Linear convergence of an adaptive finite element method for the p-Laplacian equation

Lars Diening, Christian Kreuzer

Angaben zur Veröffentlichung / Publication details:

Diening, Lars, and Christian Kreuzer. 2007. "Linear convergence of an adaptive finite element method for the p-Laplacian equation." Augsburg: Universität Augsburg.

Universität Augsburg

Institut für Mathematik

Lars Diening, Christian Kreuzer

Linear convergence of an adaptive finite element method for the p-Laplacian equation

Preprint Nr. 024/2007 — 03. Juli 2007

Institut für Mathematik, Universitätsstraße, D-86 135 Augsburg http://www.math.uni-augsburg.de/

Impressum:

Herausgeber:

Institut für Mathematik
Universität Augsburg
86135 Augsburg
`http://www.math.uni-augsburg.de/forschung/preprint/`

ViSdP:

Christian Kreuzer
Institut für Mathematik
Universität Augsburg
86135 Augsburg

LINEAR CONVERGENCE OF AN ADAPTIVE FINITE ELEMENT METHOD FOR THE p-LAPLACIAN EQUATION

LARS DIENING AND CHRISTIAN KREUZER

ABSTRACT. We study an adaptive finite element method for the p-Laplacian like PDE's using piecewise linear, continuous functions. The error is measured by means of the quasi-norm of Barrett and Liu. We provide residual based error estimators without a gap between the upper and lower bound. We show linear convergence of the algorithm which is similar to the one of Morin, Nochetto, and Siebert. All results are obtained without extra marking for the oscillation.

1. INTRODUCTION

Let Ω be a polyhedral, bounded domain in $\mathbb{R}^d$, $d \in \mathbb{N}$. We consider the following system of nonlinear structure

$$\begin{aligned} -\operatorname{div}(\mathbf{A}(\nabla u)) &= f && \text{in } \Omega, \\ u &= 0 && \text{on } \partial\Omega. \end{aligned} \tag{1.1}$$

Our considerations include in particular the case of the p-Laplacian, where

$$\mathbf{A}(\nabla u) = (\kappa + |\nabla u|)^{p-2}\, \nabla u, \tag{1.2}$$

with $1 < p < \infty$, $\kappa \geq 0$, $f \in L^{p'}(\Omega)$, and $\frac{1}{p} + \frac{1}{p'} = 1$.

The purpose of this paper is to present a linear convergence result for an adaptive finite element method AFEM applied to the nonlinear Laplace equation (1.1). As is common practice the adaptive finite element method consists of a loop

$$\text{SOLVE} \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \rightarrow \text{REFINE} \tag{AFEM}$$

starting from a initial triangulation of Ω. To be more specific, the finite element problem on the current mesh is solved, then the aposteriori error estimator is computed and finally with its help elements are marked for refinement. The marking strategy resorts to *Dörfler's marking* introduced in [Dör96]. The algorithm uses piecewise linear, continuous finite elements, whereas the refinement is realized by newest vertex bisection. This produces a sequence of weak finite element solutions u_k of (1.1) in nested finite element spaces V_k.

The main result states linear convergence of u_k to the weak solution u of (1.1). In particular, we show that there exists $\alpha \in (0,1)$, $C > 0$ with

$$\|\mathbf{F}(\nabla u_k) - \mathbf{F}(\nabla u)\|_2^2 + \operatorname{osc}_k^2(f) \leq \alpha^{2k} C,$$

where the vector field $\mathbf{F}$ arises from the vector field $\mathbf{A}$ by $\mathbf{F}(\mathbf{a}) := |\mathbf{A}(\mathbf{a})|^{\frac{1}{2}}\, |\mathbf{a}|^{-\frac{1}{2}}\mathbf{a}$. The L^2 norm of the error $\|\mathbf{F}(\nabla u_k) - \mathbf{F}(\nabla u)\|_2^2$ measured in terms of $\mathbf{F}$ is equivalent to the so called *quasi norm* $\|\nabla u_k - \nabla u\|_{(p)}^2$ introduced by Barrett and Liu,

Date: July 4, 2007.

2000 *Mathematics Subject Classification.* 65N15, 65N30, 35J60, 35J70.

Key words and phrases. Convergence, a posteriori error estimator, p-Laplacian, quasi norms, error reduction, oscillation.

L. Diening is indebted to the Landesstiftung Baden-Württemberg for facilitating the analysis entailed in this paper.

C. Kreuzer has been supported by Project C.1 of the DFG-Research-Unit "Nonlinear Partial Differential Equations" Generalized Newtonian fluids and electrorheological fluids.

cf. [BL94a] and Remark 4. The *quasi norm* was a breakthrough in the numerical investigation of (1.1). In particular, Barrett and Liu obtained the best approximation property of the conforming, finite element solution $u_h \in V_h$ in terms of quasi norms, i.e.

$$\|\nabla v - \nabla v_h\|_{(p)}^2 \leq c \min_{\psi_h \in V_h} \|\nabla v - \nabla \psi_h\|_{(p)}^2,$$

In [EL05] it has been proved by Ebmeyer and Liu that for piecewise linear, continuous finite elements and $p > \frac{2d}{d+2}$ the best approximation error can be estimated as

$$\min_{\psi_h \in V_h} \|\nabla v - \nabla \psi_h\|_{(p)}^2 \leq c\, h^2 \int_\Omega (\kappa + |\nabla u|)^{p-2} \left|\nabla^2 u\right|^2 dx. \tag{1.3}$$

Recently, Diening and Růžička improved these results in [DR06] to the case $p > 1$ admitting also more general finite element spaces. In particular, they showed

$$\|\mathbf{F}(\nabla v) - \mathbf{F}(\nabla v_h)\|_2^2 \leq c \min_{\psi_h \in V_h} \|\mathbf{F}(\nabla v) - \mathbf{F}(\nabla \psi_h)\|_2^2$$

and

$$\|\mathbf{F}(\nabla v) - \mathbf{F}(\nabla \Pi_h v)\|_2^2 \leq c\, h^2 \, \|\nabla \mathbf{F}(\nabla u)\|_2^2, \tag{1.4}$$

where Π_h is a suitable interpolation operator, e.g. the Scott-Zhang operator. We want to mention that the right hand sides of (1.3) and (1.4) are proportional. They express the natural regularity of a *strong* solution of (1.1) (cf. [Giu03], [BL93a], [ELS05], [Ebm05]).

