

# RELIABLE AND EFFICIENT A POSTERIORI ERROR ESTIMATES FOR FINITE ELEMENT APPROXIMATIONS OF THE PARABOLIC $p$ -LAPLACIAN

CHRISTIAN KREUZER

ABSTRACT. We generalize the a posteriori techniques for the linear heat equation in [Ver03] to the case of the nonlinear parabolic  $p$ -Laplace problem thereby proving reliable and efficient a posteriori error estimates for a fully discrete implicit Euler Galerkin finite element scheme. The error is analyzed using the so-called quasi-norm and a related dual error expression. This leads to equivalence of the error and the residual, which is the key property for proving the error bounds.

## 1. INTRODUCTION

In this work we shall prove reliable and efficient residual based error bounds for fully discrete finite element approximations to the solution  $u : I \times \Omega \rightarrow \mathbb{R}$  of the problem

$$(1.1) \quad \begin{aligned} \partial_t u - \operatorname{div} \mathbf{A}(\nabla u) &= f && \text{in } I \times \Omega, \\ u(0) &= u_0 && \text{in } \Omega, \\ u &= 0 && \text{on } I \times \partial\Omega, \end{aligned}$$

where  $\Omega \subset \mathbb{R}^d$ ,  $d \in \mathbb{N}$ , is a bounded open polyhedral spatial domain and  $I := (0, T)$  is the temporal interval with some finite time  $T > 0$ . Precise assumptions on the initial datum  $u_0 : \Omega \rightarrow \mathbb{R}$  and the external force  $f : I \times \Omega \rightarrow \mathbb{R}$  will be specified later in §2.

The nonlinear vector field  $\mathbf{A} : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is supposed to have power-law structure. Common representatives are e.g. the power law or the Carreau law

$$(1.2) \quad \begin{aligned} \mathbf{A}(\nabla u) &= |\nabla u|^{p-1} \nabla u, \\ \mathbf{A}(\nabla u) &= (\nu_\infty + (\nu_0 - \nu_\infty)(\kappa^2 + |\nabla u|^2)^{\frac{p-2}{2}}) \nabla u, \end{aligned}$$

respectively. Hereby  $\kappa \geq 0$ ,  $\nu_0 > \nu_\infty \geq 0$ , and  $p \geq \frac{2d}{d+2}$ . The latter restriction on  $p$  is due to analytical reasons and ensures the unique weak solvability of the problem; compare with §2.

Finite element approximations of the steady state of (1.1), i. e., to the  $p$ -Laplacian problem, have been widely investigated. Primary a priori results, relying on energy norm techniques, yield suboptimal convergence rates; see e.g. [Cia78, GM75, Cho89]. Introducing the so called quasi-norm, Barrett and Liu overcame this drawback and proved sharp a priori bounds in [BL93a]. The advantage of their approach over energy norm techniques is, that in quasi-norms the residual is equivalent to the error; we elaborate on this issue in Lemma 11 and Remark 12 below. As for linear elliptic problems this implies that, up to a constant, the Galerkin approximation is a best-approximation and whence optimal convergence estimates could be deduced from interpolation estimates [EL05, DR07a].

---

2010 *Mathematics Subject Classification.* 35K55, 35K59, 65N60, 65M15.

*Key words and phrases.*  $p$ -Laplace, nonlinear parabolic pde, a posteriori error analysis, quasi-norm.

Thanks to its outstanding success in the a priori analysis of the nonlinear Laplace problem, the quasi-norm technique could be expected to play a crucial role in finite element estimates for nonlinear parabolic problems (1.1) as well. This expectation is supported by the results in [BL94, DER07].

Recently, the quasi-norm technique also found its way into the a posteriori analysis of the nonlinear Laplacian [LY01, LY02, CK03, CLY06a, DK08, BDK11]. Moreover, those estimators can be used to design adaptive finite element methods with optimal convergence rates; see [DK08, BDK11].

There have been also attempts to generalize those results to the a posteriori analysis of nonlinear parabolic problems in [CLY06b]. Although therein the concept of distance is based on the quasi-norm, it appears to be too weak. It particularly lacks the control of the temporal derivative. As a consequence the obtained lower bounds are somewhat suboptimal. In the linear case this issue has been solved by Verfürth [Ver03, Ver05] proving fully reliable and efficient a posteriori error bounds. It is the purpose of this work to extend this approach, by means of the quasi-norm technique, to parabolic problems with power law structure. We would like mention that in the linear case the temporal derivative of a given function is controlled in the dual norm of some Sobolev space. Since we deal with quasi-norms, which are not norms, we need to generalize this error concept to a corresponding dual error concept. This is the key for a generalization of the a posteriori analysis in [Ver03] for the linear case. The plan of the paper is as follows.

In section §2 we revisit problem (1.1) in a more general setting. Based on so-called N-functions we state the weak formulation of (1.1). The benefit of this general setting is twofold. On the one hand it simplifies the proofs significantly and on the other hand it allows us to simultaneously treat the cases  $p \in [2, \infty)$  and  $p \in (1, 2)$  as well as miscellaneous power law like models. Section §3 is concerned with introducing the above mentioned error concept, which finally leads to a first abstract a posteriori error estimate, namely the equivalence of error and residual. This is then exploited in §4 in order to prove reliable and efficient estimates for the error of a standard implicit Euler Galerkin finite element method. We finish the work in section §5 dealing with the computability of the error bound and applying the results to the concrete case of the power law (1.2).

## 2. THE WEAK FORMULATION AND PRELIMINARY RESULTS

In this section we shall introduce so called N-functions and related Orlicz–Sobolev spaces. Among others, this abstract setting covers all nonlinearities in (1.2) for all  $p \in (1, \infty)$ . Then, having the analytical framework at hand, we introduce the weak formulation of (1.1). We conclude this section by collecting some preliminary facts on N-functions.

**2.1. The Weak Formulation.** A convenient way of treating most nonlinearities with power law structure is to utilize the concept of so-called N-functions. In particular, a ‘nice’ Young function, termed an N-function, is a continuous, convex, and strictly monotone function  $\phi : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ , such that

$$\phi(0) = 0 \text{ and } \phi(s) > 0, \text{ if } s > 0, \quad \lim_{s \rightarrow 0} \frac{\phi(s)}{s} = 0, \quad \text{and} \quad \lim_{s \rightarrow \infty} \frac{\phi(s)}{s} = \infty,$$

where  $\mathbb{R}^+$  denotes the non-negative real semi-axis. For more detailed information on N-functions and related Orlicz and Orlicz-Sobolev spaces consider e. g. [RR91, KK91, KR61, Mus83, Kre08].

Thanks to the convexity, for each N-function  $\phi$  there exists a unique nondecreasing and right continuous function  $\phi' : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  such that  $\phi'(0) = 0$  and  $\int_0^s \phi'(r) dr = \phi(s)$ ; see [RR91]. However, for a quantitative numerical approach, we

require more regularity: In particular, we assume

$$(2.1) \quad \phi \in C^1([0, \infty)) \cap C^2((0, \infty)) \quad \text{and} \quad c s \phi''(s) \leq \phi'(s) \leq C s \phi''(s), \quad s \geq 0,$$

for some constants  $C, c > 0$ .

An N-function  $\phi$  is said to satisfy the  $\Delta_2$ -condition if there exists a constant  $K$  such that  $\phi(2s) \leq K\phi(s)$ . The smallest constant is then denoted by  $\Delta_2(\phi)$  and for a family of N-functions  $\{\phi_a\}_{a \in A}$  we define  $\Delta(\{\phi_a\}_{a \in A}) := \max_{a \in A} \Delta_2(\phi_a)$ . Let  $\phi$  satisfy assumption (2.1), therefore  $(\ln \phi'(s))' \leq \frac{1}{cs}$  and thus

$$(2.2) \quad \ln \phi'(2s) - \ln \phi'(s) \leq \int_s^{2s} \frac{1}{cr} dr = \frac{\ln 2}{c} \quad \Rightarrow \quad \phi'(2s) \leq 2^{1/c} \phi'(s).$$

By the fundamental theorem of calculus and the transformation theorem, this implies that an N-functions  $\phi$  satisfying assumption (2.1) already satisfies the  $\Delta_2$ -condition.

For a measurable subset  $\omega \subset \mathbb{R}^d$ , we denote the classical Orlicz and Orlicz-Sobolev spaces by  $L^\phi(\omega)$  and  $W^{1,\phi}(\omega)$ , respectively. In other words,  $f \in L^\phi(\omega)$  if  $\int_\omega \phi(|f|) dx < \infty$  and  $f \in W^{1,\phi}(\omega)$  if  $f, \nabla f \in L^\phi(\omega)$ . The so-called Luxemburg norm  $\|f\|_{(\phi),\omega} := \inf\{\lambda > 0 : \int_\omega \phi(|f|/\lambda) dx \leq 1\}$  turns  $L^\phi(\omega)$  into a reflexive Banach space and  $W^{1,\phi}(\omega)$  is a reflexive Banach space with the norm  $\|\cdot\|_{1,(\phi),\omega} := \|\cdot\|_{(\phi),\omega} + \|\nabla \cdot\|_{(\phi),\omega}$ . By  $W_0^{1,\phi}(\omega)$  we denote the closure of  $C_0^\infty(\omega)$  in  $W^{1,\phi}(\omega)$ . As usual  $L^p(\omega)$ ,  $W^{1,p}(\Omega)$ , and  $W_0^{1,p}(\omega)$ ,  $p \in [1, \infty]$ , denote the common Lebesgue and Sobolev spaces with norms  $\|\cdot\|_{p;\omega}$  and  $\|\cdot\|_{1,p;\omega}$ . For  $p \in (1, \infty)$  and  $\frac{1}{p} + \frac{1}{p'} = 1$  the dual of  $W_0^{1,p}(\omega)$  is denoted by  $W_0^{-1,p'}(\omega)$  with norm  $\|\cdot\|_{-1,p';\omega}$ . For  $\omega = \Omega$  we skip the domain in the notion of the norm, e.g.  $\|\cdot\|_p = \|\cdot\|_{p,\Omega}$ . We denote  $\langle v, w \rangle_\Omega = \int_\Omega vw dx$  and by  $\langle \cdot, \cdot \rangle$  all dual pairings regardless of the space-pairings that are considered.

We define the nonlinear vector-field  $\mathbf{A} : \mathbb{R}^d \rightarrow \mathbb{R}^d$  by

$$(2.3) \quad \mathbf{A}(\mathbf{Q}) := \phi'(|\mathbf{Q}|) \frac{\mathbf{Q}}{|\mathbf{Q}|} \quad \text{for all } \mathbf{Q} \in \mathbb{R}^d.$$

For the weak formulation of the problem it is crucial that the spaces  $W_0^{1,\phi}(\Omega)$  together with  $L^2(\Omega)$ , and its dual  $(W_0^{1,\phi}(\Omega))^*$  form a Gelfand triple. For this reason we assume there exists  $C_0, c_0 > 0$ ,  $s_0 \geq 0$  and  $p \geq \frac{2d}{d+2}$  such that

$$(2.4) \quad c_0 s^p \leq \phi(s) \leq C_0 s^p \quad \text{for all } s \geq s_0.$$

Consequently, we get with  $\frac{1}{p} + \frac{1}{p'} = 1$  that

$$W_0^{1,\phi}(\Omega) = W_0^{1,p}(\Omega) \hookrightarrow L^2(\Omega) \hookrightarrow W^{-1,p'}(\Omega) = (W_0^{1,\phi}(\Omega))^*;$$

compare e.g. with [KR61, §3]. Let  $L^p(I; X)$  be the usual Bochner space with norm  $\|\cdot\|_{L^p(I; X)}$ , where  $X$  is some Banach space. We can thus define the weak solution space for (1.1) by

$$(2.5) \quad \mathbb{W}(I) := \{v \in L^p(I; W_0^{1,p}(\Omega)) \mid \partial_t v \in L^{p'}(I; W^{-1,p'}(\Omega))\} \hookrightarrow C(I; L^2(\Omega)).$$

Here the derivative  $\partial_t v$  has to be understood in a distributional sense; compare with [Zei90, GGZ74]. Note that we suppress the dependence of functions on variables  $t, x$  if there is no risk of confusion. For  $\tilde{I} \subset \mathbb{R}$  being an open interval define

$$\|v\|_{\mathbb{W}(\tilde{I})} := \|v\|_{L^p(\tilde{I}; W_0^{1,p}(\Omega))} + \|\partial_t v\|_{L^{p'}(\tilde{I}; W^{-1,p'}(\Omega))} + \text{ess sup}_{t \in \tilde{I}} \|v(t)\|_2,$$

for  $v \in \mathbb{W}(\tilde{I})$ , whence  $\mathbb{W}(I)$  becomes a reflexive Banach space with the norm  $\|\cdot\|_{\mathbb{W}(I)}$ .

