, u_h(\Pi_h \overline{q}))\,. \tag{7.2}$$

Using $W^{1,p}(\Omega) \subset L^2(\Omega)$ for $p \ge \frac{2d}{d+2}$, we can then infer

$$\|u_h(\Pi_h \overline{q}) - u(\overline{q})\|_2 \le \underbrace{\|u_h(\Pi_h \overline{q}) - u_h(\overline{q})\|_2}_{\to 0 \text{ due to Lemma 4.4,(4.9),(4.7)}} + \underbrace{\|u_h(\overline{q}) - u(\overline{q})\|_2}_{\to 0 \text{ due to Lemma 6.4}} \overset{h\to 0}{\longrightarrow} 0\,.$$

Therefore, inequality (7.2) implies

$$\limsup_{h\to 0} J(\overline{q}_h, \overline{u}_h) \le \limsup_{h\to 0} J(\Pi_h \overline{q}, u_h(\Pi_h \overline{q})) = J(\overline{q}, u(\overline{q})).$$

Hence, any weak limit (q, u) of $(\overline{q}_h, \overline{u}_h)$ in $L^{\max\{p',2\}}(\Omega) \times W^{1,p}(\Omega)$ satisfies

$$J(q, u) \le \liminf_{h\to 0} J(\overline{q}_h, \overline{u}_h) \le \limsup_{h\to 0} J(\overline{q}_h, \overline{u}_h) \le J(\bar{q}, \bar{u}),$$

as J is weakly lower semi-continuous. Analogously to the proof of Theorem 3.3, one can show that (q, u) solves (3.1). Hence, (q, u) is a global minimizer of (P) and

$$J(\overline{q}_h, \overline{u}_h) \to J(q, u) \qquad (h \to 0)\,. \tag{7.3}$$

Further, $\overline{q}_h \rightharpoonup q$ weakly in $L^{\max\{p',2\}}(\Omega)$ implies $\overline{q}_h \to q$ strongly in $W^{-1,p'}(\Omega)$. We apply Lemma 6.4, Lemma 4.4 (together with the bound (4.9) in the case $p \le 2$), to see

$$\|\overline{u}_h - u\|_{1,p} \le \underbrace{\|u_h(\overline{q}_h) - u_h(q)\|_{1,p}}_{\to 0 \text{ due to Lemma 4.4,(4.9)}} + \underbrace{\|u_h(q) - u(q)\|_{1,p}}_{\to 0 \text{ due to Lemma 6.4}} \overset{h\to 0}{\longrightarrow} 0,$$

in order to obtain strong convergence $\overline{u}_h \to u$ in $W_0^{1,p}(\Omega)$. By the parallelogram law (3.4), for $\hat{q} = \frac{1}{2}(\overline{q}_h + q)$ and $\hat{u} = \frac{1}{2}(\overline{u}_h + u)$ we get

$$\frac{1}{2}\|\overline{u}_h - u\|^2 + \frac{\alpha}{2}\|\overline{q}_h - q\|_2^2 \le \frac{1}{2}J(q,u) + \frac{1}{2}J(\overline{q}_h, \overline{u}_h) - J(\hat{q}, \hat{u}).$$

We set $\tilde{u} = u(\hat{q})$. Then there holds $J(q,u) \le J(\hat{q}, \tilde{u})$ and hence

$$\frac{1}{2}\|\overline{u}_h - u\|^2 + \frac{\alpha}{2}\|\overline{q}_h - q\|_2^2 \le \frac{1}{2}J(\overline{q}_h, \overline{u}_h) - \frac{1}{2}J(q,u) + \Big(J(\hat{q},\tilde{u}) - J(\hat{q},\hat{u})\Big). \tag{7.4}$$

In view of (7.3) the first difference on the right-hand side of (7.4) goes to zero for $h \to 0$. For the last sum, in parenthesis, we notice

$$2J(\hat{q},\tilde{u}) - 2J(\hat{q},\hat{u}) = \|\tilde{u} - u_d\|_2^2 - \|\hat{u} - u_d\|_2^2\,. \tag{7.5}$$