The technique of quasi-norms founds its way into a posteriori analysis in the work of Liu and Yan [LY01, LY02]. They show that

$$c\, \eta_h^2 - C \operatorname{osc}_h^2(f) \leq \|\nabla u - \nabla u_h\|_{(p)}^2 \leq C \left(\eta_h^2 + \widetilde{\eta}^2\right)$$

for residual based estimators η_h. Numerical experiments [LY01, CK06] indicate, that these new estimators are indeed sharper and lead to more efficient meshes than existing ones. However for convergence analysis the additional term $\widetilde{\eta}^2$ causes problems, since it forms a gap between the left and the right hand side. In this work we are able to overcome this drawback and prove estimates avoiding $\widetilde{\eta}^2$ (see Lemma 8 and Corollary 11):

$$c\, \eta_h^2 - C \operatorname{osc}_h^2(f) \leq \|\mathbf{F}(\nabla u) - \mathbf{F}(\nabla u_h)\|_2^2 \leq C\, \eta_h^2. \tag{1.5}$$

Recently a posteriori error estimators by gradient recovery have been studied in [CLY06].

Dörfler was the first who proved in [Dör96] linear error reduction of (AFEM) for the linear Laplacian, if the data oscillation is small enough. Later, this additional assumption has been removed by Morin, Nochetto, and Siebert in [MNS00] by additional marking for oscillation. The first convergence result for the nonlinear Laplacian is stated by Veeser [Ve02]. There residual based estimators for the $W^{1,p}$ norm are used. Since there appears a gap in the power between the upper and lower estimates this prevents to prove linear convergence.

The convergence results in the linear case are heavily based on Galerkin orthogonality and the Pythagorean Theorem which yield

$$|||u_h - u|||^2 = |||u_H - u|||^2 - |||u_H - u_h|||^2. \tag{1.6}$$

in the energy norm. To overcome the lack of orthogonality in the non-linear case we proceed as follows: We prove that the energy difference of weak solutions in nested spaces $V_1 \subset V_2$ is proportional to the quasi-norm distance, i.e.

$$\mathcal{J}(u_1) - \mathcal{J}(u_2) \sim \|\nabla u_1 - \nabla u_2\|_{(p)}^2 \sim \|\mathbf{F}(\nabla u_1) - \mathbf{F}(\nabla u_2)\|_2^2,$$

where $\mathcal{J}(u)$ is the energy functional of (1.1), $u_1 \in V_1$, and $u_2 \in V_2$, see (2.10) and Lemma 16. This property and the trivial equality

$$\mathcal{J}(u_h) - \mathcal{J}(u) = \big(\mathcal{J}(u_H) - \mathcal{J}(u)\big) - \big(\mathcal{J}(u_H) - \mathcal{J}(u_h)\big)$$

is our substitute for the orthogonality of the error (1.6).

In the linear, symmetric case it is possible to consider the reduction of the error and the oscillation independently, since the oscillation is solely dependent on the data f. Mekchay and Nochetto showed linear reduction of the sum of error and oscillation for non-symmetric second order linear elliptic PDE in [MN05]. In this case oscillation and error are coupled. A similar effect appears in our non-linear setting. We introduce a new proof for error reduction, which enables us to manage without extra marking for oscillation, see Remark 14. Our proof permits to use the fact that oscillation is dominated by the error indicator. Moreover, we prove a strict reduction of the difference of energies plus the oscillation in each step.

An essential tool in our calculations in the use of *shifted* N-functions, namely φ_a. They are closely related to the quasi-norms, which is best expressed by the relation

$$\big(\mathbf{A}(\mathbf{a}) - \mathbf{A}(\mathbf{b})\big) \cdot (\mathbf{a} - \mathbf{b}) \sim \big|\mathbf{F}(\mathbf{a}) - \mathbf{F}(\mathbf{b})\big|^2 \sim \varphi_{|\mathbf{a}|}(|\mathbf{a} - \mathbf{b}|)$$

for $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$. See Lemma 3 for more details. The shifted N-functions enable us to handle more general non-linear equations than the p-Laplacian, namely the φ-Laplacian from (2.1). But most important, the shifted N-functions simplify and clarify the calculations significantly also in the case of the p-Laplacian.

2. Preliminaries

We first introduce the nonlinear Dirichlet problem. Thereby the nonlinear partial differential operator called φ-Laplacian is defined via a certain function $\varphi : \mathbb{R}^{\geq 0} \to \mathbb{R}^{\geq 0}$. The most popular case of such operators is the p-Laplacian which corresponds to the function $\varphi(t) := \int_0^t (\kappa + s)^{p-2} s\, ds$. As mentioned before the treatment of the nonlinear Laplacian via *N-functions* simplifies and clarifies calculations. Assumptions on φ and related properties are discussed subsequently. Afterwards the weak formulation of the problem is stated along with the corresponding minimizing problem.

Let Ω be a polyhedral, bounded domain in $\mathbb{R}^d$, $d \in \mathbb{N}$. In the center of our considerations are solutions of the φ-Laplacian problem, i.e.

$$\begin{aligned} -\operatorname{div}(\mathbf{A}(\nabla u)) &= f && \text{in } \Omega, \\ u &= 0 && \text{on } \partial\Omega, \end{aligned} \tag{2.1}$$

with

$$\mathbf{A}(\nabla u) = \varphi'(|\nabla u|)\frac{\nabla u}{|\nabla u|}. \tag{2.2}$$

2.1. **Assumptions on φ and resulting properties.** Now we shed light on the considered function φ. To go not beyond the scope of this work we give only a short sketch of the underlying theory. The following definitions and results are standard in the theory of Orlicz functions and can for example be found in [RR91]. A few assertions are also proved at the Appendix 5.

We use $c, C > 0$ (no index) as generic constants, i.e. their value may change from line to line but does not depend on the important variables. Furthermore, we write $f \sim g$ iff $c\, f \leq g \leq C\, f$.

A continuous function $\varphi : \mathbb{R}^{\geq 0} \to \mathbb{R}^{\geq 0}$ is said to be an *N-function*, iff

- φ is continuous and convex with $\varphi(0) = 0$;
- there exists a derivative φ' of φ which is right continuous, non-decreasing and satisfies $\varphi'(0) = 0$, $\varphi'(t) > 0$ for $t > 0$, and $\lim_{t\to\infty} \varphi'(t) = \infty$.

The N in N-function stands for *nice* function. Furthermore, φ satisfies the Δ_2-condition, iff it holds

- $\varphi(2t) \leq C\,\varphi(t)$ uniformly in $t \geq 0$.

We denote the smallest such constant by $\Delta_2(\varphi)$. Since $\varphi(t) \leq \varphi(2t)$ the Δ_2-condition means that $\varphi(t)$ and $\varphi(2t)$ are proportional. Note that if $\Delta_2(\varphi) < \infty$ then $\varphi(t) \sim \varphi(a\,t)$ uniformly in $t \geq 0$ for any fixed $a > 1$.