Now we are in the position to state the weak formulation of (1.1): For  $f \in L^{p'}(I; L^{p'}(\Omega)) \hookrightarrow L^{p'}(I; W^{-1,p'}(\Omega))$  and  $u_0 \in L^2(\Omega)$ , find  $u \in \mathbb{W}(I)$  such that, a.e.

in  $I$ ,

$$(2.6) \quad \begin{aligned} \langle \partial_t u, w \rangle + \langle \mathbf{A}(\nabla u), \nabla w \rangle_\Omega &= \langle f, w \rangle_\Omega & \text{for all } w \in W_0^{1,p}(\Omega), \\ u(0) &= u_0 & \text{in } L^2(\Omega). \end{aligned}$$

This is equivalent to

$$(2.7) \quad \int_I \langle \partial_t u, w \rangle + \langle \mathbf{A}(\nabla u), \nabla w \rangle_\Omega dt + \langle u(0), w_0 \rangle_\Omega = \int_I \langle f, w \rangle_\Omega dt + \langle u_0, w_0 \rangle_\Omega$$

for all  $(w, w_0) \in L^p(I; W_0^{1,p}(\Omega)) \times L^2(\Omega)$ .

By [DE08, Lemma 20] we have for all  $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^d$  that

$$(2.8) \quad \begin{aligned} (\mathbf{A}(\mathbf{P}) - \mathbf{A}(\mathbf{Q})) \cdot (\mathbf{P} - \mathbf{Q}) &\gtrsim \phi''(|\mathbf{P}| + |\mathbf{Q}|) |\mathbf{P} - \mathbf{Q}|^2, \\ |\mathbf{A}(\mathbf{P}) - \mathbf{A}(\mathbf{Q})| &\lesssim \phi''(|\mathbf{P}| + |\mathbf{Q}|) |\mathbf{P} - \mathbf{Q}| \end{aligned}$$

and thus the operator  $-\operatorname{div} \mathbf{A}(\nabla \cdot) : W_0^{1,p}(\Omega) \rightarrow W^{-1,p'}(\Omega)$  is strictly monotone. Consequently existence and uniqueness of the solution  $u \in \mathbb{W}(I)$  follow by means of the theory of monotone operators; see [GGZ74].

It is worth noticing that for  $p \geq \frac{2d}{d+2}$  the power law as well as the Carreau law perfectly fit into the introduced framework by defining

$$\phi(s) := \frac{1}{p} s^p \quad \text{respectively} \quad \phi(s) := \int_0^s (\nu_\infty + (\nu_0 - \nu_\infty)(\kappa^2 + r^2)^{\frac{p-2}{2}}) r dr.$$

We will come back to this issue later in Section §5.2.

**2.2. Additional Facts on N-Functions.** In this section we collect some well-known facts on N-functions and introduce the so-called shifted N-functions, which play a crucial role in the subsequent analysis. For detailed proofs we refer to the literature, e.g. [RR91, KK91, KR61, Mus83, DE08, Kre08].

N-functions come in mutually complementary pairs. In particular, for an N-function  $\phi$  we can define its dual by

$$(2.9) \quad \phi^*(s) := \max\{rs - \phi(r) : r \geq 0\}.$$

It follows that  $\phi^*$  itself is an N-function and  $(\phi^*)^* = \phi$ . If  $\phi$  satisfies (2.1) then also  $\phi^*$  does; see [DE08, Kre08]. Therefore, thanks to (2.2),  $\phi^*$  also satisfies the  $\Delta_2$ -condition. An immediate consequence of the definition of the complementary function is the Young inequality

$$(2.10) \quad sr \leq \phi(s) + \phi^*(r) \quad \text{for all } s, r \geq 0,$$

where equality is obtained if  $r = \phi'(s)$  or  $s = (\phi^*)'(r)$ . Thanks to the  $\Delta_2$ -condition, a scaled Young inequality also holds, i.e., for each  $\delta > 0$  there exists a constant  $C_\delta > 0$  depending on  $\Delta_2(\phi)$  (and hence on the constants in (2.1)), such that

$$(2.11) \quad sr \leq C_\delta \phi(s) + \delta \phi^*(r) \quad \text{for all } s, r \geq 0.$$

Note that by means of the  $\Delta_2$ -condition  $(L^\phi(\Omega))^* = L^{\phi^*}(\Omega)$ . In analogy to the dual Sobolev spaces  $W^{-1,p'}(\Omega)$  we denote by  $W^{-1,\phi^*}(\Omega)$  the dual of  $W_0^{1,\phi}(\Omega)$ .

**Proposition 1.** *Let  $\phi, \psi$  be N-functions. Then, for all  $s \geq 0$ ,*

$$(2.12a) \quad \phi(\alpha s) \leq \alpha \phi(s) \quad \text{for all } \alpha \in [0, 1],$$

$$(2.12b) \quad \frac{s}{2} \phi'\left(\frac{s}{2}\right) \leq \phi(s) \leq s \phi'(s),$$

$$(2.12c) \quad s \leq (\phi^*)^{-1}(s) \phi^{-1}(s) \leq 2s,$$

$$(2.12d) \quad \phi\left(\frac{\phi^*(s)}{s}\right) \leq \phi^*(s) \leq \phi\left(2 \frac{\phi^*(s)}{s}\right),$$

$$(2.12e) \quad \phi(s) \leq \psi(s) \Rightarrow \psi^*(s) \leq \phi^*(s).$$

For N-functions  $\phi$  satisfying the  $\Delta_2$ -condition, (2.12b) and (2.12d) imply that there exists constants  $c, C > 0$ , such that

$$(2.13) \quad c\phi(s) \leq s\phi'(s) \leq C\phi(s) \quad \text{and} \quad c\phi^*(s) \leq \phi((\phi^*)'(s)) \leq C\phi^*(s)$$

for all  $s \geq 0$ . For each constant  $C \leq 2^k$ ,  $k \in \mathbb{N}$  we have

$$(2.14) \quad \phi(Ct) \leq \Delta_2(\phi)^k \phi(t) \quad \text{for all } t \geq 0,$$

and by the convexity of  $\phi$  it follows that

$$(2.15) \quad \phi(s+t) \leq \frac{1}{2}\phi(2t) + \frac{1}{2}\phi(2s) \leq \frac{\Delta_2(\phi)}{2}(\phi(s) + \phi(t)) \quad \text{for all } s, t \geq 0.$$

The  $\Delta_2$ -condition of an N-function also affects its dual function. In particular, for all constants  $\Delta_2(\phi^*) \leq C < \infty$ , we have that

$$(2.16) \quad \phi(s) \leq \frac{1}{C}\phi\left(\frac{C}{2}s\right), \quad s \geq 0.$$

The smallest such constant is called  $\nabla_2(\phi) = \Delta_2(\phi^*)$ .

### 3. A BALANCED ERROR CONCEPT

In order to prove reliable and efficient residual based estimates we need an error concept that complies with the nonlinear structure of the problem. For that reason we shall introduce the so called quasi-norm and define a related dual quasi-norm. Based on this error concept we can finally conclude the equivalence of the error and the residual following the approach of Verfürth in [Ver03].

For the rest of this paper we use the notation  $v \lesssim w$  to indicate  $v \leq Cw$ , with a generic constant  $C > 0$  solely depending on some fixed parameters like the constants in (2.1) and the domain  $\Omega$ . We denote  $v \lesssim w \lesssim v$  as  $v \approx w$ .

**3.1. Shifted N-functions and the Quasi-Norm.** Several studies on steady partial differential equations with power law structure show that the quasi-norm is the optimal expression to deal with the error of nonlinear problems with power law structure; see e.g. [BL93a, BL93b, DE08, DK08, Kre08, BDK11].

For a fixed  $\phi$  satisfying (2.1), we introduce the family of shifted N-functions  $\{\phi_a\}_{a \geq 0}$  by

$$\phi'_a(s) := \frac{\phi'(a+s)}{a+s} s \quad \text{and} \quad \phi_a(s) := \int_0^s \phi'_a(r) dr.$$

Those functions are N-functions as well. In the following we present properties of this family of N-functions without proof. The interested reader is referred to [DE08, DK08, Kre08, Kre12, DR07b].

**Lemma 2.** *Let  $\phi$  be an N-function that satisfies (2.1). Then  $\Delta_2(\{\phi_a, (\phi_a)^*\}_{a \geq 0})$  is uniformly bounded only depending on  $\Delta_2(\{\phi, \phi^*\})$  and thus on the constants in (2.1). Moreover, for all  $s \geq 0$  we have*

$$(\phi^*)_{\phi'(a)}(s) \approx (\phi_a)^*(s).$$

We next introduce the nonlinear vector field  $\mathbf{F} : \mathbb{R}^d \rightarrow \mathbb{R}^d$  by

$$(3.1) \quad \mathbf{F}(\mathbf{Q}) := \sqrt{|\mathbf{A}(\mathbf{Q})||\mathbf{Q}|} \frac{\mathbf{Q}}{|\mathbf{Q}|} \approx \sqrt{\phi(|\mathbf{Q}|)} \frac{\mathbf{Q}}{|\mathbf{Q}|} \quad \text{for } \mathbf{Q} \in \mathbb{R}^d.$$

The following Lemma from [DE08] is concerned with the relation between  $\mathbf{A}$ ,  $\mathbf{F}$ , and  $\{\phi_a\}_{a \geq 0}$ .

**Proposition 3.** *Let  $\phi$  be an N-function that satisfies (2.1). For all  $\mathbf{Q}, \mathbf{P} \in \mathbb{R}^d$  we have that*

$$\begin{aligned} (\mathbf{A}(\mathbf{P}) - \mathbf{A}(\mathbf{Q})) : (\mathbf{P} - \mathbf{Q}) &\approx |\mathbf{F}(\mathbf{P}) - \mathbf{F}(\mathbf{Q})|^2 \\ &\approx \phi''(|\mathbf{P}| + |\mathbf{P} - \mathbf{Q}|) |\mathbf{P} - \mathbf{Q}|^2 \\ &\approx \phi_{|\mathbf{P}|}(|\mathbf{P} - \mathbf{Q}|); \end{aligned}$$

and

$$|\mathbf{A}(\mathbf{P}) - \mathbf{A}(\mathbf{Q})| \approx \phi'_{|\mathbf{P}|}(|\mathbf{P} - \mathbf{Q}|).$$

This directly implies the next result.

**Corollary 4** (Quasi-Norm). *Let  $\phi$  be an N-function that satisfies (2.1). Then, for all  $v, w \in W_0^{1,\phi}(\Omega)$  we have*

$$\begin{aligned} \int_{\Omega} (\mathbf{A}(\nabla v) - \mathbf{A}(\nabla w)) \cdot (\nabla v - \nabla w) \, dx \\ &\approx \|\mathbf{F}(\nabla v) - \mathbf{F}(\nabla w)\|_2^2 \\ &\approx \int_{\Omega} \phi''(|\nabla v| + |\nabla(v - w)|) |\nabla v - \nabla w|^2 \, dx \\ &\approx \int_{\Omega} \phi_{|\nabla v|}(|\nabla v - \nabla w|) \, dx =: \|v - w\|_{(\nabla v)}^2. \end{aligned}$$

The three equivalent quantities of Corollary 4 naturally arise from the nonlinearity of the problem. For historical reasons we refer to them as the *quasi-norm* of  $v - w$ ; cf. also Remark 7. We note that  $W_0^{1,\phi}(\Omega) = W_0^{1,\phi_{|\nabla v|}}(\Omega)$  and that in particular the topology induced to  $W_0^{1,\phi}(\Omega)$  by the quasi norm is equivalent to the topology induced by the norm  $\|\cdot\|_{1,(\phi)}$ . The expression  $\mathbf{d}(v, w) := \|\mathbf{F}(\nabla v) - \mathbf{F}(\nabla w)\|_2$  is a metric on  $W_0^{1,\phi}(\Omega)$ .

The next results deal with the question how the N-functions of the family  $\{\phi_a\}_{a \geq 0}$  are related to each other.

**Lemma 5.** *Let  $\phi$  be an N-function satisfying (2.1). Then for all  $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^d$  and  $t \geq 0$  we have*

$$\phi'_{|\mathbf{P}|}(t) \lesssim \phi'_{|\mathbf{Q}|}(t) + \phi'_{|\mathbf{P}|}(|\mathbf{P} - \mathbf{Q}|).$$

Moreover, for all  $\delta > 0$  there exists a constant  $C_\delta > 0$  such that, for all  $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^d$ ,

$$\phi_{|\mathbf{P}|}(s) \lesssim (1 + C_\delta) \phi_{|\mathbf{Q}|}(s) + \delta \phi_{|\mathbf{P}|}(|\mathbf{P} - \mathbf{Q}|)$$

and

$$(\phi_{|\mathbf{P}|})^*(s) \approx (1 + C_\delta) (\phi_{|\mathbf{Q}|})^*(s) + \delta \phi_{|\mathbf{P}|}(|\mathbf{P} - \mathbf{Q}|).$$

**Corollary 6.** *Let  $\phi$  be an N-function that satisfies (2.1). We then have for all  $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^d$  that*

$$\phi'_{|\mathbf{P}|}(|\mathbf{P} - \mathbf{Q}|) \approx \phi'_{|\mathbf{Q}|}(|\mathbf{P} - \mathbf{Q}|) \quad \text{and} \quad \phi_{|\mathbf{P}|}(|\mathbf{P} - \mathbf{Q}|) \approx \phi_{|\mathbf{Q}|}(|\mathbf{P} - \mathbf{Q}|).$$

**Remark 7.** A distance measure named *quasi-norm* was first introduced by Barrett and Liu in [BL93a, BL93b] into a priori analysis of finite elements for nonlinear problems. For this error concept they proved optimal a priori error estimates. In particular, for the power law  $\phi(s) = \frac{1}{p} s^p$  with  $p \in (1, \infty)$ , they use

$$\int_{\Omega} (|\nabla u| + |\nabla u - \nabla v|)^{p-2} |\nabla u - \nabla v|^2 \, dx$$

which in turn is equivalent to

$$\approx \int_{\Omega} \phi''(|\nabla u| + |\nabla(u-v)|) |\nabla u - \nabla v|^2 dx.$$

Consequently, our approach generalizes this concept to  $N$ -functions satisfying assumption (2.1).