As already shown, $\overline{u}_h \to u$ in $W^{1,p}(\Omega)$, hence $\hat{u} \to u$ in $W^{1,p}(\Omega)$. Moreover, as $\overline{q}_h \to q$ in $W^{-1,p'}(\Omega)$, we have $\hat{q} \to q$ in $W^{-1,p'}(\Omega)$, and thus from Lemma 3.2

$$\tilde{u} = u(\hat{q}) \to u = u(q) \text{ strongly in } W^{1,p}(\Omega).$$

By our assumption on p, $p \ge \frac{2d}{d+2}$, we therefore obtain

$$\tilde{u} \to u \text{ strongly in } L^2(\Omega), \qquad \hat{u} \to u \text{ strongly in } L^2(\Omega)\,.$$

Combining this, (7.5), (7.3) and (7.4), we conclude $\overline{q}_h \to q$ strongly in $L^2(\Omega)$. □

In order to prove rates of convergence, we follow the approach presented in [36]. First let us deal with the variational discretization, i.e., only the state space is discretized. For brevity of presentation, we name this problem

(P_s) Minimize $J(q_h, u_h)$ subject to (4.8) and $(q_h, u_h) \in \mathcal{Q}_{\mathrm{ad}} \times \mathcal{V}_h$.

Theorem 7.2 (Convergence rates for variational discretization)**.** *For $\varepsilon \in [0,\infty)$ and $p \in [\frac{2d}{d+2}, \infty)$ let Assumptions 2.1 and 2.4 be satisfied. For each $h > 0$ let $(\overline{q}_h, \overline{u}_h) \in \mathcal{Q}_{\mathrm{ad}} \times \mathcal{V}_h$ be a global solution of the semi-discretized optimization problem (P_s) and let $(\overline{q}, \overline{u}) \in \mathcal{Q}_{\mathrm{ad}} \times \mathcal{V}$, be a global solution of (P). Then there exists a constant $c > 0$ independent of h such that*

$$|J(\overline{q},\overline{u}) - J(\overline{q}_h, \overline{u}_h)| \le c h^{\min\left\{1, \frac{2}{p}\right\}}\,. \tag{7.6}$$

Proof. We define $u_h \in \mathcal{V}_h$ the solution to (4.8), for the control $\overline{q}$. Corollary 6.3 provides us for $p \in (1,\infty)$ with the estimate

$$\|\nabla \overline{u} - \nabla u_h\|_p \le c h^{\min\left\{1, \frac{2}{p}\right\}}\,.$$

The continuous embedding $W^{1,p}(\Omega) \subset L^2(\Omega)$ for $p \ge \frac{2d}{d+2}$ implies

$$\|\overline{u} - u_h\|_2 \le c\|\overline{u} - u_h\|_{1,p} \le c h^{\min\left\{1, \frac{2}{p}\right\}}\,. \tag{7.7}$$

We notice the following elementary inequality that holds for all $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\eta} \in \mathbb{R}^d$:

$$\big||\boldsymbol{\xi}_1 - \boldsymbol{\eta}|^2 - |\boldsymbol{\xi}_2 - \boldsymbol{\eta}|^2\big| = |(\boldsymbol{\xi}_1 + \boldsymbol{\xi}_2 - 2\boldsymbol{\eta}) \cdot (\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2)| \le 2\left(|\boldsymbol{\xi}_1| + |\boldsymbol{\xi}_2| + |\boldsymbol{\eta}|\right)|\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2|$$

From this and (7.7) we infer the estimate

$$\begin{aligned} |J(\overline{q},\overline{u}) - J(\overline{q},u_h)| &= \left|\frac{1}{2}\int_\Omega |\overline{u} - u_d|^2\,\mathrm{d}x - \frac{1}{2}\int_\Omega |u_h - u_d|^2\,\mathrm{d}x\right| \\ &\le \int_\Omega \left(|\overline{u}| + |u_h| + |u_d|\right)|\overline{u} - u_h|\,\mathrm{d}x \\ &\le \left(\|\overline{u}\|_2 + \|u_h\|_2 + \|u_d\|_2\right)\|\overline{u} - u_h\|_2 \\ &\le c h^{\min\left\{1, \frac{2}{p}\right\}}\,, \end{aligned} \tag{7.8}$$