Define $(\varphi')^{-1} : \mathbb{R}^{\geq 0} \to \mathbb{R}^{\geq 0}$ as

$$(\varphi')^{-1}(t) := \sup\{u \in \mathbb{R}^{\geq 0} : \varphi'(u) \leq t\}.$$

If φ' is strictly increasing then $(\varphi')^{-1}$ is the inverse function of φ'. By the definition

$$\varphi^*(t) := \int_0^t (\varphi')^{-1}(s)\,ds$$

the so called complementary function φ^* of φ is again an N-function and $(\varphi^*)'(t) = (\varphi')^{-1}(t)$ for $t > 0$. Note that $(\varphi^*)^* = \varphi$.

Assume that $\Delta_2(\varphi), \Delta_2(\varphi^*) < \infty$. Then for all $\delta > 0$ there exists $C_\delta > 0$ (only depending on $\Delta_2(\varphi)$, and $\Delta_2(\varphi^*)$), such that for all $s, t \geq 0$ hold

$$\begin{aligned} s\,t &\leq C_\delta\,\varphi(s) + \delta\,\varphi^*(t), \\ s\,t &\leq \delta\,\varphi(s) + C_\delta\,\varphi^*(t). \end{aligned} \tag{2.3}$$

This inequalities are called Young's inequality. Moreover, for all $t > 0$ there exists $s > 0$ such that

$$s\,t = \varphi(s) + \varphi^*(t). \tag{2.4}$$

Further basic inequalities are for all $t \geq 0$

$$\begin{aligned} t \leq \varphi^{-1}(t)\,(\varphi^*)^{-1}(t) &\leq 2t, \\ \frac{t}{2}\varphi'\Big(\frac{t}{2}\Big) \leq \varphi(t) &\leq t\,\varphi'(t), \\ \varphi\bigg(\frac{\varphi^*(t)}{t}\bigg) \leq \varphi^*(t) &\leq \varphi\bigg(\frac{2\,\varphi^*(t)}{t}\bigg). \end{aligned} \tag{2.5}$$

Therefor, uniformly in $t \geq 0$

$$\varphi^{-1}(t)\,(\varphi^*)^{-1}(t) \sim t, \qquad \varphi(t) \sim \varphi'(t)\,t, \qquad \varphi^*\big(\varphi'(t)\big) \sim \varphi(t), \tag{2.6}$$

where the constants only depend on $\Delta_2(\varphi)$ and $\Delta_2(\varphi^*)$.

As in [DE05, DR06] we require the following properties about our function φ:

Assumption 1. Let φ be an N-function with $\Delta_2(\varphi) < \infty$, $\Delta_2(\varphi^*) < \infty$, and $\varphi \in C^2(0,\infty)$ such that

$$\varphi'(t) \sim t\,\varphi''(t) \tag{2.7}$$

uniformly in $t \geq 0$.

It is shown in [DE05] that φ satisfies Assumption 1 if and only if φ^* satisfies Assumption 1. Moreover, it is shown that we have for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$

$$\begin{aligned} \big(\mathbf{A}(\mathbf{a}) - \mathbf{A}(\mathbf{b})\big) \cdot (\mathbf{a} - \mathbf{b}) &\geq c\,\varphi''(|\mathbf{a}| + |\mathbf{b}|)\,|\mathbf{a} - \mathbf{b}|^2, \\ |\mathbf{A}(\mathbf{a}) - \mathbf{A}(\mathbf{b})| &\leq C\,\varphi''(|\mathbf{a}| + |\mathbf{b}|)\,|\mathbf{a} - \mathbf{b}|, \end{aligned} \tag{2.8}$$

where c, C only depend on $\Delta_2(\varphi)$, $\Delta_2(\varphi)$, and the constant in (2.7).

Remark 2. *The most important example of such functions is certainly the p-Laplacian. Thereby $\varphi(t) := \int_0^t (\kappa + s)^{p-2}\,s\,ds$ with $1 < p < \infty$ and $\kappa \geq 0$. This* *function satisfies Assumption 1. If $\kappa = 0$, then Young's inequality (2.3) coincides with the well known classical Young's inequality*

$$s\,t \leq \delta \frac{1}{p} t^p + \delta^{\frac{1}{p-1}} \frac{1}{q} s^q,$$

where $q \in (1,\infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$. Moreover, with $\mathbf{A}(\mathbf{a}) = \varphi'(|\mathbf{a}|)\frac{\mathbf{a}}{|\mathbf{a}|} = (\kappa + |\mathbf{a}|)^{p-2}\mathbf{a}$ (2.8) corresponds to the well known monotonicity and coercivity inequalities

$$\begin{aligned}\big((\kappa+|\mathbf{a}|)^{p-2}\mathbf{a} - (\kappa+|\mathbf{b}|)^{p-2}\mathbf{b}\big)(\mathbf{a}-\mathbf{b}) &\geq c\,(\kappa+|\mathbf{a}|+|\mathbf{b}|)^{p-2}\,|\mathbf{a}-\mathbf{b}|^2,\\ |(\kappa+|\mathbf{a}|)^{p-2}\mathbf{a} - (\kappa+|\mathbf{b}|)^{p-2}\mathbf{b}| &\leq C\,(\kappa+|\mathbf{a}|+|\mathbf{b}|)^{p-2}\,|\mathbf{a}-\mathbf{b}|.\end{aligned}$$

for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$ (see e.g. [GM75, BL93a]).

2.2. **Weak formulation of the φ-Laplacian problem and corresponding minimizing problem.** We introduce the weak form of the nonlinear equation (2.1). To proceed so we first have to introduce analytical background. For details we refer to [RR91].

In the following the function φ will be a fixed N-function as stated in Assumption 1. By L^φ and $W^{1,\varphi}$ we denote the classical Orlicz and Sobolev-Orlicz spaces i.e. $g \in L^\varphi$ iff $\int \varphi(|g|)\,dx < \infty$ and $g \in W^{1,\varphi}$ iff $g, \nabla g \in L^\varphi$. The norm on L^φ is given by $\|f\|_\varphi = \inf\{\lambda > 0 : \int \varphi(f/\lambda)\,dx \leq 1\}$. By $W_0^{1,\varphi}(\Omega)$ we denote the closure of $C_0^\infty(\Omega)$ in $W^{1,\varphi}(\Omega)$.