**3.2. The Dual Quasi-Norm.** We recall that, for  $v \in \mathbb{W}(I)$ , we have that  $\partial_t v \in W^{-1,\phi^*}(\Omega)$  a.e. in  $I$ . Consequently, in order to control the temporal derivative, a dual distance on  $W^{-1,\phi^*}(\Omega)$  is required.

To motivate this so-called dual quasi-norm we first take a closer look at the dual norm  $\|\cdot\|_{-1,p'}$  on  $W^{-1,p'}(\Omega)$ ,  $p \in (1, \infty)$  and  $\frac{1}{p} + \frac{1}{p'} = 1$ . We observe that the  $N$ -functions  $s \mapsto \frac{1}{p}s^p$  and  $s \mapsto \frac{1}{p'}s^{p'}$  are dual in the sense of (2.9). From this one can deduce that

$$\frac{1}{p'} \|R\|_{-1,p'}^{p'} = \sup_{w \in W_0^{1,p}(\Omega)} \langle R, w \rangle - \frac{1}{p} \|w\|_{1,p}^p$$

for  $R \in W^{-1,p'}(\Omega)$ . Recalling Corollary 4 for a fixed  $N$ -function  $\phi$  satisfying (2.1), this expression can be easily generalized to the quasi-norm context by defining

$$\begin{aligned} \|R\|_{*,(\nabla u)}^2 &:= \sup_{w \in W_0^{1,\phi}(\Omega)} \langle R, w \rangle - \|w\|_{(\nabla u)}^2 \\ (3.2) \quad &= \sup_{w \in W_0^{1,\phi}(\Omega)} \langle R, w \rangle - \int_{\Omega} \phi_{|\nabla u|}(|\nabla w|) dx, \end{aligned}$$

for  $R \in W^{-1,\phi^*}(\Omega)$  and  $u \in W_0^{1,\phi}(\Omega)$ . Note the analogy to the definition of the dual  $N$ -function (2.9). The operator on the right-hand side is Fréchet differentiable and thanks to the theory of monotone operators its supremum is actually attained at  $w_R \in W_0^{1,\phi}(\Omega)$  satisfying the nonlinear equation

$$(3.3) \quad -\operatorname{div} \phi'_{|\nabla u|}(\nabla w_R) \frac{\nabla w_R}{|\nabla w_R|} = R \quad \text{in } W^{-1,\phi^*}(\Omega).$$

It turns out that the dual quasi-norm of  $R$  is equivalent to the quasi-norm of  $w_R$ . In particular, thanks to (2.10), (3.3), and (2.13), we have that

$$\begin{aligned} \|R\|_{*,(\nabla u)}^2 &= \langle R, w_R \rangle - \int_{\Omega} \phi_{|\nabla u|}(|\nabla w_R|) dx \\ (3.4) \quad &= \int_{\Omega} (\phi_{|\nabla u|})^* (\phi'_{|\nabla u|}(|\nabla w_R|)) dx \approx \int_{\Omega} \phi_{|\nabla u|}(|\nabla w_R|) dx. \end{aligned}$$

Some properties of the dual quasi-norm are collected in the following statements.

**Lemma 8** (Convexity). *Let the  $N$ -function  $\phi$  satisfy (2.1). For  $R, S \in W^{-1,\phi^*}(\Omega)$  and  $u \in W_0^{1,\phi}(\Omega)$ , we have that*

$$\|(1-\theta)R + \theta S\|_{*,(\nabla u)} \lesssim (1-\theta)\|R\|_{*,(\nabla u)} + \theta\|S\|_{*,(\nabla u)}$$

for all  $\theta \in [0, 1]$ .

*Proof.* Thanks to the definition of the dual quasi-norm we have

$$\begin{aligned} \|(1-\theta)R + \theta S\|_{*,(\nabla u)}^2 &= \sup_{h \in W_0^{1,\phi}(\Omega)} \langle (1-\theta)R + \theta S, h \rangle - \int_{\Omega} \phi_{|\nabla u|}(|\nabla h|) dx \\ &\leq (1-\theta) \left\{ \sup_{h \in W_0^{1,\phi}(\Omega)} \langle R, h \rangle - \int_{\Omega} \phi_{|\nabla u|}(|\nabla h|) dx \right\} \\ &\quad + \theta \left\{ \sup_{h \in W_0^{1,\phi}(\Omega)} \langle S, h \rangle - \int_{\Omega} \phi_{|\nabla u|}(|\nabla h|) dx \right\}. \end{aligned}$$

This is the assertion.  $\square$

**Lemma 9.** *Let the  $N$ -function  $\phi$  satisfy (2.1). For  $R \in W^{-1,\phi^*}(\Omega)$  we have*

$$\|2R\|_{*,(\nabla u)}^2 \leq C \|R\|_{*,(\nabla u)}^2,$$

with  $C = \Delta_2(\{(\phi_a)^*\}_{a \geq 0}) < \infty$  solely depending on the constants in (2.1).

*Proof.* Thanks to Lemma 2 we have  $C < \infty$  and with (2.16) it follows that

$$\begin{aligned} \sup_{h \in W_0^{1,\phi}(\Omega)} \langle R, 2h \rangle - \int_{\Omega} \phi_{|\nabla u|}(|\nabla h|) dx \\ \leq \sup_{h \in W_0^{1,\phi}(\Omega)} \langle R, 2h \rangle - \int_{\Omega} C \phi_{|\nabla u|}\left(\frac{2}{C} |\nabla h|\right) dx \\ = C \left\{ \sup_{h \in W_0^{1,\phi}(\Omega)} \langle R, \frac{2}{C} h \rangle - \int_{\Omega} \phi_{|\nabla u|}\left(\frac{2}{C} |\nabla h|\right) dx \right\} = C \|R\|_{*,(\nabla u)}^2. \end{aligned}$$

This proves the assertion.  $\square$

Combining Lemmas 8 and 9 we can conclude the following result.

**Corollary 10** (Quasi-Triangle Inequality). *Let  $\phi$  be an  $N$ -function satisfying (2.1). For  $R, S \in W^{-1,\phi^*}(\Omega)$ ,  $u \in W_0^{1,\phi}(\Omega)$ , we have that*

$$\|R + S\|_{*,(\nabla u)} \lesssim \|R\|_{*,(\nabla u)} + \|S\|_{*,(\nabla u)}.$$

The constants hidden in  $\lesssim$  only depend on the constants in (2.1).

*Proof.* Thanks to Lemma 8 we have

$$\|R + S\|_{*,(\nabla u)}^2 \leq \frac{1}{2} \|2R\|_{*,(\nabla u)}^2 + \frac{1}{2} \|2S\|_{*,(\nabla u)}^2$$

and hence the assertion follows from Lemma 9.  $\square$

The next lemma essentially states that for the steady state version of (1.1), the nonlinear Poisson problem, the dual quasi-norm of the residual is equivalent to the quasi-norm of the error. This is the underlying main principle in [BL93a, DK08, BDK11]; compare with Remark 12.

**Lemma 11.** *Let the  $N$ -function  $\phi$  satisfy (2.1). For  $u, w \in W_0^{1,\phi}(\Omega)$  set*

$$R = -\operatorname{div} \mathbf{A}(\nabla u) \in W^{-1,\phi^*}(\Omega) \quad \text{and} \quad S = -\operatorname{div} \mathbf{A}(\nabla w) \in W^{-1,\phi^*}(\Omega);$$

then,

$$\|\nabla u - \nabla w\|_{(\nabla u)} \approx \|R - S\|_{*,(\nabla u)}.$$

*Proof.* Thanks to (3.3) and (3.4) there exists  $w_{R-S} \in W_0^{1,\phi}(\Omega)$  such that

$$\|R - S\|_{*,(\nabla \nabla u)}^2 = \int_{\Omega} (\phi_{|\nabla u|})^* (\phi'_{|\nabla u|}(|\nabla w_{R-S}|)) dx$$

and

$$(\mathbf{A}(\nabla u) - \mathbf{A}(\nabla w), \nabla v)_0 = (\phi'_{|\nabla u|}(|\nabla w_{R-S}|))_{|\nabla w_{R-S}|}^{\nabla w_{R-S}}, \nabla v)_0,$$

for all  $v \in W_0^{1,\phi}(\Omega)$ . Testing with  $v = w_{R-S}$  and applying Young's inequality (2.11) for  $\phi_{|\nabla u|}$  on the left-hand side together with (2.13) and Proposition 3 proves  $\gtrsim$ . On the other hand testing with  $v = u - w$ , applying Young's inequality on the right-hand side, and Corollary 4 proves  $\lesssim$ .  $\square$



**Remark 12.** We shall briefly discuss the relevance of Lemma 11 in the case of the nonlinear Laplace problem: Find  $u \in W_0^{1,\phi}(\Omega)$ , such that

$$-\operatorname{div} \mathbf{A}(\nabla u) = f \quad \text{in } \Omega \quad \text{and} \quad u = 0 \quad \text{on } \partial\Omega.$$

For an approximation  $V \in W_0^{1,\phi}(\Omega)$  to  $u$  the residual in  $V$  is then given by  $f + \operatorname{div} \mathbf{A}(\nabla V) \in W^{-1,\phi^*}(\Omega)$  and we have by Lemma 11 the equivalence

$$\|f + \operatorname{div} \mathbf{A}(\nabla V)\|_{*,(\nabla u)} \approx \|u - V\|_{(\nabla u)}$$

of error and residual. This is a crucial characteristic of the quasi-norm, which allows for proving optimal estimates as in [BL93a, BL94, LY02, CLY06a, CLY06b, DK08, DER07, BDK11, BBDR10].

**3.3. Equivalence of Error and Residual.** In this section we show that the above introduced error concept is suitable for the nonlinear heat equation. In particular, it provides the building blocks for an error measure on  $\mathbb{W}(I)$ , which yields equivalence of the error and the residual of (2.7). To this end, we fix an  $N$ -function  $\phi$  satisfying (2.1) and (2.4) for some  $p \geq \frac{2d}{d+2}$ ,  $p' > 1$ , with  $\frac{1}{p} + \frac{1}{p'} = 1$ . We recall, that under these assumptions there exists a unique weak solution to (1.1), which we denote by  $u \in \mathbb{W}(I)$ .

The residual  $\operatorname{Res}(v) \in L^{p'}(I; W^{-1,p'}(\Omega))$  of (2.6) in  $v \in \mathbb{W}(I)$  is given by

$$(3.5) \quad \langle \operatorname{Res}(v), w \rangle := \langle f, w \rangle - \langle \partial_t v, w \rangle - \langle \mathbf{A}(\nabla v), \nabla w \rangle_\Omega, \quad w \in W_0^{1,p}(\Omega),$$

a.e. in  $I$ . We define the quasi- $\mathbb{W}$ -norm on  $\mathbb{W}(I)$  by

$$(3.6) \quad \|v\|_{\mathbb{W},(\nabla u)}^2 := \int_I \|\partial_t v\|_{*,(\nabla u)}^2 + \|v\|_{(\nabla u)}^2 dt + \sup_{t \in I} \|v(t)\|_2^2.$$

The following result states that integrating the dual quasi-norm of the residual in time, it is equivalent to the corresponding error in the quasi- $\mathbb{W}$ -norm.

**Theorem 13.** *Supposing the conditions of this section, we have for all  $v \in \mathbb{W}$  that*

$$\|u - v\|_{\mathbb{W},(\nabla u)}^2 \lesssim \int_I \|\operatorname{Res}(v)\|_{*,(\nabla u)}^2 dt + \|u_0 - v(0)\|_2^2$$

and

$$\int_{\tilde{I}} \|\operatorname{Res}(v)\|_{*,(\nabla u)}^2 dt \lesssim \int_{\tilde{I}} \|\partial_t(u - v)\|_{*,(\nabla u)}^2 + \|u - v\|_{(\nabla u)}^2 dt$$

for all  $\tilde{I} \subset I$ . The constants hidden in  $\lesssim$  depend solely on the constants in (2.1).

*Proof.* We prove the second claim first. Recalling the definitions of the residual (3.5) and the dual norm (3.2), we observe from (2.6) that, a.e. in  $I$ , we have

$$\begin{aligned} \|\operatorname{Res}(v)\|_{*,(\nabla u)}^2 &= \sup_{w \in W_0^{1,p}(\Omega)} \langle \operatorname{Res}(v), w \rangle - \|w\|_{(\nabla u)}^2 \\ &= \sup_{w \in W_0^{1,p}(\Omega)} \langle \partial_t(u - v), w \rangle + \langle \mathbf{A}(\nabla u) - \mathbf{A}(\nabla v), \nabla w \rangle_\Omega - \|w\|_{(\nabla u)}^2. \end{aligned}$$

Hence, the claim follows from Corollary 10 and Lemma 11.