where for the last inequality we have also used (3.2) and (4.9). As the pair $(\overline{q}, u_h)$ is admissible for (P_s), the inequality

$$J(\overline{q}_h, \overline{u}_h) \le J(\overline{q}, u_h) \tag{7.9}$$

is fulfilled, consequently

$$J(\overline{q}_h, \overline{u}_h) - J(\overline{q}, \overline{u}) \le J(\overline{q}, u_h) - J(\overline{q}, \overline{u}) \overset{(7.8)}{\le} ch^{\min\left\{1, \frac{2}{p}\right\}}. \tag{7.10}$$

Note that $\|\overline{q}_h\|_{\max\{p',2\}} \le C$ uniformly in $h \in (0, 1]$. In order to obtain the reverse inequality of (7.10), starting from $(\overline{q}_h, \overline{u}_h)$ we construct $(\overline{q}_h, \hat{u})$ by defining $\hat{u} \in \mathcal{V}$ as the solution to (3.1). Note that $(\overline{q}_h, \hat{u})$ are feasible for the exact optimal control problem (P), although both $\overline{q}_h$ and $\hat{u}$ depend on h. As a result, there holds

$$J(\overline{q}, \overline{u}) \le J(\overline{q}_h, \hat{u}). \tag{7.11}$$

We can precisely use the same arguments as for (7.8) in order to obtain

$$|J(\overline{q}_h, \overline{u}_h) - J(\overline{q}_h, \hat{u})| \le ch^{\min\left\{1, \frac{2}{p}\right\}}. \tag{7.12}$$

Combining the inequalities (7.8), (7.9), (7.11) and (7.12), we finally arrive at

$$\begin{aligned} -ch^{\min\left\{1, \frac{2}{p}\right\}} &\overset{(7.8)}{\le} J(\overline{q}, \overline{u}) - J(\overline{q}, u_h) \overset{(7.9)}{\le} J(\overline{q}, \overline{u}) - J(\overline{q}_h, \overline{u}_h) \\ &\overset{(7.11)}{\le} J(\overline{q}_h, \hat{u}) - J(\overline{q}_h, \overline{u}_h) \overset{(7.12)}{\le} ch^{\min\left\{1, \frac{2}{p}\right\}}. \end{aligned}$$

This establishes the assertion. □

Now let us deal with the case that the control space is discretized. To this end, we adapt the theory presented in [36] to our situation. In order to quantify the order of convergence, one usually requires some regularity of the optimal control $\overline{q}$. In the linear setting, additional regularity of $\overline{q}$ can be proven by deriving additional regularity of the adjoint state z. As we have seen in our discussion in Section 5 such additional regularity can only be shown in the case of the variational discretization.

Theorem 7.3 (Convergence rates for piecewise constant controls)**.** *For $\varepsilon \in [0, \infty)$ and $p \in [\frac{2d}{d+2}, \infty)$ let Assumptions 2.1 and 2.4 be satisfied. For each $h > 0$ and $\mathcal{Q}_{h,\text{ad}} = \mathcal{Q}^0_{h,\text{ad}}$ let $\overline{q}_h \in \mathcal{Q}_{h,\text{ad}}$ be a discrete optimal control and $\overline{u}_h = u_h(\overline{q}_h) \in \mathcal{V}_h$ the corresponding discrete optimal state, i.e., $(\overline{q}_h, \overline{u}_h)$ are global solutions of (P_h). Further, let $(\overline{q}, \overline{u}) \in \mathcal{Q}_{\text{ad}} \times \mathcal{V}$, be a global solution of (P). Then there exists a constant $c > 0$ independent of h such that*

$$|J(\overline{q}, \overline{u}) - J(\overline{q}_h, \overline{u}_h)| \le ch^{\min\left\{1, \frac{1}{p-1}\right\}}. \tag{7.13}$$

Proof. We have already proven the existence of an accumulation point $(\overline{q}, \overline{u})$ in Theorem 7.1, and assume from now on that $(\overline{q}_h, \overline{u}_h)$ converges to this limit.