The weak formulation reads as follows: For $f \in L^{\varphi^*}(\Omega) \subset (W_0^{1,\varphi}(\Omega))^*$ find $u \in W_0^{1,\varphi}(\Omega)$ with

$$\langle \mathbf{A}(\nabla u), \nabla v\rangle = \langle f, v\rangle \qquad \text{for all } v \in W_0^{1,\varphi}(\Omega). \tag{2.9}$$

The theory of monotone operators ensures the unique existence of u. Moreover, u is the unique minimizer of the *energy functional*

$$\mathcal{J}(u) := \int_\Omega \varphi(|\nabla u|)\,dx - \int_\Omega u\,f\,dx \to \min! \tag{2.10}$$

3. A Posteriori Analysis

In this section we first dicuss our concept of distance. In particular we introduce shifted N-functions and construct an measure of distance related to the nonlinear problem (2.1). In case of the p-Laplacian this concept is equivalent to the quasi-norm introduced by Barrett and Liu [BL93a]. In the following we introduce the finite element spaces and note an interpolation inequality. In the second part of this section we construct a posteriori upper and lower bounds for the error of a finite element solution to the continuous solution. Finally, we discuss a discrete local lower bound, i.e., a lower bound for the distance between two different finite element solutions.

3.1. **Concept of Distance.** Let φ be again a fixed N-function. We define another N-function ψ by

$$\psi'(t) := \sqrt{\varphi'(t)\,t}$$

and set for $\mathbf{a} \in \mathbb{R}^d$

$$\mathbf{F}(\mathbf{a}) := \psi'(|\mathbf{a}|)\,\frac{\mathbf{a}}{|\mathbf{a}|} = \sqrt{\varphi'(|\mathbf{a}|)\,|\mathbf{a}|}\;\frac{\mathbf{a}}{|\mathbf{a}|} = \sqrt{|\mathbf{A}(\mathbf{a})|\,|\mathbf{a}|}\;\frac{\mathbf{a}}{|\mathbf{a}|}. \tag{3.1}$$

It is shown in [DE05] that since φ satisfies Assumption 1 also ψ, ψ^*, and φ^* satisfy Assumption 1. It is also shown that as a consequence (2.8) holds with $\mathbf{A}, \varphi$ replaced by $\mathbf{F}, \psi$. In addition we introduce a family of N-function $\{\varphi_a\}_{a\geq 0}$ by

$$\frac{\varphi_a'(t)}{t} := \frac{\varphi'(a+t)}{a+t} \tag{3.2}$$

which owing to (2.7) implies $\varphi_a''(t) \sim \varphi''(a+t)$ uniformly in $a, t \geq 0$. The functions φ_a are called *shifted* N-functions. The basic properties of φ_a are summarized in the appendix. The connection between $\mathbf{A}$, $\mathbf{F}$, and $\{\varphi_a\}_{a\geq 0}$ is best reflected in the following lemma from [DE05].

Lemma 3. *Let φ satisfy Assumption 1 and let $\mathbf{A}$ and $\mathbf{F}$ be defined by (2.2) and (3.1). Then*

$$\begin{aligned}
&\text{(3.3a)} & \big(\mathbf{A}(\mathbf{a}) - \mathbf{A}(\mathbf{b})\big) \cdot \big(\mathbf{a} - \mathbf{b}\big) &\sim \big|\mathbf{F}(\mathbf{a}) - \mathbf{F}(\mathbf{b})\big|^2 \\
&\text{(3.3b)} & &\sim \varphi_{|\mathbf{a}|}(|\mathbf{a} - \mathbf{b}|), \\
&\text{(3.3c)} & &\sim |\mathbf{a} - \mathbf{b}|^2\, \varphi''\big(|\mathbf{a}| + |\mathbf{b}|\big),
\end{aligned}$$

uniformly in $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$. If $\varphi''(0)$ does not exists, the expression in (3.3c) is continuously extended by zero for $|\mathbf{a}| = |\mathbf{b}| = \mathbf{0}$. Moreover

$$\text{(3.3d)} \qquad |\mathbf{A}(\mathbf{a}) - \mathbf{A}(\mathbf{b})| \sim \varphi'_{|\mathbf{a}|}(|\mathbf{a} - \mathbf{b}|),$$

$$\text{(3.3e)} \qquad \mathbf{A}(\mathbf{b}) \cdot \mathbf{b} \sim |\mathbf{F}(\mathbf{b})|^2 \sim \varphi(|\mathbf{b}|)$$

uniformly in $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$.

We use the equivalences (3.3) extensively in this paper without referring.
An immediately consequence of Lemma 3 is

Corollary 4. *We have for all $u, v \in W^{1,\varphi}(\Omega)$*

$$\begin{aligned}
\int_\Omega \big(\mathbf{A}(\nabla u) - \mathbf{A}(\nabla v)\big) \cdot (\nabla u - \nabla v)\, dx &\sim \|\mathbf{F}(\nabla u) - \mathbf{F}(\nabla v)\|_2^2 \\
&\sim \int_\Omega \varphi_{|\nabla u|}(|\nabla u - \nabla v|)\, dx.
\end{aligned}$$

Remark 5. *In the case of the p-Laplacian, i.e. $\mathbf{A}(\nabla v) = (\kappa + |\nabla v|)^{p-2}\nabla v$ and $\varphi'(t) = (\kappa + t)^{p-2}\, t$ with $1 < p < \infty$ and $\kappa \geq 0$ we have for all $\mathbf{a} \in \mathbb{R}^d$, $t \geq 0$*

$$\mathbf{F}(\mathbf{a}) := (\kappa + |\mathbf{a}|)^{\frac{p-2}{2}}\mathbf{a}, \qquad \psi'(t) := (\kappa + t)^{\frac{p-2}{2}}\, t.$$

Moreover, for the p-Laplacian all expressions in Corrolary 4 are proportional to the quasi-norm introduced by Barrett and Liu in [BL93a]. This follows from the relation

$$\varphi'_{|\mathbf{a}|}(t) = (\kappa + |\mathbf{a}| + t)^{p-2} t$$

and

$$\begin{aligned}
\|\nabla u - \nabla v\|_{(p)}^2 &= \int_\Omega (\kappa + |\nabla u| + |\nabla u - \nabla v|)^{p-2} |\nabla u - \nabla v|^2\, dx \\
&= \int_\Omega \varphi'_{|\nabla u|}(|\nabla u - \nabla v|)|\nabla u - \nabla v|\, dx \sim \int_\Omega \varphi_{|\nabla u|}(|\nabla u - \nabla v|)\, dx.
\end{aligned}$$

This ensures in case of the p-Laplacian that all the results below can also be expressed in terms of the quasi-norm.