In order to prove the first assertion, we observe by Corollary 4 that

$$(3.7) \quad \begin{aligned} \frac{d}{dt} \|u - v\|_2^2 + \|u - v\|_{(\nabla u)}^2 \\ \approx \langle \partial_t(u - v), u - v \rangle + \langle \mathbf{A}(\nabla u) - \mathbf{A}(\nabla v), \nabla u - \nabla v \rangle_\Omega \\ = \langle \operatorname{Res}(v), u - v \rangle \end{aligned}$$

a.e. in  $I$ . Defining  $w_t := w_R$  as in (3.3) with  $R = \text{Res}(v)(t)$ ,  $t \in I$ , we conclude with Young's inequality (2.11) for  $\delta > 0$ , (3.4), and again Corollary 4 that

$$\begin{aligned} \frac{d}{dt} \|u - v\|_2^2 + \|u - v\|_{(\nabla u)}^2 &\approx \left\langle \phi'_{|\nabla u|}(|\nabla w_t|) \frac{\nabla w_t}{|\nabla w_t|}, \nabla u - \nabla v \right\rangle_\Omega \\ &\leq \int_\Omega C_\delta (\phi'_{|\nabla u|})^* (\phi'_{|\nabla u|}(|\nabla w_t|)) + \delta \phi_{|\nabla u|}(|\nabla u - \nabla v|) dx \\ &\lesssim C_\delta \|\text{Res}(v)\|_{*,(\nabla u)}^2 + \delta \|u - v\|_{(\nabla u)}^2, \end{aligned}$$

a.e. in  $I$ . Choosing  $\delta > 0$  small enough and integrating with respect to time yields

$$(3.8) \quad \|u(t) - v(t)\|_2^2 \lesssim \int_0^t \|\text{Res}(v)\|_{*,(\nabla u)}^2 ds + \|u_0 - v(0)\|_2^2$$

for a.e.  $t \in I$  and

$$(3.9) \quad \int_0^t \|u - v\|_{(\nabla u)}^2 dt \lesssim \int_0^t \|\text{Res}(v)\|_{*,(\nabla u)}^2 ds + \|u_0 - v(0)\|_2^2,$$

for all  $t \geq 0$ . On the other hand, by the definition of the residual (3.5), we have

$$\begin{aligned} \|\partial_t(u - v)\|_{*,(\nabla u)}^2 &= \sup_{w \in W_0^{1,p}(\Omega)} \langle \partial_t(u - v), w \rangle - \|w\|_{(\nabla u)}^2 \\ &= \sup_{w \in W_0^{1,p}(\Omega)} \langle \mathbf{A}(\nabla v) - \mathbf{A}(\nabla u), \nabla w \rangle_\Omega + \langle \text{Res}(v), w \rangle - \|w\|_{(\nabla u)}^2 \\ &\lesssim \|u - v\|_{(\nabla u)}^2 + \|\text{Res}(v)\|_{*,(\nabla u)}^2, \end{aligned}$$

a.e. in  $I$ , where the last inequality follows from Corollary 10 and Lemma 11. Integrating this estimate with respect to time and combining it with (3.8) and (3.9) proves the desired assertion.  $\square$

#### 4. A POSTERIORI ERROR ESTIMATES

In this section we shall use the balanced error concept of the previous section to prove reliable and efficient residual based a posteriori estimates for the error of an implicit Euler–Galerkin scheme. As before we fix an N-function  $\phi$  satisfying assumptions (2.1) and (2.4) for some  $p \geq \frac{2d}{d+2}$ ,  $p' > 1$ , with  $\frac{1}{p} + \frac{1}{p'} = 1$  in order to ensure existence and uniqueness of the weak solution  $u \in \mathbb{W}(I)$  to (1.1). We further use the abbreviations  $\mathbb{V} := W_0^{1,p}(\Omega) = W_0^{1,\phi}(\Omega)$  and  $\mathbb{V}^* = W^{-1,p'}(\Omega) = W^{-1,\phi^*}(\Omega)$ .

**4.1. The Backward Euler Galerkin Scheme.** For the discretization in time we let  $0 = t_0 < t_1 < \dots < t_N = T$  be a partition of  $(0, T)$  into half open subintervals  $I_n = (t_{n-1}, t_n]$  with corresponding local time step sizes  $\tau_n := |I_n| = t_n - t_{n-1}$ ,  $n = 1, \dots, N$ .

For the spatial discretization we use Lagrange finite elements of first order on conforming simplicial grids. To be more precise, to each time  $t_n$ ,  $n = 0, \dots, N$ , we assign a partition  $\mathcal{G}_n$  of  $\Omega$  into  $d$ -simplices. The intersection of any two elements of such a partition  $\mathcal{G}_n$ ,  $n \in \{0, \dots, N\}$  is either empty or a complete  $k$ -dimensional sub-simplex,  $0 \leq k < n$ . Moreover, we assume that the family  $\{\mathcal{G}_n\}_{n=0, \dots, N}$  of grids is uniformly shape-regular, i.e., for each element  $E \in \mathcal{G}_n$ ,  $n = 0, \dots, N$ , the ratio of its diameter to the diameter of its largest inscribed ball is bounded by a constant independent of  $E \in \mathcal{G}_n$  and  $n = 0, \dots, N$ . For  $n = 0, \dots, N$  the Lagrange finite element space  $\mathbb{V}_n := \mathbb{V}(\mathcal{G}_n)$  consists of all continuous piecewise affine functions over  $\mathcal{G}_n$  that vanish on the boundary  $\partial\Omega$ .

On any interval  $I_n$ ,  $n = 1, \dots, N$ , we let

$$f_n := \oint_{I_n} f dt = \frac{1}{|I_n|} \int_{I_n} f dt$$

be the local mean value of  $f \in L^{p'}(I; L^{p'}(\Omega))$  on  $I_n$ .

The discrete scheme is defined recursively: Starting from an approximation  $U_0 \in \mathbb{V}_0$  of the initial value  $u_0$ , we define  $U_n \in \mathbb{V}_n$ ,  $n = 1, \dots, N$ , by

$$(4.1a) \quad \left\langle \frac{U_n - U_{n-1}}{\tau_n}, V_n \right\rangle_{\Omega} + \langle \mathbf{A}(\nabla U_n), \nabla V_n \rangle_{\Omega} = \langle f_n, V_n \rangle_{\Omega} \quad \text{for all } V_n \in \mathbb{V}_n.$$

The unique solvability of this finite-dimensional problem follows immediately from Brouwer's fix-point theorem and the strict monotonicity of the nonlinear operator. We denote the linear interpolation at  $(U_n, t_n)$ ,  $n = 0, \dots, N$ , by

$$(4.1b) \quad \mathcal{U}|_{I_n}(t) := \frac{t - t_{n-1}}{\tau_n} U_n + \frac{t_n - t}{\tau_n} U_{n-1}.$$

With this definition  $\mathcal{U}$  is continuous and piece-wise affine in time, which implies that  $\mathcal{U} \in \mathbb{W}(I)$  and hence we deduce from Theorem 13 the following result.

**Corollary 14** (Abstract Error Bound). *Supposing the assumptions and definitions of this section we have that*

$$\|u - \mathcal{U}\|_{\mathbb{W},(\nabla u)}^2 \lesssim \int_I \|\text{Res}(\mathcal{U})\|_{*,(\nabla \mathcal{U})}^2 dt + \|u_0 - U_0\|_2^2$$

and

$$\int_{\tilde{I}} \|\text{Res}(\mathcal{U})\|_{*,(\nabla \mathcal{U})}^2 dt \lesssim \int_{\tilde{I}} \|\partial_t(u - \mathcal{U})\|_{*,(\nabla u)}^2 + \|u - \mathcal{U}\|_{(\nabla u)}^2 dt,$$

for all  $\tilde{I} \subset I$ .

*Proof.* In order to prove the above assertions we observe from Lemma 5 that for  $\delta > 0$  we have

$$\begin{aligned} \|R\|_{*,(\nabla u)}^2 &\leq \sup_{v \in \mathbb{V}} \langle R, v \rangle - \frac{1}{C_\delta} \|v\|_{(\nabla u)}^2 + \frac{\delta}{C_\delta} \|u - \mathcal{U}\|_{(\nabla u)}^2 \\ &= \frac{1}{C_\delta} \|C_\delta R\|_{*,(\nabla u)}^2 + \frac{\delta}{C_\delta} \|u - \mathcal{U}\|_{(\nabla u)}^2 \\ &\leq C(\delta) \|R\|_{*,(\nabla u)}^2 + \frac{\delta}{C_\delta} \|u - \mathcal{U}\|_{(\nabla u)}^2 \end{aligned}$$

for any  $R \in \mathbb{V}^*$ . Analogously,

$$(4.2) \quad \|R\|_{*,(\nabla \mathcal{U})}^2 \leq C(\delta) \|R\|_{*,(\nabla u)}^2 + \frac{\delta}{C_\delta} \|u - \mathcal{U}\|_{(\nabla u)}^2.$$

The constant  $C(\delta) > 0$  results from a repeated application of Lemma 9 in the spirit of (2.14). Together with Theorem 13 we obtain

$$\begin{aligned} \|u - \mathcal{U}\|_{\mathbb{W},(\nabla u)}^2 &= \int_I \|\partial_t(u - \mathcal{U})\|_{*,(\nabla u)}^2 + \|u - \mathcal{U}\|_{(\nabla u)}^2 dt + \sup_{t \in I} \|u - \mathcal{U}\|_2^2 \\ &\lesssim \int_I \|\text{Res}(\mathcal{U})\|_{*,(\nabla \mathcal{U})}^2 dt + \|u_0 - U_0\|_2^2 \\ &\lesssim \int_I C(\delta) \|\text{Res}(\mathcal{U})\|_{*,(\nabla u)}^2 + \frac{\delta}{C_\delta} \|u - \mathcal{U}\|_{(\nabla u)}^2 dt + \|u_0 - U_0\|_2^2. \end{aligned}$$

Recalling that  $C_\delta$  is large for small  $\delta$ , the term  $\frac{\delta}{C_\delta} \int_I \|u - \mathcal{U}\|_{(\nabla u)}^2 dt$  can be absorbed into  $\|u - \mathcal{U}\|_{\mathbb{W},(\nabla u)}^2$  by choosing  $\delta > 0$  small enough. This proves the first assertion and the second one follows analogously.  $\square$

**4.2. Decomposition of the Residual.** We fix  $n \in \{1, \dots, N\}$ . Recalling (3.5), we can decompose the residual according to

$$(4.3) \quad \text{Res}(\mathcal{U}) = \text{Res}_h(\mathcal{U}) + \text{Res}_\tau(\mathcal{U}) + f - f_n$$

with a spatial contribution

$$\text{Res}_h(\mathcal{U}) := f_n - \frac{U_n - U_{n-1}}{\tau_n} + \text{div } \mathbf{A}(\nabla U_n) \quad \text{on } I_n,$$

a temporal contribution

$$\text{Res}_\tau(\mathcal{U}) := \text{div } \mathbf{A}(\nabla \mathcal{U}) - \text{div } \mathbf{A}(\nabla U_n) \quad \text{on } I_n,$$

and a data error. By the definition of the linear interpolant  $\mathcal{U}$ , we have

$$U_n - \mathcal{U}|_{I_n} = \left(1 - \frac{t - t_{n-1}}{\tau_n}\right)(U_n - U_{n-1})$$

on  $I_n$ . Consequently, by Lemmas 11 and 5 it follows that

$$(4.4) \quad \begin{aligned} \|\text{Res}_\tau(\mathcal{U})\|_{*,(\nabla \mathcal{U})}^2 &\approx \|U_n - \mathcal{U}\|_{(\nabla \mathcal{U})}^2 \approx \|U_n - \mathcal{U}\|_{(\nabla U_n)}^2 \\ &= \int_{\Omega} \phi_{|\nabla U_n|} \left( \left(1 - \frac{t - t_{n-1}}{\tau_n}\right) |\nabla U_n - \nabla U_{n-1}| \right) dx, \end{aligned}$$

hence (2.12a) yields

$$\begin{aligned} \int_{I_n} \|\text{Res}_\tau(\mathcal{U})\|_{*,(\nabla \mathcal{U})}^2 dt &\lesssim \int_{I_n} \int_{\Omega} \left(1 - \frac{t - t_{n-1}}{\tau_n}\right) \phi_{|\nabla U_n|} (|\nabla U_n - \nabla U_{n-1}|) dx dt \\ &= \frac{1}{2} \tau_n \|U_n - U_{n-1}\|_{(\nabla U_n)}^2. \end{aligned}$$

On the other hand we get from (4.4) with Jensen's inequality that

$$\begin{aligned} \int_{I_n} \|\text{Res}_\tau(\mathcal{U})\|_{*,(\nabla \mathcal{U})}^2 dt &\gtrsim \tau_n \int_{\Omega} \phi_{|\nabla U_n|} \left( \frac{1}{\tau_n} \int_{I_n} \left(1 - \frac{t - t_{n-1}}{\tau_n}\right) |\nabla U_n - \nabla U_{n-1}| dt \right) dx \\ &= \tau_n \int_{\Omega} \phi_{|\nabla U_n|} \left( \frac{1}{2} |\nabla U_n - \nabla U_{n-1}| \right) dx \\ &\gtrsim \tau_n \|U_n - U_{n-1}\|_{(\nabla U_n)}^2, \end{aligned}$$

where we used that  $\Delta_2(\{\phi_a\}_{a \geq 0})$  is bounded; see Lemma 2. We have thus proved the following result.