Let $(\hat{q}_h, \hat{u}_h) \in \mathcal{Q}_{h,\text{ad}} \times \mathcal{V}$ be a global solution of the following auxiliary problem in which only the control variable is discretized:

$$\text{Minimize } J(q_h, u_h) \quad \text{subject to (3.1) and } (q_h, u_h) \in \mathcal{Q}_{h,\text{ad}} \times \mathcal{V} \tag{7.14}$$

In order to derive the stated error estimate, we split the error as follows:

$$|J(\overline{q}, \overline{u}) - J(\overline{q}_h, \overline{u}_h)| \le |J(\overline{q}, \overline{u}) - J(\hat{q}_h, \hat{u}_h)| + |J(\hat{q}_h, \hat{u}_h) - J(\overline{q}_h, \overline{u}_h)| \tag{7.15}$$

By repeating the proof of Theorem 7.2, we can estimate the second term on the right hand side of (7.15) as

$$|J(\hat{q}_h, \hat{u}_h) - J(\overline{q}_h, \overline{u}_h)| \le ch^{\min\left\{1, \frac{2}{p}\right\}}. \tag{7.16}$$

This is possible as all constants appearing in this proof are only dependent of $\|\hat{q}_h\|_{\max\{p',2\}}$ (and the regularity of $\mathbb{T}_h$ and the characteristics of $\boldsymbol{S}$). Note that $\|\hat{q}_h\|_{\max\{p',2\}}$ is uniformly bounded in $h \in (0, 1]$.

Thus it is sufficient to estimate the first term on the right hand side of (7.15). To this end, we use again similar arguments as in the proof of Theorem 7.2. Let us set $q_h := \Pi_h \overline{q}$, where Π_h stands for the L^2-projection onto Q_h^0. It is clear that $\Pi_h \colon \mathcal{Q}_{\text{ad}} \to \mathcal{Q}_{h,\text{ad}}^0$.

Let $u_h \in \mathcal{V}$ be the solution to the state equation (3.1) for control q_h. From Lemma 3.2 we deduce the estimate

$$\|\nabla \overline{u} - \nabla u_h\|_p \lesssim \begin{cases} \|\varepsilon + |\nabla \overline{u}| + |\nabla u_h|\|_p^{2-p} \|\overline{q} - \Pi_h \overline{q}\|_{-1,p'} & \text{for } p \le 2, \\ \|\overline{q} - \Pi_h \overline{q}\|_{-1,p'}^{\frac{1}{p-1}} & \text{for } p \ge 2. \end{cases} \tag{7.17}$$

Due to the uniform a priori bounds (3.2), (4.9) and the stability of Π_h, in the case $p \le 2$ there exists a constant $C > 0$ independent of h such that

$$\|\varepsilon + |\nabla \overline{u}| + |\nabla u_h|\|_p^{2-p} \le C.$$

Employing Lemma 4.2, we can bound the right-hand side of (7.17) by

$$\|\nabla \overline{u} - \nabla u_h\|_p \le c h^{\min\left\{1, \frac{1}{p-1}\right\}} . \tag{7.18}$$

From the convexity of J we conclude

$$\begin{aligned} J(\overline{q}, \overline{u}) &\ge J(q_h, u_u) + \langle J'(q_h, u_h), (\overline{q} - q_h, \overline{u} - u_h) \rangle \\ &= J(q_h, u_u) + \alpha(q_h, \overline{q} - q_h)_\Omega + (u_h - u_d, \overline{u} - u_h)_\Omega . \end{aligned}$$

Since (q_h, u_h) is feasible for (7.14), the inequality

$$\begin{aligned} 0 \le J\left(\hat{q}_h, \hat{u}_h\right) - J\left(\overline{q}, \overline{u}\right) &\le J\left(q_h, u_h\right) - J\left(\overline{q}, \overline{u}\right) \\ &\le \alpha(q_h, q_h - \overline{q})_\Omega + (u_h - u_d, u_h - \overline{u})_\Omega \end{aligned}$$