Addtionally, we need the following direct consequence of Lemma 3:

Corollary 6. *Let $\mathbf{A}, \varphi, \mathbf{F}$ be as in Lemma 3. Then for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$*

$$\text{(3.4)} \qquad (\varphi_{|\mathbf{a}|})^*\big(\big|\mathbf{A}(\mathbf{a}) - \mathbf{A}(\mathbf{b})\big|\big) \sim \varphi_{|\mathbf{a}|}\big(\big|\mathbf{a} - \mathbf{b}\big|\big) \sim \big|\mathbf{F}(\mathbf{a}) - \mathbf{F}(\mathbf{b})\big|^2.$$

Proof. The second relation is contained in Lemma 3. The first relation follows from (3.3d), (2.6), and $\Delta_2(\varphi) < \infty$. □

3.2. **Finite element spaces.** Let $\mathcal{T}_H$ be a conforming triangulation of Ω consisting of closed simplices $T \in \mathcal{T}_H$. Let h_T denote the diameter of the (closed) simplex $T \in \mathcal{T}_H$ and ρ_T the maximal radius of a ball that is contained in T. The maximal quotient h_T/ρ_T with $T \in \mathcal{T}_h$ is called the *shape regularity* (chunkiness) of $\mathcal{T}_H$.

Let $V_H := V(\mathcal{T}_H)$ be the space of continuous, piecewise linear finite elements over $\mathcal{T}_H$ with boundary values zero, then $V_H \subset V$. By $u_H \in V_H$ we denote the finite element solution of (2.9) with respect to V_H, i.e.

$$\langle \mathbf{A}(\nabla u_H), \nabla v_H \rangle = \langle f, v_H \rangle \qquad \text{for all } v_H \in V_H, \tag{3.5}$$

where $\langle f, v_H \rangle = \int_\Omega f\, v_H\, dx$. The theory of monotone operators ensures the existence of a unique solution.

We denote by Γ_H the set of interior faces (sides) of the triangulation $\mathcal{T}_H$. For $\gamma \in \Gamma_H$ we define N_γ as the set of elements sharing γ and S_γ as the union of these elements, i.e.

$$N_\gamma := \{T_1, T_2 \in \mathcal{T}_H \,:\, T_1 \cap T_2 = \gamma\}, \qquad S_\gamma := \bigcup_{T \in N_\gamma} T.$$

For $T \in \mathcal{T}_H$ define the set of neighbours N_T and the neighbourhood S_T by

$$N_T := \{T' \in \mathcal{T}_H \,:\, T' \cap T \in \Gamma_H\}, \qquad S_T := \bigcup_{T' \in N_T} T'.$$

For interpolation estimates (see (3.6)) we additionally need to define the patch Ω_T around T and the set of its elements ω_T by

$$\omega_T := \{T' \in \mathcal{T}_H \,:\, T' \cap T \neq \emptyset\}, \qquad \Omega_T := \bigcup_{T' \in \omega_T} T'.$$

For $\gamma \in \Gamma_H$ let $h_\gamma := \operatorname{diam}(\gamma)$. For $T \in \mathcal{T}_H$ holds $h_T \sim h_\gamma$ for each face (side) $\gamma \subset \partial T$ depending only on the shape regularity of $\mathcal{T}_H$.

Let $\Pi_H : V \to V_H$ be the Scott-Zhang interpolation operator which respects zero boundary values, see [SZ90]. Additionally Π_H satisfies $\Pi_H v_H = v_H$ for all $v_H \in V_H$. It was shown in [DR06] that for all $v \in W^{1,\varphi}(\Omega)$, $a \geq 0$, and $T \in \mathcal{T}_h$

$$\int_T \varphi_a(|v - \Pi_H v|)\, dx + \int_T \varphi_a(h_T\, |\nabla v - \nabla \Pi_H v|)\, dx \leq C \int_{\Omega_T} \varphi_a(h_T\, |\nabla v|)\, dx, \tag{3.6}$$

where C only depends on $\Delta_2(\varphi)$ and the shape regularity of $\mathcal{T}_H$.

We introduce residual based error estimators for our system (2.1). For $\gamma \in \Gamma_H$ and $T \in \mathcal{T}_H$ define the (local) *interior and the jump estimators* by

$$\eta_E^2(u_H, T) := \int_T (\varphi_{|\nabla u_H|})^* \big(h_T\, |f|\big)\, dx,$$
$$\eta_J^2(u_H, \gamma) := \int_\gamma h_\gamma\, \big|[\![\mathbf{F}(\nabla u_H)]\!]_\gamma\big|^2\, dx.$$

where $[\![\mathbf{F}(\nabla u_H)]\!]_\gamma$ denotes the jump of $\mathbf{F}(\nabla u_H)$ over the face γ. Furthermore we define for $T \in \mathcal{T}_H$ the (local) element based error indicators and the oscillation as

$$\begin{aligned} \eta^2(u_H, T) &:= \eta_E^2(u_H, T) + \sum_{\gamma \in \Gamma_H, \gamma \subset \partial T} \eta_J^2(u_H, \gamma), \\ \operatorname{osc}^2(u_H, T) &:= \inf_{f_T \in \mathbb{R}} \int_T (\varphi_{|\nabla u_H|})^* \big(h_T\, |f - f_T|\big)\, dx. \end{aligned} \tag{3.7}$$

For a subset $\hat{\mathcal{T}}_H \subset \mathcal{T}_H$ we define the *total error estimator* over $\hat{\mathcal{T}}_H$ by

$$\eta^2(u_k, \hat{\mathcal{T}}_H) := \sum_{T \in \hat{\mathcal{T}}_H} \eta^2(u_k, T).$$

Note that $\eta^2(u_H, \{T\}) = \eta^2(u_H, T)$, so there is no confusion between this definition of η^2 and (3.7). Similarly, we define the *total oscillation* on subsets of $\mathcal{T}_H$.

Furthermore we use $e_H := u - u_H$ for the difference of the solutions.

Remark 7. *In the case of the p-Laplacian we can translate the above definitions of the estimators and the oscillation. For the jump estimator we obtain by Corollary 4, Lemma 22, and Remark 5*

$$\begin{aligned}\eta_J^2(u_H, \gamma) &= \int\limits_\gamma h_\gamma \left|[\![\mathbf{F}(\nabla u_H)]\!]_\gamma\right|^2 dx \sim \int\limits_{S_\gamma} \varphi_{|\nabla u_H|}\big([\![\nabla u_H]\!]_\gamma\big) \\ &\sim \int\limits_{S_\gamma} \big(\kappa + |\nabla u_H| + |[\![\nabla u_H]\!]_\gamma|\big)^{p-2} |[\![\nabla u_H]\!]_\gamma|\, dx.\end{aligned}$$

The element residual an the oscillation can be treated in the same way. We use Lemma 22, (2.6), and Remark 5 to obtain

$$\begin{aligned}\eta_E^2(u_H, T) &= \int\limits_T (\varphi_{|\nabla u_H|})^*\big(h_T\,|f|\big)\,dx \sim \int\limits_T (\varphi^*)_{\varphi'(|\nabla u_H|)}\big(h_T\,|f|\big)\,dx \\ &\sim \int\limits_T \big(\kappa + |\nabla u_H|^{p-1} + h_T\,|f|\big)^{q-2} h_T^2\,|f|^2\,dx.\end{aligned}$$

Let f_T denote the mean value of f over T. Then the same calculations yield for the oscillation

$$\operatorname{osc}^2(u_H, T) \sim \int\limits_T \big(\kappa + |\nabla u_H|^{p-1} + h_T\,|f - f_T|\big)^{q-2} h_T^2\,|f - f_T|^2\,dx,$$

where we have used that for any N-function ϱ with $\Delta_2(\varrho) < \infty$ holds

$$\inf_{f_T} \int\limits_T \varrho(|f - f_T|)\,dx \le \int\limits_T \varrho(|f - f_T|)\,dx \le c \inf_{f_T} \int\limits_T \varrho(|f - f_T|)\,dx$$

with constants c only depending on $\Delta_2(\varphi)$. Thus our a posteriori estimators improve the one in [LY02].