**Lemma 15** (Temporal Indicator). *Suppose that the conditions of this section are satisfied. Then we have for the temporal contribution to the residual in (4.3) that*

$$(4.5) \quad \int_{I_n} \|\text{Res}_\tau(\mathcal{U})\|_{*,(\nabla \mathcal{U})}^2 dt \approx \tau_n \|U_n - U_{n-1}\|_{(\nabla U_n)}^2 =: \tau_n \Theta_n^2,$$

where the hidden constants depend solely on  $\Delta_2(\{\phi, \phi^*\})$  and thus on the constants in (2.1).

The decomposition (4.3) is proper in the sense that the sum of the single contributions is equivalent to the residual up to the term  $f - f_n$  on  $I_n$ ,  $n = 1, \dots, N$ , which measures the resolution of the right-hand side.

**Lemma 16.** *For the decomposition (4.3) of the residual we have the upper bound*

$$\begin{aligned} \int_I \|\text{Res}(\mathcal{U})\|_{*,(\nabla \mathcal{U})}^2 dt &\lesssim \sum_{n=1}^N \int_{I_n} \|\text{Res}_h(\mathcal{U})\|_{*,(\nabla U_n)}^2 + \|\text{Res}_\tau(\mathcal{U})\|_{*,(\nabla U_n)}^2 + \|f - f_n\|_{*,(\nabla U_n)}^2 dt \end{aligned}$$

as well as the lower bound

$$\begin{aligned} \int_{I_n} \|\text{Res}_h(\mathcal{U})\|_{*,(\nabla U_n)}^2 + \|\text{Res}_\tau(\mathcal{U})\|_{*,(\nabla U_n)}^2 dt \\ \lesssim \int_{I_n} \|\text{Res}(\mathcal{U})\|_{*,(\nabla \mathcal{U})}^2 + \|f - f_n\|_{*,(\nabla U_n)}^2 dt \end{aligned}$$

for all  $n = 1, \dots, N$ .

*Proof.* Before we turn to proving the asserted estimates we first investigate the change of the dual quasi-norm from  $\|\cdot\|_{*,(\nabla \mathcal{U})}$  to  $\|\cdot\|_{*,(\nabla U_n)}$  and vice versa. To this end we fix  $n \in \{1, \dots, N\}$  and observe, as in (4.2), that for  $\delta > 0$  there exist constants  $C_\delta, C(\delta) > 0$  such that

$$(4.6) \quad \|\text{Res}(\mathcal{U})\|_{*,(\nabla \mathcal{U})}^2 \leq C(\delta) \|\text{Res}(\mathcal{U})\|_{*,(\nabla U_n)}^2 + \frac{\delta}{C_\delta} \|U_n - \mathcal{U}\|_{*,(\nabla U_n)}^2$$

and analogously

$$(4.7) \quad \|\text{Res}(\mathcal{U})\|_{*,(\nabla U_n)}^2 \leq C(\delta) \|\text{Res}(\mathcal{U})\|_{*,(\nabla \mathcal{U})}^2 + \frac{\delta}{C_\delta} \|U_n - \mathcal{U}\|_{*,(\nabla \mathcal{U})}^2.$$

Now, thanks to the first estimate, the upper bound follows from (4.4), the decomposition (4.3), and the generalized triangle inequality in Corollary 10.

It remains to prove the lower bound. For the temporal residual we obtain by Corollary 10 and (4.3) that

$$(4.8) \quad \|\text{Res}_\tau(\mathcal{U})\|_{*,(\nabla U_n)}^2 \lesssim \|\text{Res}(\mathcal{U})\|_{*,(\nabla U_n)}^2 + \|f - f_n\|_{*,(\nabla U_n)}^2 + \|\text{Res}_h(\mathcal{U})\|_{*,(\nabla U_n)}^2,$$

almost everywhere in  $I_n$ . For the spatial residual we use techniques from [Ver03]. To this end we observe that  $\|\text{Res}_h(\mathcal{U})\|_{*,(\nabla U_n)}^2$  is constant on  $I_n$  and therefore

$$\int_{I_n} \|\text{Res}_h(\mathcal{U})\|_{*,(\nabla U_n)}^2 dt = \int_{I_n} a(t) \|\text{Res}_h(\mathcal{U})\|_{*,(\nabla U_n)}^2 dt,$$

where

$$a(t) := (\alpha + 1) \left( \frac{t - t_{n-1}}{\tau_n} \right)$$

for some  $\alpha \geq 0$  to be determined later. Note that  $a \geq 0$  on  $I_n$ . Thanks to (4.3) and Corollary 10 this implies

$$\begin{aligned} \int_{I_n} \|\text{Res}_h(\mathcal{U})\|_{*,(\nabla U_n)}^2 dt &\lesssim \int_{I_n} a(t) \|\text{Res}(\mathcal{U})\|_{*,(\nabla U_n)}^2 + a(t) \|f - f_n\|_{*,(\nabla U_n)}^2 \\ &\quad + a(t) \|\text{Res}_\tau(\mathcal{U})\|_{*,(\nabla U_n)}^2 dt \\ (4.9) \quad &\leq (\alpha + 1) \int_{I_n} \|\text{Res}(\mathcal{U})\|_{*,(\nabla U_n)}^2 + \|f - f_n\|_{*,(\nabla U_n)}^2 dt \\ &\quad + \int_{I_n} a(t) \|\text{Res}_\tau(\mathcal{U})\|_{*,(\nabla U_n)}^2 dt. \end{aligned}$$

According to (4.4) the latter term can be estimated by

$$\begin{aligned} \int_{I_n} a(t) \|\text{Res}_\tau(\mathcal{U})\|_{*,(\nabla U_n)}^2 dt \\ \approx \int_{I_n} a(t) \int_{\Omega} \phi_{|\nabla U_n|} \left( \left( 1 - \frac{t - t_{n-1}}{\tau_n} \right) |\nabla(U_n - U_{n-1})| \right) dx dt \\ \leq \int_{I_n} \int_{\Omega} a(t) \left( 1 - \frac{t - t_{n-1}}{\tau_n} \right) \phi_{|\nabla U_n|} (|\nabla(U_n - U_{n-1})|) dx dt, \end{aligned}$$

where we used (2.12a) in the last step. By Fubini's Theorem and Lemma 15 this yields

$$\int_{I_n} a(t) \|\text{Res}_\tau(\mathcal{U})\|_{*,(\nabla U_n)}^2 dt \lesssim \left(1 - \frac{\alpha+1}{\alpha+2}\right) \int_{I_n} \|\text{Res}_\tau(\mathcal{U})\|_{*,(\nabla U_n)}^2 dt.$$

Therefore, combining this with (4.8) and (4.9), we get

$$\begin{aligned} \int_{I_n} \|\text{Res}_\tau(\mathcal{U})\|_{*,(\nabla U_n)}^2 + \|\text{Res}_h(\mathcal{U})\|_{*,(\nabla U_n)}^2 dt \\ \lesssim (\alpha+1) \int_{I_n} \|\text{Res}(\mathcal{U})\|_{*,(\nabla U_n)}^2 + \|f - f_n\|_{*,(\nabla U_n)}^2 dt \\ + \left(1 - \frac{\alpha+1}{\alpha+2}\right) \int_{I_n} \|\text{Res}_\tau(\mathcal{U})\|_{*,(\nabla U_n)}^2 dt, \end{aligned}$$

and hence with (4.7) and (4.4), for  $\delta > 0$ , we arrive at

$$\begin{aligned} \lesssim (\alpha+1) \int_{I_n} C(\delta) \|\text{Res}(\mathcal{U})\|_{*,(\nabla \mathcal{U})}^2 + \|f - f_n\|_{*,(\nabla U_n)}^2 dt \\ + \left(1 - \frac{\alpha+1}{\alpha+2} + (\alpha+1) \frac{\delta}{C_\delta}\right) \int_{I_n} \|\text{Res}_\tau(\mathcal{U})\|_{*,(\nabla U_n)}^2 dt. \end{aligned}$$

Choosing  $\alpha \geq 0$  large enough and  $\delta > 0$  small enough the latter term can be absorbed into the left-hand side. This finishes the proof.  $\square$

**4.3. The Spatial Residual  $\text{Res}_h$ .** In the light of the previous section, namely Lemmas 16 and 15, we need to investigate the spatial residual. In particular, fixing  $n \in \{1, \dots, N\}$ , in this section we shall prove that the spatial residual on the interval  $I_n$  can be estimated by

(4.10)

$$\mathcal{E}_n^2 := \sum_{E \in \mathcal{G}_n} \mathcal{E}_n^2(E) := \sum_{E \in \mathcal{G}_n} \int_E (\phi_{|\nabla U_n|})^* (h_E |R_n|) dx + h_E \|\llbracket \mathbf{F}(\nabla U_n) \rrbracket\|_{2,\delta E}^2.$$

Here  $R_n := f_n - \frac{U_n - U_{n-1}}{\tau_n}$  and  $\llbracket \mathbf{F}(\nabla U_n) \rrbracket|_S$  denotes the jump of  $\mathbf{F}(\nabla U_n)$  across inter-element sides  $S \in \mathcal{S}_n$  of  $\mathcal{G}_n$ , respectively  $\llbracket \mathbf{F}(\nabla U_n) \rrbracket = 0$  on  $\partial\Omega$ . The local mesh-sizes are given by  $h_E := |E|^{1/d}$ ,  $E \in \mathcal{G}_n$ . In particular, we have the following bounds.

**Lemma 17** (Spatial Indicator). *Suppose that the assumptions of this section are satisfied. Then we have, with the notation introduced above, that*

$$\|\text{Res}_h(\mathcal{U})\|_{*,(\nabla U_n)}^2 \lesssim \mathcal{E}_n^2 \lesssim \|\text{Res}_h(\mathcal{U})\|_{*,(\nabla U_n)}^2 + \text{osc}_n^2,$$

where

$$(4.11) \quad \text{osc}_n^2 = \sum_{E \in \mathcal{G}_n} \text{osc}_n^2(E) := \sum_{E \in \mathcal{G}_n} \min_{R_E \in \mathbb{R}} \int_E (\phi_{|\nabla U_n|})^* (h_E |R_n - R_E|) dx.$$

The constants hidden in  $\lesssim$  depend solely on  $\Delta_2(\{\phi, \phi^*\})$  and on the shape-regularity of  $\mathcal{G}_n$ .

*Proof.* The proof is similar to the proof of the error bounds in [DK08]. For this reason we only sketch it and point out the differences.

We start by showing the first inequality. For arbitrary  $v \in \mathbb{V}$  we denote the Scott Zhang quasi-interpolation into the finite element space  $\mathbb{V}_n$  by  $\Pi_n : \mathbb{V} \rightarrow \mathbb{V}_n$ ; see [SZ90]. From (4.1a) we observe that the spatial residual is orthogonal to  $\mathbb{V}_n$ . Consequently, we have

$$\begin{aligned} \langle \text{Res}_h(U), v \rangle &= \langle \text{Res}_h(U), v - \Pi_n v \rangle \\ &= \langle R_n, v - \Pi_n v \rangle_\Omega - \langle \mathbf{A}(\nabla U_n), \nabla(v - \Pi_n v) \rangle_\Omega. \end{aligned}$$

Note that  $\mathbf{A}(\nabla U_n)$  is piece wise constant; hence, using integration by parts, we obtain

$$\begin{aligned} \langle \text{Res}_h(U), v \rangle &= \langle R_n, v - \Pi_n v \rangle_\Omega - \frac{1}{2} \sum_{E \in \mathcal{G}_n} \int_{\partial E} \llbracket \mathbf{A}(\nabla U_n) \cdot \mathbf{n} \rrbracket (v - \Pi_n v) d\sigma \\ &= \text{Upper}_1 + \text{Upper}_2, \end{aligned}$$

where  $\mathbf{n}$  denotes the unit outer normal on  $\partial E$ . We estimate the two terms separately. For the first term we conclude with Young's inequality (2.11) and interpolation estimates for N-functions [DR07a] for arbitrary  $\delta > 0$ , that

$$\begin{aligned} \text{Upper}_1 &\leq \sum_{E \in \mathcal{G}} \int_E C_\delta (\phi_{|\nabla U_n|})^* (h_E |R_n|) + \delta \phi_{|\nabla U_n|} \left( \left| \frac{v - \Pi_n v}{h_E} \right| \right) dx \\ &\lesssim \sum_{E \in \mathcal{G}} C_\delta \int_E (\phi_{|\nabla U_n|})^* (h_E |R_n|) dx + \delta \int_{\Omega_E} \phi_{|\nabla U_n|} (|\nabla v|) dx, \end{aligned}$$

where  $\Omega_E := \bigcup \{E' \in \mathcal{G}_n \mid E' \cap E \neq \emptyset\}$ . Note that the shift of the last term is fixed to  $|\nabla U_n|_E$  although the domain of integration  $\Omega_E$  covers the neighboring elements of  $E$  as well. For this reason we apply Lemma 5 in order to obtain

$$\begin{aligned} \text{Upper}_1 &\lesssim \sum_{E \in \mathcal{G}} \left\{ C_\delta \int_E (\phi_{|\nabla U_n|})^* (h_E |R_n|) dx + \delta \int_{\Omega_E} \phi_{|\nabla U_n|} (|\nabla v|) dx \right. \\ &\quad \left. + \delta \sum_{E' \subset \Omega_E} \int_{E'} \phi_{|\nabla U_n|} (|\nabla U_n|_E - \nabla U_n|_{E'}|) dx \right\}. \end{aligned}$$