follows. The last term on the right-hand side is bounded by

$$(u_h - u_d, u_h - \overline{u})_\Omega \le \|u_h - u_d\|_2 \|u_h - \overline{u}\|_2 \overset{(7.18)}{\le} c h^{\min\left\{1, \frac{1}{p-1}\right\}}$$

for $p \ge \frac{2d}{d+2}$. Moreover, there holds

$$(q_h, q_h - \overline{q})_\Omega = (q_h, \Pi_h \overline{q} - \overline{q})_\Omega = 0$$

due to the definition of the L^2-projection. To sum up, we get

$$0 \le J\left(\hat{q}_h, \hat{u}_h\right) - J\left(\overline{q}, \overline{u}\right) \le c h^{\min\left\{1, \frac{1}{p-1}\right\}} . \tag{7.19}$$

Combining (7.15), (7.16) and (7.19), we conclude the assertion, noting that $h^{\min\left\{1, \frac{1}{p-1}\right\}} \le h^{\min\left\{1, \frac{2}{p}\right\}}$. □

Acknowledgement The authors would like to thank the anonymous referees for their valuable comments helping to shorten and improve the manuscript.

W. Wollner is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Projektnummer 314067056

References

[1] J.W. Barrett and W.B. Liu. Quasi-norm error bounds for the finite element approximation of a non-Newtonian flow. *Numer. Math.*, 68:437–456, 1994.

[2] L. Belenki, L. C. Berselli, L. Diening, and M. Růžička. On the finite element approximation of p-Stokes systems. *SIAM J. Numer. Anal.*, (2):373–397, 2012.

[3] L. Belenki, L. Diening, and C. Kreuzer. Optimality of an adaptive finite element method for the p-Laplacian equation. *IMA J. Numer. Anal.*, 32(2):484–510.

[4] L. C. Berselli, L. Diening, and M. Růžička. Existence of strong solutions for incompressible fluids with shear dependent viscosities. *J. Math. Fluid Mech.*, 12:101–132, 2010.

[5] T. Bhattacharya and F. Leonetti. A new Poincaré inequality and its application to the regularity of minimizers of integral functionals with nonstandard growth. *Nonlinear Anal.*, 17:833–839, 1991.

[6] L. Bonifacius and I. Neitzel. Second order optimality conditions for optimal control of quasilinear parabolic equations. *Math. Control Relat. Fields*, 8(1):1–34, 2018.

[7] S. Brenner and R. L. Scott. *The Mathematical Theory of Finite Element Methods*. Springer Verlag, Berlin Heidelberg New York, 1994.

[8] E. Casas and K. Chrysafinos. Analysis and optimal control of some quasilinear parabolic equations. *Math. Control Relat. Fields*, 8(3 & 4):607–623, 2018.

[9] E. Casas and V. Dhamo. Optimality conditions for a class of optimal boundary control problems with quasilinear elliptic equations. *Control and Cybernet.*, 40(2):457–490, 2011.

[10] E. Casas, P.I. Kogut, and G. Leugering. Approximation of optimal control problems in the coefficient for the p-Laplace equation. I. convergence result. *SIAM J. Control Optim.*, 54(3):1406–1422, 2016.

[11] E. Casas and F. Tröltzsch. First- and second-order optimality conditions for a class of optimal control problems with quasilinear elliptic equations. *SIAM J. Control Optim.*, 48(2):688–718, 2009.

[12] S.-S. Chow. Finite element error estimates for non-linear elliptic equations of monotone type. *Numer. Math.*, 54:373–393, 1989.

[13] A. Cianchi and V. Maz'ya. Second-Order Two-Sided Estimates in Nonlinear Elliptic Problems. *Arch. Rational Mech. Anal.*, 229:569–599, 2018.

[14] A. Cianchi and V. G. Maz'ya. Global Lipschitz regularity for a class of quasilinear elliptic equations. *Comm. Partial Differential Equations*, 36(1):100–133, 2010.

[15] P. G. Ciarlet. *The finite elements methods for elliptic problems*. North-Holland, 1980.

[16] F. Crispo, C. R. Grisanti, and P. Maremonti. On the high regularity of solutions to the p-Laplacian boundary value problem in exterior domains. *P. Annali di Matematica*, 195:821–834, 2016.