3.3. **Upper Bound.** To obtain the upper bound we use Lemma 3, the Galerkin orthogonality, and $\Pi_H : V \to V_H$:

$$\begin{aligned}\big\|\mathbf{F}(\nabla u_H) - \mathbf{F}(\nabla u)\big\|_2^2 &\sim \big\langle \mathbf{A}(\nabla u) - \mathbf{A}(\nabla u_H), \nabla(u - u_H)\big\rangle \\ &= \big\langle \mathbf{A}(\nabla u) - \mathbf{A}(\nabla u_H), \nabla(e_H - \Pi_H e_H)\big\rangle.\end{aligned}$$

By integration by parts on each $T \in \mathcal{T}_h$ we get

$$\begin{aligned}&\big\|\mathbf{F}(\nabla u_H) - \mathbf{F}(\nabla u)\big\|_2^2 \\ &\quad\sim -\sum_{T \in \mathcal{T}_H} \int\limits_{\partial T} \big(\mathbf{A}(\nabla u_H)\cdot \mathbf{n}\big)\big(e_H - \Pi_H e_H\big)\,dx + \big\langle f, e_H - \Pi_H e_H\big\rangle \\ &\quad= -\sum_{\gamma \in \Gamma_H} \int\limits_\gamma [\![\mathbf{A}(\nabla u_H)\cdot \mathbf{n}]\!]_\gamma \big(e_H - \Pi_H e_H\big)\,dx + \big\langle f, e_H - \Pi_H e_H\big\rangle \\ &\quad= (\mathtt{Upper}_1) + (\mathtt{Upper}_2),\end{aligned} \tag{3.8}$$

where $\mathbf{n}$ is the outernormal of ∂T. We handle the two terms jump residual $(\mathtt{Upper}_1)$ and element residual $(\mathtt{Upper}_2)$ separately. First we estimate

$$(\mathtt{Upper}_1) \le \sum_{\gamma \in \Gamma_H} \big|[\![\mathbf{A}(\nabla u_H)]\!]_\gamma\big| \int\limits_\gamma \big|e_H - \Pi_H e_H\big|\,dx, \tag{3.9}$$

where we have used that ∇u_H is constant on each $T \in \mathcal{T}_H$. The trace theorem $W^{1,1}(S_\gamma) \hookrightarrow L^1(\gamma)$ and the $W^{1,1}$-approximability of Π_H (see [SZ90] or use (3.6) with $\varphi(t) = t$ and $a = 0$) imply

$$\begin{aligned}\int\limits_\gamma \big|e_H - \Pi_H e_H\big|\,dx &\le C\,h_\gamma \int\limits_{S_\gamma} |e_H - \Pi_H e_H|\,dx + C\int\limits_{S_\gamma} |\nabla e_H - \nabla \Pi_H e_H|\,dx \\ &\le c \sum_{T\in N_\gamma} \int\limits_{\Omega_T} |\nabla e_h|\,dx.\end{aligned}$$

If we combine this with (3.9), then we get

$$(\mathtt{Upper}_1) \le C \sum_{\gamma\in\Gamma_H} \Bigg(\big|[\![\mathbf{A}(\nabla u_H)]\!]_\gamma\big| \sum_{T\in N_\gamma} \int\limits_{\Omega_T} |\nabla e_H|\,dx\Bigg).$$

Now Young's inequality (2.3) for $\varphi_{|\nabla u_H|}$ on each element yields

$$(\mathtt{Upper}_1) \le \sum_{\gamma\in\Gamma_H} \sum_{T\in N_\gamma} \int\limits_{\Omega_T} C_\delta\,(\varphi_{|\nabla u_H|})^*\big(\big|[\![\mathbf{A}(\nabla u_H)]\!]_\gamma\big|\big) + \delta\,\varphi_{|\nabla u_H|}\big(\big|\nabla e_H\big|\big)\,dx$$

Due to (3.4) we have

$$(\varphi_{|\nabla u_H|})^*\big(\big|[\![\mathbf{A}(\nabla u_H)]\!]_\gamma\big|\big) \sim \varphi_{|\nabla u_H|}\big(\big|[\![\nabla u_H]\!]_\gamma\big|\big) \sim \big|[\![\mathbf{F}(\nabla u_H)]\!]_\gamma\big|^2. \tag{3.10}$$

Note that due to shape regularity it holds $h_{T'} \sim h_T \sim h_\gamma$ for $T' \in \omega_T$, $T \in N_\gamma$. With the help of (3.10), Lemma 3, the finite overlapping of the Ω_T, and $\#N_T = 2$ we get

$$\begin{aligned}(\mathtt{Upper}_1) &\le C_\delta\,C \sum_{\gamma\in\Gamma_H} \sum_{T\in N_\gamma} \int\limits_{\Omega_T} \big|[\![\mathbf{F}(\nabla u_H)]\!]_\gamma\big|^2 \\ &\quad + \delta\,C \sum_{\gamma\in\Gamma_H} \sum_{T\in N_\gamma} \int\limits_{\Omega_T} \big|\mathbf{F}(\nabla u) - \mathbf{F}(\nabla u_H)\big|^2\,dx \\ &\le C_\delta\,C \sum_{\gamma\in\Gamma_H} \int\limits_\gamma h_\gamma\,\big|[\![\mathbf{F}(\nabla u_H)]\!]_\gamma\big|^2 \\ &\quad + \delta\,C \sum_{\gamma\in\Gamma_H} \sum_{T\in N_\gamma} \int\limits_{\Omega_T} \big|\mathbf{F}(\nabla u) - \mathbf{F}(\nabla u_H)\big|^2\,dx \\ &\le C_\delta\,C \sum_{\gamma\in\Gamma_H} \eta_J^2(u_H,\gamma) + \delta\,C\,\big\|\mathbf{F}(\nabla u_H) - \mathbf{F}(\nabla u)\big\|_2^2.\end{aligned} \tag{3.11}$$