Proposition 3 then implies

$$\phi_{|\nabla U_n|} (|\nabla U_n|_E - \nabla U_n|_{E'}|) \approx |\mathbf{F}(\nabla U_n|_E) - \mathbf{F}(\nabla U_n|_{E'})|^2 \leq \sum_{S \in \Omega_E} |\llbracket \mathbf{F}(\nabla U_n) \rrbracket_S|^2.$$

Here we used the triangle inequality and that we can find a path from  $E$  to  $E'$  only crossing sides  $S \in \mathcal{S}$ ,  $S \in \Omega_E$ . Thanks to the shape-regularity of  $\mathcal{G}_n$ , the numbers of elements and inter-element sides contained in  $\Omega_E$  are bounded uniformly in  $E \in \mathcal{G}_n$ . Moreover, for each  $E \in \mathcal{G}_n$  we have  $h_E |S| \approx |E'|$  for all  $E' \in \mathcal{G}_n$  and all  $S \in \mathcal{S}$  with  $E' \subset \Omega_E$ ,  $S \subset E$ . Therefore, by the finite overlap of the  $\Omega_E$  and using that  $\llbracket \mathbf{F}(\nabla U_n) \rrbracket$  is constant on each interior side, we arrive at

$$\begin{aligned} \text{Upper}_1 &\lesssim \sum_{E \in \mathcal{G}} \left\{ C_\delta \int_E (\phi_{|\nabla U_n|})^* (h_E |R_n|) dx + \delta \int_{\partial E} h_E |\llbracket \mathbf{F}(\nabla U_n) \rrbracket_S|^2 d\sigma \right\} \\ (4.12) \quad &+ \delta \int_{\Omega} \phi_{|\nabla U_n|} (|\nabla v|) dx; \end{aligned}$$

for details see [DK08, Kre08]. We turn to the estimate of  $\text{Upper}_2$  observing, by a standard scaled trace theorem and Young's inequality (2.11), that

$$\begin{aligned} \text{Upper}_2 &\leq \frac{1}{2} \sum_{E \in \mathcal{G}_n} \sum_{S \subset \partial E} |\llbracket \mathbf{A}(\nabla U_n) \rrbracket_S| \int_S |v - \Pi_n v| d\sigma \\ &\lesssim \sum_{E \in \mathcal{G}_n} \sum_{S \subset \partial E} |\llbracket \mathbf{A}(\nabla U_n) \rrbracket_S| \int_{\Omega_E} |\nabla v| dx \\ &\leq \sum_{E \in \mathcal{G}_n} \sum_{S \subset \partial E} \int_{\Omega_E} C_\delta (\phi_{|\nabla U_n|})^* (|\llbracket \mathbf{A}(\nabla U_n) \rrbracket_S|) + \delta \phi_{|\nabla U_n|} (|\nabla v|) dx, \end{aligned}$$

for any  $\delta > 0$ . To estimate the first term, we recall from Proposition 3 and (2.13) that on  $\partial E$  we have

$$(4.13) \quad \begin{aligned} \left( \phi_{|\nabla U_n|E} \right)^* (|\llbracket \mathbf{A}(\nabla U_n) \rrbracket|) &\approx \left( \phi_{|\nabla U_n|E} \right)^* \left( \phi'_{|\nabla U_n|E} (|\llbracket \nabla U_n \rrbracket|) \right) \\ &\approx \phi_{|\nabla U_n|E} (|\llbracket \nabla U_n \rrbracket|) \approx |\llbracket \mathbf{F}(\nabla U) \rrbracket|^2. \end{aligned}$$

For the second term, we are in the same situation as in the estimate of  $\text{Upper}_1$ , that the shift on whole  $\Omega_E$  is fixed to  $|\nabla U_n|E$ . Consequently, using the same techniques as for  $\text{Upper}_1$ , we arrive at

$$\text{Upper}_2 \lesssim \sum_{E \in \mathcal{G}} \left\{ C_\delta \int_{\partial E} h_E |\llbracket \mathbf{F}(\nabla U_n) \rrbracket|^2 d\sigma \right\} + \delta \int_{\Omega} \phi_{|\nabla U_n|} (|\nabla v|) dx.$$

In combination with (4.12) and the definition of the error indicators in (4.10), this yields

$$\langle \text{Res}_h(\mathcal{U}), v \rangle \lesssim C_\delta \mathcal{E}_n^2 + \delta \|v\|_{(\nabla U_n)}^2.$$

Consequently, for suitable chosen  $\delta$ , we arrive at

$$\langle \text{Res}_h(U), v \rangle - \|v\|_{(\nabla U_n)}^2 \lesssim \mathcal{E}_n^2.$$

In view of the definition of the dual quasi-norm (3.2), taking the supremum over all  $v \in \mathbb{V}$  proves the first estimate.

In order to prove the second inequality we recall the set of interior sides  $\mathcal{S}_n$  of  $\mathcal{G}_n$ , i.e., for each  $S \in \mathcal{S}_n$ , there exists  $E_1, E_2 \in \mathcal{G}_n$  such that  $S = E_1 \cap E_2$ . We set  $\omega_S := E_1 \cup E_2$ . For  $E \in \mathcal{G}_n$  and  $S \in \mathcal{S}_n$  we define bubble functions  $\psi_E \in W_0^{1,p}(E)$  and  $\psi_S \in W_0^{1,p}(\omega_S)$ , as the product of the corresponding barycentric coordinates scaled by  $\frac{1}{d!}$  and  $\frac{1}{(d-1)!}$  respectively. Then there exists a constant  $C > 0$  only depending on the shape-regularity of  $\mathcal{G}_n$  such that

$$(4.14a) \quad \int_E \psi_E dx = |E|, \quad \|\psi_E\|_\infty \leq C, \quad \|\nabla \psi_E\|_\infty \leq \frac{C}{h_E},$$

and

$$(4.14b) \quad \int_S \psi_S d\sigma = |S|, \quad \|\psi_S\|_\infty \leq C, \quad \|\nabla \psi_S\|_\infty \leq \frac{C}{h_S},$$

where  $h_S := |S|^{\frac{1}{d-1}}$ ; compare e.g. with [Kre08]. We investigate the single contributions to  $\mathcal{E}_n$  separately. Fix  $E \in \mathcal{G}_n$ , and recall that  $|\nabla U_n|E| \in \mathbb{R}$  is constant. Then for each  $R_E \in \mathbb{R}$  taking  $s_E := \text{sgn}(R_E) (\phi_{|\nabla U_n|E})^{*'} (|h_E R_E|) \in \mathbb{R}$  we have that

$$h_E R_E s_E = \left( \phi_{|\nabla U_n|E} \right)^* (|h_E R_E|) + \phi_{|\nabla U_n|E} (s_E);$$

compare with (2.10). Consequently, by (4.14a), (2.11), and (2.14) we obtain for  $\delta > 0$  that

$$\begin{aligned} \int_E (\phi_{|\nabla U_n|})^* (|h_E R_E|) + \phi_{|\nabla U_n|} (|s_E|) dx &= \int_E h_E R_E s_E dx = \int_E h_E R_E s_E \psi_E dx \\ &= \int_E h_E R_n s_E \psi_E dx + \int_E h_E (R_E - R_n) s_E \psi_E dx \\ &\leq \langle \text{Res}_h(\mathcal{U}), h_E s_E \psi_E \rangle_\Omega \\ &\quad + \int_E C_\delta (\phi_{|\nabla U_n|})^* (|h_E (R_E - R_n)|) + \delta \phi_{|\nabla U_n|} (|s_E \psi_E|) dx \\ &\lesssim \langle \text{Res}_h(\mathcal{U}), h_E s_E \psi_E \rangle_\Omega \\ &\quad + \int_E C_\delta (\phi_{|\nabla U_n|})^* (|h_E (R_E - R_n)|) + \delta \phi_{|\nabla U_n|} (|s_E|) dx. \end{aligned}$$



The triangle inequality (2.15) and choosing  $\delta > 0$  small enough then yields

$$(4.15) \quad \int_E (\phi_{|\nabla U_n|})^* (|h_E R_n|) dx \lesssim \langle \text{Res}_h(\mathcal{U}), h_E s_E \psi_E \rangle_\Omega + \int_E (\phi_{|\nabla U_n|})^* (|h_E (R_E - R_n)|) dx.$$

For the side contribution to  $\mathcal{E}_n$  fix  $S \in \mathcal{S}_n$ ,  $S \subset E$ , and recall (4.13). Like above set

$$s_S^E := \text{sgn} \left( \llbracket \mathbf{A}(\nabla U_n) \mathbf{n} \rrbracket|_S \right) (\phi_{|\nabla U_n|})^* \left( \left| \llbracket \mathbf{A}(\nabla U_n) \mathbf{n} \rrbracket|_S \right| \right) \in \mathbb{R},$$

where  $\mathbf{n}$  is the unit outer normal with respect to  $S \subset E$ . Consequently, by (2.10), (2.11), and (4.14b) we have for  $\delta > 0$  that

$$\begin{aligned} & \int_S h_E (\phi_{|\nabla U_n|})^* (|\llbracket \mathbf{A}(\nabla U_n) \mathbf{n} \rrbracket|) + h_E (\phi_{|\nabla U_n|})^* (|s_S^E|) d\sigma \\ &= \int_S \llbracket \mathbf{A}(\nabla U_n) \mathbf{n} \rrbracket h_E s_S^E d\sigma = \int_S \llbracket \mathbf{A}(\nabla U_n) \mathbf{n} \rrbracket s_S^E h_E \psi_S d\sigma \\ &= \langle \text{Res}_h(\mathcal{U}), s_S^E h_E \psi_S \rangle_\Omega - \langle h_E R_n, s_S^E \psi_S \rangle_\Omega \\ &\leq \langle \text{Res}_h(\mathcal{U}), s_S^E h_E \psi_S \rangle_\Omega \\ &\quad + \int_{\omega_S} C_\delta (\phi_{|\nabla U_n|})^* (h_E |R_n|) + \delta (\phi_{|\nabla U_n|})^* (|s_S^E \psi_S|) dx. \end{aligned}$$

Shape-regularity of  $\mathcal{G}_n$  and (4.14b) imply that

$$\int_{\omega_S} (\phi_{|\nabla U_n|})^* (|s_S^E \psi_S|) dx \lesssim \int_S h_E (\phi_{|\nabla U_n|})^* (|s_S^E|) d\sigma.$$

Note from [DK08] that  $|\llbracket \mathbf{A}(\nabla U_n) \mathbf{n} \rrbracket| \approx |\llbracket \mathbf{A}(\nabla U_n) \rrbracket|$ . Therefore, choosing  $\delta > 0$  small enough yields

$$\begin{aligned} & \int_S h_E (\phi_{|\nabla U_n|})^* (|\llbracket \mathbf{A}(\nabla U_n) \rrbracket|) d\sigma \\ &\lesssim \langle \text{Res}_h(\mathcal{U}), s_S^E h_E \psi_S \rangle_\Omega + \int_{\omega_S} (\phi_{|\nabla U_n|})^* (h_E |R_n|) dx. \end{aligned}$$

We recall that  $\omega_S = E \cup \tilde{E}$  for some  $\tilde{E} \in \mathcal{G}_n$  and note that the last term in the estimate has the wrong shift on  $\tilde{E}$ . The shift can be changed to  $|\nabla U_n|_{\tilde{E}}$  using Lemma 5 and observing by (4.13) that the resulting correction term can be absorbed into the left-hand side of the estimate. Define  $\omega_E := \bigcup_{S \subset \partial E} \omega_S$  and note that, depending on the shape-regularity of  $\mathcal{G}_n$ , we have that  $h_E \approx h_{\tilde{E}}$  for all  $\tilde{E} \in \mathcal{G}_n$ ,  $\tilde{E} \subset \omega_E$ . Together with (4.15), (4.13), and the finite overlap of the patches, we thus get

$$(4.16) \quad \begin{aligned} & \int_E (\phi_{|\nabla U_n|})^* (|h_E R_n|) dx + \int_{\partial E} h_E |\llbracket \mathbf{F}(\nabla U_n) \rrbracket|^2 d\sigma \\ &\lesssim \left\langle \text{Res}_h(\mathcal{U}), \sum_{S \in \mathcal{S}_n, S \subset E} s_S^E h_E \psi_S + \sum_{\tilde{E} \in \mathcal{G}_n, \tilde{E} \subset \omega_E} s_{\tilde{E}} h_{\tilde{E}} \psi_{\tilde{E}} \right\rangle_\Omega \\ &\quad + \sum_{\tilde{E} \in \mathcal{G}_n, \tilde{E} \subset \omega_E} \int_{\tilde{E}} (\phi_{|\nabla U_n|})^* (h_{\tilde{E}} |R_n - R_{\tilde{E}}|) dx. \end{aligned}$$