[17] K. Deckelnick and M. Hinze. Convergence of a finite element approximation to a state-constrained elliptic control problem. *SIAM J. Numer. Anal.*, 45(5):1937–1953, 2007.

[18] L. Diening, C. Ebmeyer, and M. Růžička. Optimal convergence for the implicit space-time discretization of parabolic systems with p-structure. *SIAM J. Numer. Anal.*, 45(2):457–472, 2007.

[19] L. Diening and F. Ettwein. Fractional estimates for non-differentiable elliptic systems with general growth. *Forum Math.*, 20:523–556, 2008.

[20] L. Diening, D. Kröner, M. Růžička, and I. Toulopoulos. A local discontinuous Galerkin approximation for systems with p-structure. *IMA J. Numer. Anal.*, 34(4):1447–1488, 2014.

[21] L. Diening and M. Růžička. Interpolation operators in Orlicz–Sobolev spaces. *Numer. Math.*, 107(1):107–129, 2007.

[22] L. Diening, B. Stroffolini, and A. Verde. Everywhere regularity of functionals with φ-growth. *Manuscripta Math.*, 129:449–481, 2009.

[23] C. Ebmeyer and W.B. Liu. Quasi-norm interpolation error estimates for piecewise linear finite element approximation of p-Laplacian problems. *Numer. Math.*, 100(204):233–258, 2005.

[24] B. Endtmayer, U. Langer, I. Neitzel, T. Wick, and W. Wollner. Multigoal-oriented optimal control problems with nonlinear PDE constraints. *Comput. Math. Appl.*, 79(10):3001–3026, 2020.

[25] C. Helanow and J. Ahlkrona. Stabilized equal low-order finite elements in ice sheet modeling – accuracay and robustness. *Computat. Geosci.*, 22:951–974, 2018.

[26] R. Herzog and W. Wollner. A conjugate direction method for linear systems in Banach spaces. *J. Inverse Ill-Posed Probl.*, 25(5):553–572, 2017.

[27] M. Hinze. A variational discretization concept in control constrained optimization: The linear-quadratic case. *Comput. Optim. Appl.*, 30(1):45–61, 2005.

[28] A. Hirn. Approximation of the p-Stokes equations with equal-order finite elements. *J. Math. Fluid Mech.*, 15:65–88, 2013.

[29] A. Hirn. Finite element approximation of singular power-law systems. *Math. Comp.*, 82:1247–1268, 2013.

[30] J. L. Lions. *Optimal Control of Systems Governed by Partial Differential Equations*. Die Grundlehren der mathematischen Wissenschaften. Springer, Berlin, 1971.

[31] W.B. Liu. Degenerate quasilinear elliptic equations arising from bimaterial problems in elastic–plastic mechanics. *Nonlin. Anal.*, 35:517–529, 1999.

[32] W.B. Liu and J.W. Barrett. A remark on the regularity of the solutions of the p-Laplacian and its applications to their finite element approximation. *J. math. Anal. Appl.*, 178(2):470–487, 1993.

[33] J. C. De los Reyes. *Numerical PDE-Constrained Optimization*. Springer International Publishing, 2015.

[34] J. Málek, J. Nečas, M. Rokyta, and M. Růžička. *Weak and measure-valued solutions to evolutionary PDEs*, volume 13 of *Applied Mathematics and Mathematical Computation*. Chapman and Hall, London, 1996.

[35] C. Meyer. Error estimates for the finite-element approximation of an elliptic control problem with pointwise state and control constraints. *Control Cybernet.*, 37(1):51–85, 2008.

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

[37] R. Rannacher. *Numerik 2. Numerik partieller Differentialgleichungen*. Lecture Notes Mathematik. Heidelberg University Publishing, 2017.

[38] M. Růžička. *Nichtlineare Funktionalanalysis*. Springer-Verlag, 2004.

[39] L. R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.*, 54(190):483–493, 1990.

[40] F. Tröltzsch. *Optimal control of partial differential equations*, volume 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010.

[41] E. Zeidler. *Nonlinear Functional Analysis and its Applications. II/B: Nonlinear Monotone Operators*. Springer-Verlag, New York, 1990.