We treat the element residual as follows

$$\begin{aligned}(\mathtt{Upper}_2) &\le \sum_{T\in\mathcal{T}_H} \int\limits_T |f|\,|e_H - \Pi_H e_H|\,dx \\ &\le \sum_{T\in\mathcal{T}_H} \int\limits_T C_\delta\,(\varphi_{|\nabla u_H|})^*\big(h_T\,|f|\big) + \delta\,(\varphi_{|\nabla u_H|})\bigg(\frac{|e_H - \Pi_H e_H|}{h_T}\bigg)\,dx,\end{aligned}$$

where we have used Young's inequality (2.3). Since ∇u_H is constant on each $T \in \mathcal{T}_H$ we can use (3.6) to obtain

$$(\mathtt{Upper}_2) \le C_\delta\,C \sum_{T\in\mathcal{T}_H} \eta_E^2(u_H,T) + \delta\,C \sum_{T\in\mathcal{T}_H} \int\limits_{\Omega_T} \varphi_{|\nabla u_H(T)|}(|\nabla e_H|)\,dx,$$

where we write $\nabla u_H(T)$ to indicate that the shift on the whole Ω_T depends on the value of ∇u_H on the triangle T. In order to get $\varphi_{|\nabla u_H|}(|\nabla e_H|)$ instead of

$\varphi_{|\nabla u_H(T)|}(|\nabla e_H|)$ we need a *change of shift*. We apply Corollary 26 on each $T' \in \omega_T$ and get

$$(\mathtt{Upper}_2) \leq C_\delta\, C \sum_{T\in\mathcal{T}_H} \eta_E^2(u_H,T) + \delta\, C \sum_{T\in\mathcal{T}_H} \int_{\Omega_T} \varphi_{|\nabla u_H|}(|\nabla e_H|)\,dx$$
$$+ \delta\, C \sum_{T\in\mathcal{T}_H} \sum_{T'\in\omega_T} \int_{T'} \big|\mathbf{F}(\nabla u_H(T)) - \mathbf{F}(\nabla u_H(T'))\big|^2\,dx.$$

Now we transform the last term. Since one can reach T' from T by passing through a finite number of faces (depending on the shape regularity of $\mathcal{T}_H$), we can estimate each $\big|\mathbf{F}(\nabla u_H(T)) - \mathbf{F}(\nabla u_H(T'))\big|$ for $T' \in \omega_T$ by a sum of jumps $\big|[\![\mathbf{F}(\nabla u_H)]\!]_\gamma\big|$ over these faces. In particular,

$$\sum_{T\in\mathcal{T}_H} \sum_{T'\in\omega_T} \int_{T'} \big|\mathbf{F}(\nabla u_H(T)) - \mathbf{F}(\nabla u_H(T'))\big|^2\,dx \sim \sum_{\gamma\in\Gamma_H} \int_{S_\gamma} \big|[\![\mathbf{F}(\nabla u_H)]\!]_\gamma\big|^2\,dx$$
$$\sim \sum_{\gamma\in\Gamma_H} h_\gamma \int_\gamma \big|[\![\mathbf{F}(\nabla u_H)]\!]_\gamma\big|^2\,dx \;=\; \sum_{\gamma\in\Gamma_H} \eta_J^2(u_H,\gamma)$$

using that ∇u_H is piecewise constant and $|S_\gamma| \sim h_\gamma\,|\gamma|$. Hence, with Lemma 3

$$(\mathtt{Upper}_2)$$
$$\leq C_\delta\, C \sum_{T\in\mathcal{T}_H} \eta_E^2(u_H,T) + \delta\, C \sum_{T\in\mathcal{T}_H} \int_{\Omega_T} \varphi_{|\nabla u_H|}(|\nabla e_H|)\,dx + \delta\, C \sum_{\gamma\in\Gamma_H} \eta_J^2(u_H,\gamma)$$
$$\leq C_\delta\, C \sum_{T\in\mathcal{T}_H} \eta_E^2(u_H,T) + \delta\, C\,\big\|\mathbf{F}(\nabla u_H) - \mathbf{F}(\nabla u)\big\|_2^2 + \delta\, C \sum_{\gamma\in\Gamma_H} \eta_J^2(u_H,\gamma).$$

Now, taking $\delta > 0$ small enough we obtain from (3.8), (3.11), the last inequality, (3.7) and the fact that each side $\gamma \in \Gamma_H$ is shared by at most two elements:

Lemma 8 (Upper Bound). *For finite element solutions u_H of (3.5) it holds*

$$\big\|\mathbf{F}(\nabla u_H) - \mathbf{F}(\nabla u)\big\|_2^2 \leq C\,\eta^2(u_H,\mathcal{T}_H), \tag{3.12}$$

where the constant C only depend on $\Delta_2(\varphi)$, $\Delta_2(\varphi^*)$, and the shape regularity of $\mathcal{T}_H$.

3.4. **Lower Bound.** In this section we show that the error can be locally estimated from below by the error estimators. We begin with the element estimator. As is well known, for each $T \in \mathcal{T}_H$ there exists a *bubble function* $w_T \in W_0^{1,\varphi}(T)$ with $w_T \geq 0$ and

$$\int_T w_T\,dx = |T|, \qquad \|w_T\|_\infty \leq C, \qquad \|\nabla w_T\|_\infty \leq \frac{C}{h_T}, \tag{3.13}$$

where $C > 0$ depends only on the shape regularity of $\mathcal{T}_H$. Then for $s \in \mathbb{R}$

$$\big\langle \mathbf{A}(\nabla u) - \mathbf{A}(\nabla u_H), \nabla(s\,w_T)\big\rangle = \langle f, s\,w_T\rangle, \tag{3.14}$$

where we have used that ∇u_H is constant on T. For $f_T \in \mathbb{R}$ by (2.4) applied to $\varphi_{|\nabla u_H|}$ there exists $s_T \in \mathbb{R}$ such that

$$s_T\,(h_T f_T) = (\varphi_{|\nabla u_H(T)|})^*(h_T|f_T|) + \varphi_{|\nabla u_H|}(|s_T|), \tag{3.15}$$

i.e. Young's inequality is sharp. We obtain with (3.15) and (3.14) taking $s = h_T s_T$

$$\begin{aligned}
&|T|\,(\varphi_{|\nabla u_H(T)|})^*(h_T|f_T|) + |T|\,\varphi_{|\nabla u_H(T)|}(|s_T|) = |T|\,f_T\,h_T\,s_T\\
&= \big\langle \mathbf{A}(\nabla u) - \mathbf{A}(\nabla u_H(T)), \nabla(h_T\,s_T\,w_T)\big\rangle + \langle f_T - f, h_T\,s_T\,w_T\rangle\\
&= (\mathtt{Lower}_1) + (\mathtt{Lower}_2).
\end{aligned} \tag{3.16}$$