Recall that  $R_E \in \mathbb{R}$ ,  $E \in \mathcal{G}_n$  was chosen arbitrarily. Whence summing over all  $E \in \mathcal{G}_n$  and again accounting for the finite overlap of the  $\omega_{\tilde{E}}$ ,  $\tilde{E} \in \mathcal{G}_n$ , we arrive at

$$\mathcal{E}_n^2 \lesssim \langle \text{Res}_h(\mathcal{U}), v \rangle_\Omega + \text{osc}_n^2,$$

where

$$v = \sum_{E \in \mathcal{G}_n} \left\{ s_E h_E \psi_E + \sum_{S \in \mathcal{S}_n, S \subset E} s_S^E h_E \psi_S \right\}.$$

Since  $\psi_S, \psi_E \in \mathbb{V}$  for all  $E \in \mathcal{G}_n$  and  $S \in \mathcal{S}$  we have  $v \in \mathbb{V}$ . Moreover, we can bound

$$\begin{aligned} \|v\|_{(\nabla U_n)}^2 &= \int_{\Omega} \phi_{|\nabla U_n|} (|\nabla v|) \, dx \\ &\lesssim \sum_{E \in \mathcal{G}_n} \left\{ \int_E \phi_{|\nabla U_n|} (|s_E h_E \nabla \psi_E|) \, dx \right. \\ &\quad \left. + \sum_{S \in \mathcal{S}_n, S \subset E} \int_{\omega_S} \phi_{|\nabla U_n|} (|s_S^E h_E \nabla \psi_S|) \, dx \right\} \\ &\lesssim \sum_{E \in \mathcal{G}_n} \left\{ \int_E \phi_{|\nabla U_n|} (|s_E|) \, dx + \sum_{S \in \mathcal{S}_n, S \subset E} \int_{\omega_S} \phi_{|\nabla U_n|} (|s_S^E|) \, dx \right\}, \end{aligned}$$

where we used once again the finite overlap of patches, the quasi triangle inequality (2.15), and (4.14). For  $S \subset \partial E$  changing the shift of the last addends to  $\nabla U_n|_E$  on the whole domain  $\omega_S$ , we obtain with Lemma 5, the definitions of  $s_E, s_S^E$ , (2.13), and (4.13), that

$$\|v\|_{(\nabla U_n)}^2 \lesssim \mathcal{E}_n^2.$$

In other words, there exists a constant  $C > 0$  such that

$$\mathcal{E}_n^2 \leq \langle C \operatorname{Res}_h(\mathcal{U}), v \rangle_{\Omega} - \|v\|_{(\nabla U_n)} + C \operatorname{osc}_n^2 \leq \|C \operatorname{Res}_h(\mathcal{U})\|_{*, (\nabla u)}^2 + C \operatorname{osc}_n^2.$$

The assertion follows now from Lemma 9, arguing as in (2.14).  $\square$

Actually, in the proof of Lemma 17 we proved even a local lower bound for the spatial residual. To see this, for each element  $E \in \mathcal{G}_n$ , we define the patch of adjacent elements by

$$\omega_E := \bigcup \{T' \in \mathcal{G}_n : T' = T \text{ or } T' \cap T \text{ is a common side}\} \subset \Omega.$$

Note that this definition complies with the definition in the proof of Lemma 17.

**Corollary 18.** *Suppose that the conditions of Lemma 17 are satisfied. Then, for each  $E \in \mathcal{G}_n$  there exists  $v_E \in W_0^{1,p}(\omega_E)$  such that*

$$\mathcal{E}_n^2(E) \lesssim \langle \operatorname{Res}_h(\mathcal{U}), v_E \rangle + \sum_{\tilde{E} \in \mathcal{G}_n, \tilde{E} \subset \omega_E} \operatorname{osc}_n^2(E)$$

and

$$\|v_E\|_{(\nabla U_n)}^2 \lesssim \sum_{\tilde{E} \in \mathcal{G}_n, \tilde{E} \subset \omega_E} \mathcal{E}_n^2(E).$$

*Proof.* With the same notation as in the proof of Lemma 17, we choose

$$v_E := \sum_{S \in \mathcal{S}_n, S \subset E} s_S^E h_E \psi_S + \sum_{\tilde{E} \in \mathcal{G}_n, \tilde{E} \subset \omega_E} s_{\tilde{E}} h_{\tilde{E}} \psi_{\tilde{E}}.$$

This function can be estimated in the same way as the test function  $v$  in the proof of Lemma 17. Therefore, the assertion is an immediate consequence of (4.16).  $\square$

**4.4. Reliable and efficient a posteriori estimates.** We now assemble the observations previously made in order to prove a reliable and efficient a posteriori estimate for the error of the approximations (4.1) to (2.6).

**Theorem 19.** *Assume the conditions and notations of this section. We can estimate the error between the solution  $u \in \mathbb{W}(I)$  of (2.6) and its approximation  $\mathcal{U} \in \mathbb{W}(I)$  from (4.1) by*

$$\|u - \mathcal{U}\|_{\mathbb{W},(\nabla u)}^2 \lesssim \|u_0 - U_0\|_2^2 + \sum_{n=1}^N \tau_n (\mathcal{E}_n^2 + \Theta_n^2) + \int_{I_n} \|f - f_n\|_{*,(\nabla U_n)}^2 dt$$

and

$$\tau_n (\mathcal{E}_n^2 + \Theta_n^2) \lesssim \int_{I_n} \|\partial_t(u - \mathcal{U})\|_{*,(\nabla u)}^2 + \|u - \mathcal{U}\|_{(\nabla u)}^2 + \|f - f_n\|_{*,(\nabla U_n)}^2 dt + \tau_n \text{osc}_n^2.$$

The constants hidden in  $\lesssim$  depend solely on the shape-regularity of the family of grids  $\{\mathcal{G}_n\}_{n=0,\dots,N}$  and the constants in (2.1). The quantities  $\mathcal{E}_n$ ,  $\Theta_n$ , and  $\text{osc}_n$ ,  $n = 1, \dots, N$ , are defined in (4.10), (4.5), and (4.11) respectively.

*Proof.* The estimates follow from Corollary 14 and Lemmas 16, 15, and 17.  $\square$

Several remarks are now in order.

1. Our result does not rely on the 'transition condition' (compare with [Ver03, Ver05]), which basically restricts the coarseness of a grid relative to the grid of the previous time-step. This condition is crucial to handle simultaneously the functions  $U_n$  and  $U_{n-1}$ , which are defined on different meshes. When using A-stable  $\theta$  schemes as in [Ver03, Ver05] this issue appears in the diffusion and therefore the condition enters in both the proof of the reliability as well as the efficiency estimate. For the implicit Euler scheme the transition condition can be avoided at the price of the oscillation  $\text{osc}_n$  in Lemma 17, which comprises a control of the resolution of  $U_{n-1}$  on  $\mathcal{G}_n$ ; compare with [KMSS11]. This term however can be a controlled a posteriori and only affects the efficiency of the estimate in Theorem 19.
2. The term  $\int_{I_n} \|f - f_n\|_{*,(\nabla U_n)}^2 dt$  is a data error measuring the resolution of the right-hand side with respect to the temporal variable. We can further estimate this term assuming a more regular right-hand side  $f$ . Observe that

$$|f - f_n| \leq \int_{I_n} |\partial_t f| dt$$

almost everywhere in  $\Omega$  provided  $f \in C^1(I_n; L^{q'}(\Omega))$ . The same estimate holds for  $f \in W^{1,q'}(I; L^{q'}(\Omega))$  by a density argument. Consequently, by the convexity of the dual quasi-norm Lemma 8 and Jensen's inequality, we get

$$\begin{aligned} \int_{I_n} \|f - f_n\|_{*,(\nabla U_n)}^2 dt &\leq \int_{I_n} \left\| \int_{I_n} |\partial_t f| ds \right\|_{*,(\nabla U_n)}^2 dt \\ &\leq \int_{I_n} \|\tau_n \partial_t f\|_{*,(\nabla U_n)}^2 dt. \end{aligned}$$

Note that for this estimate it is crucial that  $\nabla U_n$  and thus also the dual quasi-norm do not depend on time.

3. The error concept in [CLY06b] does not lead to equivalence of error and residual since it does not control the temporal derivative. Consequently the error and the residual are not equivalent, which is demonstrated in their non optimal efficiency estimate [CLY06b, Theorem 5.2].
4. As in [Ver03] the terms  $\int_{I_n} \|f - f_n\|_{*,(\nabla U_n)}^2 dt$  or  $\int_{I_n} \|\tau_n \partial_t f\|_{*,(\nabla U_n)}^2 dt$  are not computable, but unlike the linear case it is not clear how to estimate them

by computable quantities; compare with [KMSS11, CF06]. Although, by the definition of the dual quasi-norm, we can estimate

$$\|f - f_n\|_{*,(\nabla U_n)}^2 \leq \sup_{w \in \mathbb{V}} \left\{ \int_{\Omega} C_{\delta}(\phi_{|\nabla U_n|})^* (|f - f_n|) dt + \delta \phi_{|\nabla U_n|}(|w|) - \phi_{|\nabla U_n|}(|\nabla w|) dx \right\},$$

a.e. in  $I_n$ , the shift  $\nabla U_n$  is not constant in space. To the best of our knowledge, there is no Friedrichs inequality known in this case (for constant shift see e.g. [Kre08] for a Friedrichs inequality). Hence in general the last two terms will not cancel. In Section §5.1 below, we prove a fully computable bound for a modified error quantity similar to the one in [CLY06b].

For some particular nonlinearities other estimates may apply, compare with Remark 21.

## 5. APPLICATIONS

This final section focuses on the practical applicability of the above theory. First we shall prove a fully computable error bound for a modified error quantity. Finally we illustrate by example concrete estimators for the power law (1.2).

**5.1. Fully Computable Error Estimate.** As remarked before, the error estimator in Theorem 19 is not fully computable; compare also with the remarks below Theorem 19. In fact, although the term  $\int_{I_n} \|f - f_n\|_{*,(\nabla U_n)}^2 dt$  can be interpreted as data oscillation, determining the dual quasi-norm is an infinite dimensional problem. In the case of the (linear) heat equation one can benefit from a Friedrichs inequality in order to deduce the computable bound

$$\int_{I_n} \|f - f_n\|_{H^{-1}}^2 dt \lesssim \int_{I_n} \|f - f_n\|_2^2 dt.$$

Numerical computations in [KMSS11] indicate that in some cases this term may even dominate the temporal error indicator  $\Theta_n$ . Whence it is worthwhile to further elaborate upon this issue. Indeed, we can weaken the error measure and prove the following result.

**Theorem 20** (Fully Computable Estimate). *Suppose that the conditions of Theorem 19 hold and let additionally  $f \in L^1(I; L^2(\Omega))$ . Then, the following computable error estimate holds:*

$$\begin{aligned} \sup_{t \in I} \|u(t) - \mathcal{U}(t)\|_2^2 + \int_I \|u - \mathcal{U}\|_{(\nabla u)}^2 dt \\ \lesssim \|u(0) - \mathcal{U}(0)\|_2^2 + \sum_{n=1}^N \tau_n (\mathcal{E}_n^2 + \Theta_n^2) + \left( \sum_{n=1}^N \int_{I_n} \|f - f_n\|_2 dt \right)^2. \end{aligned}$$

*Proof.* We start from (3.7) setting  $v = \mathcal{U} \in \mathbb{W}(I)$ , i. e., for a.e.  $t \in I_n$  we have

$$\begin{aligned} \frac{d}{dt} \|u - \mathcal{U}\|_2^2 + \|u - \mathcal{U}\|_{(\nabla u)}^2 &\approx \langle \text{Res}(\mathcal{U}), u - \mathcal{U} \rangle \\ &= \langle \text{Res}_h(\mathcal{U}), u - \mathcal{U} \rangle + \langle \text{Res}_{\tau}(\mathcal{U}), u - \mathcal{U} \rangle \\ &\quad + \langle f - f_n, u - \mathcal{U} \rangle_{\Omega}. \end{aligned}$$

Integrating with respect to  $t$  yields

$$\begin{aligned} \|u(t) - \mathcal{U}(t)\|_2^2 &\lesssim \int_I \langle \text{Res}_h(\mathcal{U}), u - \mathcal{U} \rangle + \langle \text{Res}_{\tau}(\mathcal{U}), u - \mathcal{U} \rangle dt + \|u(0) - \mathcal{U}(0)\|_2^2 \\ &\quad + \sum_{n=1}^N \int_{I_n} \|f - f_n\|_2 \|u - \mathcal{U}\|_2 ds \end{aligned}$$

and

$$\begin{aligned} \int_I \|u - \mathcal{U}\|_{(\nabla u)}^2 &\lesssim \int_I \langle \text{Res}_h(\mathcal{U}), u - \mathcal{U} \rangle + \langle \text{Res}_\tau(\mathcal{U}), u - \mathcal{U} \rangle dt + \|u(0) - \mathcal{U}(0)\|_2^2 \\ &\quad + \sum_{n=1}^N \int_{I_n} \|f - f_n\|_2 \|u - \mathcal{U}\|_2 dt. \end{aligned}$$

Now similar arguments as in the proof of Theorem 13 and Young's inequality lead to

$$\begin{aligned} \sup_{t \in I} \|u(t) - \mathcal{U}(t)\|_2^2 &+ \int_I \|u - \mathcal{U}\|_{(\nabla u)}^2 dt \\ &\lesssim \int_I \|\text{Res}_h(\mathcal{U})\|_{*,(\nabla \mathcal{U})}^2 + \|\text{Res}_\tau(\mathcal{U})\|_{*,(\nabla \mathcal{U})}^2 dt \\ &\quad + \|u(0) - \mathcal{U}(0)\|_2^2 \\ &\quad + \frac{\delta}{2} \sup_{t \in I} \|u(t) - \mathcal{U}(t)\|_2^2 + \frac{1}{2\delta} \left( \sum_{n=1}^N \int_{I_n} \|f - f_n\|_2 dt \right)^2. \end{aligned}$$

Choosing  $\delta > 0$  small enough we arrive at

$$\begin{aligned} \sup_{t \in I} \|u(t) - \mathcal{U}(t)\|_2^2 &+ \int_I \|u - \mathcal{U}\|_{(\nabla u)}^2 dt \\ &\lesssim \int_I \|\text{Res}_h(\mathcal{U})\|_{*,(\nabla \mathcal{U})}^2 + \|\text{Res}_\tau(\mathcal{U})\|_{*,(\nabla \mathcal{U})}^2 dt \\ &\quad + \|u(0) - \mathcal{U}(0)\|_2^2 + \left( \sum_{n=1}^N \int_{I_n} \|f - f_n\|_2 dt \right)^2. \end{aligned}$$

Arguing as in (4.6) together with Lemmas 15 and 17 proves the assertion.  $\square$

**5.2. Application to the power law.** We turn towards deducing explicit estimates in the particular case of the power law; compare with (1.2). However, the presentation may also serve as a guideline for more complicated nonlinearities.

Obviously, the N-function  $\phi(t) := \frac{1}{p}t^p$ ,  $t \geq 0$ , corresponding to the power law satisfy the conditions of Theorems 19 and 20 for  $p \geq \frac{2d}{d+2}$ . Its dual is given by  $\phi^*(t) = \frac{1}{p'}t^{p'}$ , with  $\frac{1}{p} + \frac{1}{p'} = 1$  and we have that  $\Delta_2(\phi) = 2^p$  and  $\Delta_2(\phi^*) = 2^{p'}$ .

In order to provide a simple representation of the estimators  $\Theta_n$  and  $\mathcal{E}_n$ ,  $n = 1, \dots, N$ , we have to consider the family of shifted N-functions  $\{\phi_a\}_{a \geq 0}$ . In fact, by applying Proposition 3 to the power law, we end up with

$$(5.1) \quad \phi_{|\mathbf{P}|}(|\mathbf{Q}|) \approx (|\mathbf{P}| + |\mathbf{Q}|)^{p-2} |\mathbf{Q}|^2$$

for all  $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^d$ . Consequently (4.5) implies

$$\Theta_n^2 \approx \int_{\Omega} (|\nabla U_n| + |\nabla(U_n - U_{n-1})|)^{p-2} |\nabla(U_n - U_{n-1})|^2 dx.$$

Similarly we obtain from Lemma 2 for the dual N-function

$$(\phi_{|\mathbf{P}|})^*(|t|) \approx (\phi^*)_{\phi'(|\mathbf{P}|)}(|t|) \approx \phi^{*''}(\phi'(|\mathbf{P}|) + |t|) |t|^2,$$

for all  $\mathbf{P} \in \mathbb{R}^d$ ,  $t \in \mathbb{R}$ , and thus

$$(5.2) \quad (\phi_{|\mathbf{P}|})^*(|t|) \approx (|\mathbf{P}|^{p-1} + |t|)^{p'-2} |t|^2.$$

This expression serves to deal with the element residual of  $\mathcal{E}_n$ . For the jump residual, we recall from (3.1) that

$$(5.3) \quad \mathbf{F}(\mathbf{Q}) = |\mathbf{Q}|^{\frac{p-2}{2}} \mathbf{Q},$$

for all  $\mathbf{Q} \in \mathbb{R}^d$ . Hence, by combining (5.2) and (5.3) with (4.10), we arrive at

$$\begin{aligned} \mathcal{E}_n^2 \approx & \sum_{E \in \mathcal{G}_n} \int_E (|\nabla U_n|^{p-1} + |h_E f_n|)^{p'-2} |h_E f_n|^2 dx \\ & + \int_{\hat{e}E} h_E \left| \left[ |\nabla U_n|^{\frac{p-2}{2}} \nabla U_n \right] \right|^2 d\sigma. \end{aligned}$$

We conclude the paper with one last remark on the computability of the data error.

**Remark 21.** *In the particular case of the power law,  $p \geq 2$ , one may observe from (5.1) that  $\phi_a(t)$  is monotone increasing in  $a \geq 0$ . Thanks to the definition of the dual quasi norm (3.2) and Young's inequality (2.11), we thus have for  $\delta > 0$  that*

$$\begin{aligned} \|f - f_n\|_{*,(\nabla U_n)}^2 dt &\leq \sup_{w \in \mathbb{V}} \left\{ \int_{\Omega} C_{\delta} \phi^*(|f - f_n|) dt + \delta \phi(|w|) - \phi_{|\nabla U_n|}(|\nabla w|) dx \right\} \\ &\leq \sup_{w \in \mathbb{V}} \left\{ \int_{\Omega} C_{\delta} \phi^*(|f - f_n|) dt + \delta \phi(|w|) - \phi(|\nabla w|) dx \right\} \end{aligned}$$

a.e. in  $I_n$ . Since there is no shift involved, we can apply a Friedrichs inequality  $\int_{\Omega} \phi(|w|) dx \lesssim \int_{\Omega} \phi(|\nabla w|) dx$  (see [Kre08]) and obtain

$$(5.4) \quad \|f - f_n\|_{*,(\nabla U_n)}^2 \lesssim \int_{\Omega} \phi^*(|f - f_n|) dt = \frac{1}{p'} \int_{\Omega} |f - f_n|^{p'} dx.$$

a.e. in  $I_n$ .

**Acknowledgement:** Part of this work was carried out during a stay at the Mathematical Institute of the University of Oxford. This stay was financed by the German Research Foundation DFG within the research grant Kr 3984/1-1. Last but not least I would like to thank Prof. Endre Süli for his great hospitality and many fruitful discussions.

## REFERENCES

- [BBDR10] L. Belenki, L. Berselli, L. Diening, and M. Růžička, *On the finite element approximation of p-stokes systems*, accepted, 2010.
- [BDK11] L. Belenki, L. Diening, and C. Kreuzer, *Optimality of an adaptive finite element method for the p-laplacian equation*, IMA Journal of Numerical Analysis, doi: 10.1093/imanum/drr016 (online first).
- [BL93a] J. W. Barrett and W. B. Liu, *Finite element approximation of the p-Laplacian.*, Math. Comput. **61** (1993), no. 204, 523–537.
- [BL93b] ———, *Finite element error analysis of a quasi-newtonian flow obeying the Carreau or power law.*, Numer. Math. **64** (1993), no. 4, 433–453.
- [BL94] ———, *Finite element approximation of the parabolic p-Laplacian.*, SIAM J. Numer. Anal. **31** (1994), no. 2, 413–428.
- [CF06] Zhiming Chen and Jia Feng, *An adaptive finite element algorithm with reliable and efficient error control for linear parabolic problems*, Math. Comp. **73** (2006), 1167–1042.
- [Cho89] S.-S. Chow, *Finite element error estimates for nonlinear elliptic equations of monotone type*, Numer. Math. **54** (1989), no. 4, 373–393.
- [Cia78] P. G. Ciarlet, *The finite element method for elliptic problems*, North-Holland Publishing Co., Amsterdam, 1978, Studies in Mathematics and its Applications, Vol. 4.
- [CK03] C. Carstensen and R. Klose, *A posteriori finite element error control for the p-Laplace problem.*, SIAM J. Sci. Comput. **25** (2003), no. 3, 792–814.
- [CLY06a] C. Carstensen, W. Liu, and N. Yan, *A posteriori FE error control for p-Laplacian by gradient recovery in quasi-norm*, Math. Comp. **75** (2006), no. 256, 1599–1616 (electronic).
- [CLY06b] ———, *A posteriori error estimates for finite element approximation of parabolic p-Laplacian*, SIAM J. Numer. Anal. **43** (2006), no. 6, 2294–2319.
- [DE08] L. Diening and F. Ettwein, *Fractional estimates for non-differentiable elliptic systems with general growth*, Forum Mathematicum **3** (2008), 523–556.

- [DER07] 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** (2007), no. 2, 457–472 (electronic).
- [DK08] L. Diening and C. Kreuzer, *Convergence of an adaptive finite element method for the  $p$ -Laplacian equation*, SIAM J. Numer. Anal. **46** (2008), no. 2, 614–638.
- [DR07a] L. Diening and M. Růžička, *Interpolation operators in Orlicz-Sobolev spaces.*, Numer. Math. **107** (2007), no. 1, 107–129.
- [DR07b] ———, *Non-Newtonian fluids and function spaces.*, Nonlinear Analysis, Function Spaces and Applications **8** (2007), 95–143.
- [EL05] C. Ebmeyer and W. Liu, *Quasi-norm interpolation error estimates for the piecewise linear finite element approximation of  $p$ -Laplacian problems*, Numer. Math. **100** (2005), no. 2, 233–258.
- [GGZ74] H. Gajewski, K. Gröger, and Klaus Zacharias, *Nichtlineare Operatorgleichungen und Operator Differentialgleichungen.*, Mathematische Lehrbücher und Monographien. II. Abteilung. Band 38. Berlin: Akademie-Verlag. IX, 281 S., 1974 (German).
- [GM75] R. Glowinski and A. Marrocco, *Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité, d'une classe de problèmes de Dirichlet non linéaires*, Rev. Française Automat. Informat. Recherche Opérationnelle RAIRO Analyse Numérique **9** (1975), no. R-2, 41–76.
- [KK91] V. Kokilashvili and M. Krbeć, *Weighted inequalities in Lorentz and Orlicz spaces.*, Singapore etc.: World Scientific Publishing Co. Pte. Ltd., 1991.
- [KMSS11] C. Kreuzer, C. A. Möller, A. Schmidt, and K. G. Siebert, *Design and convergence analysis for an adaptive discretization of the heat equation*, Preprint SM-DU-724 Universität Duisburg-Essen, December 2010, to appear in IMA Journal of Numerical Analysis., 2011.
- [KR61] M. A. Krasnosel'skij and Ya. B. Rutitskij, *Convex functions and Orlicz spaces.*, Groningen-The Netherlands: P. Noordhoff Ltd., 1961.
- [Kre08] C. Kreuzer, *A convergent adaptive Uzawa finite element method for the nonlinear Stokes problem*, Ph.D. thesis, Mathematisches Institut, Universität Augsburg, 2008.
- [Kre12] ———, *Analysis of an adaptive uzawa finite element method for the nonlinear Stokes problem*, MATHEMATICS OF COMPUTATION **81** (2012), 21–55.
- [LY01] W. Liu and N. Yan, *Quasi-norm local error estimators for  $p$ -Laplacian.*, SIAM J. Numer. Anal. **39** (2001), no. 1, 100–127.
- [LY02] ———, *On quasi-norm interpolation error estimation and a posteriori error estimates for  $p$ -Laplacian.*, SIAM J. Numer. Anal. **40** (2002), no. 5, 1870–1895.
- [Mus83] J. Musielak, *Orlicz spaces and modular spaces.*, Lecture Notes in Mathematics. 1034. Berlin etc.: Springer-Verlag, 1983.
- [RR91] M. M. Rao and Z. D. Ren, *Theory of Orlicz spaces.*, Pure and Applied Mathematics, 146. New York etc.: Marcel Dekker, Inc., 1991.
- [SZ90] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Mathematics of Computation **54** (1990), no. 190, 483–493.
- [Ver03] R. Verfürth, *A posteriori error estimates for finite element discretizations of the heat equation.*, Calcolo **40** (2003), no. 3, 195–212.
- [Ver05] ———, *Robust a posteriori error estimates for nonstationary convection-diffusion equations*, SIAM J. Numer. Anal. **43** (2005), no. 4, 1783–1802 (electronic).
- [Zei90] E. Zeidler, *Nonlinear functional analysis and its applications. II/A: Linear monotone operators. Transl. from the German by the author and by Leo F. Boron.*, New York etc.: Springer-Verlag, 1990.

CHRISTIAN KREUZER, FAKULTÄT FÜR MATHEMATIK, RUHR-UNIVERSITÄT BOCHUM, UNIVERSITÄTSSTRASSE 150, D-44801 BOCHUM, GERMANY

URL: <http://www.ruhr-uni-bochum.de/ffm/Lehrstuehle/Kreuzer/index.html>

E-mail address: [christian.kreuzer@rub.de](mailto:christian.kreuzer@rub.de)