With the help of (3.13), Young's inequality (2.3), (3.4) and the fact that the integrant is constant $\varphi_{|\nabla u_H|}(|s_T|) = \varphi_{|\nabla u_H(T)|}(|s_T|)$ on T, we have

$$
\begin{aligned}
(\texttt{Lower}_1) &\leq C \int_T |\mathbf{A}(\nabla u) - \mathbf{A}(\nabla u_H)|\, |s_T|\, dx \\
&\leq C_\delta\, C \int_T |\mathbf{F}(\nabla u) - \mathbf{F}(\nabla u_H)|^2\, dx + \delta\, C\, |T|\, \varphi_{|\nabla u_H(T)|}(|s_T|).
\end{aligned}
\tag{3.17}
$$

Similarly, with (3.13) and Young's inequality (2.3) we get

$$
\begin{aligned}
(\texttt{Lower}_2) &\leq C \int_T h_T\, s_T\, |f - f_T|\; dx \\
&\leq C_\delta\, C \int_T (\varphi_{|\nabla u_H|})^*(h_T|f - f_T|)\, dx + \delta\, C|T|\, \varphi_{|\nabla u_H(T)|}(|s_T|).
\end{aligned}
\tag{3.18}
$$

Now, taking $\delta > 0$ small enough we obtain from (3.16), (3.17), and (3.18) that

$$
\begin{aligned}
|T|\, (\varphi_{|\nabla u_H(T)|})^*(h_T|f_T|) \leq C &\int_T |\mathbf{F}(\nabla u) - \mathbf{F}(\nabla u_H)|^2\, dx \\
&+ C \int_T (\varphi_{|\nabla u_H|})^*(h_T|f - f_T|)\, dx.
\end{aligned}
\tag{3.19}
$$

Observe that by convexity of $(\varphi_{|\nabla u_H(T)|})^*$, $\Delta_2(\varphi^*), \Delta_2(\varphi) < \infty$ and Lemma 22

$$
\begin{aligned}
\eta_E^2(u_H, T) &= \int_T (\varphi_{|\nabla u_H|})^*(h_T|f|) \\
&\leq \int_T (\varphi_{|\nabla u_H|})^*(2h_T|f - f_T|) + (\varphi_{|\nabla u_H|})^*(2h_T|f_T|) \\
&\leq C \int_T (\varphi_{|\nabla u_H|})^*(h_T|f - f_T|)\, dx + C\, |T|\, (\varphi_{|\nabla u_H(T)|})^*(h_T|f_T|)
\end{aligned}
$$

with $C > 0$ depending only on $\Delta_2(\varphi)$ and $\Delta_2(\varphi^*)$. This and (3.19) gives

$$
\eta_E^2(u_H, T) \leq C \int_T |\mathbf{F}(\nabla u) - \mathbf{F}(\nabla u_H)|^2\, dx + C \int_T (\varphi_{|\nabla u_H|})^*(h_T|f - f_T|)\, dx.
\tag{3.20}
$$

Taking the infimum over all $f_T \in \mathbb{R}$ proves the following assertion:

Lemma 9. *For finite element solutions u_H of (3.5) and $T \in \mathcal{T}_h$ it holds*

$$
\eta_E^2(u_H, T) \leq C\, \|\mathbf{F}(\nabla u_H) - \mathbf{F}(\nabla u)\|^2_{L^2(T)} + C\, \mathrm{osc}^2(u_H, T).
\tag{3.21}
$$

where the constant C only depends on $\Delta_2(\varphi)$, $\Delta_2(\varphi^*)$, and the shape regularity of $\mathcal{T}_H$.

Now we estimate the jump estimator. As is well known, for each $\gamma \in \Gamma_H$ there exists a *bubble function* $w_\gamma \in W_0^{1,\varphi}(S_\gamma)$ with $w_\gamma \geq 0$ and

$$
\int_\gamma w_\gamma\, dx = |\gamma|, \qquad \|w_\gamma\|_\infty \leq C, \qquad \|\nabla w_\gamma\|_\infty \leq \frac{C}{h_\gamma},
\tag{3.22}
$$

where $C > 0$ depends only on the shape regularity of $\mathcal{T}_H$. Then for $s \in \mathbb{R}$

$$
\begin{aligned}
\big\langle \mathbf{A}(\nabla u) - \mathbf{A}(\nabla u_H), \nabla(s\, w_\gamma) \big\rangle &= \langle f, s\, w_\gamma \rangle - \int_\gamma [\![\mathbf{A}(\nabla u_H) \cdot \mathbf{n}]\!]_\gamma\, s\, w_\gamma\, dx \\
&= \langle f, s\, w_\gamma \rangle - |\gamma|\, [\![\mathbf{A}(\nabla u_H) \cdot \mathbf{n}]\!]_\gamma\, s,
\end{aligned}
\tag{3.23}
$$

where we have used partial integration and that ∇u_H is piecewise constant. Let T_0, T_1 be the two triangles sharing γ. Then by (2.4) applied to $\varphi_{|\nabla u_H(T_0)|}$ there exists $s_\gamma \in \mathbb{R}$ such that

$$\llbracket \mathbf{A}(\nabla u_H)\cdot\mathbf{n}\rrbracket_\gamma\, s_\gamma = (\varphi_{|\nabla u_H(T_0)|})^*\big(\big|\llbracket \mathbf{A}(\nabla u_H)\cdot\mathbf{n}\rrbracket_\gamma\big|\big) + \varphi_{|\nabla u_H(T_0)|}(|s_\gamma|), \tag{3.24}$$

i.e. Young's inequality is sharp. We have chosen $|\nabla u_H(T_0)|$ as the shift, which puts T_0 into a special position, but we will see later that it is not important which of the two triangles is chosen. Let $s = \frac{|S_\gamma|}{|\gamma|} s_\gamma$ in (3.23), then we obtain with (3.24)

$$\begin{aligned}
&|S_\gamma|\,(\varphi_{|\nabla u_H(T_0)|})^*\big(\llbracket \mathbf{A}(\nabla u_H)\cdot\mathbf{n}\rrbracket_\gamma\big) + |S_\gamma|\,\varphi_{|\nabla u_H(T_0)|}(|s_\gamma|)\\
&\quad = |S\gamma|\,\llbracket \mathbf{A}(\nabla u_H)\cdot\mathbf{n}\rrbracket_\gamma\, s_\gamma\\
&\quad = -\frac{|S_\gamma|}{|\gamma|}\,\big\langle \mathbf{A}(\nabla u)-\mathbf{A}(\nabla u_H), \nabla(s_\gamma\, w_\gamma)\big\rangle + \frac{|S_\gamma|}{|\gamma|}\,\langle f, s_\gamma\, w_\gamma\rangle\\
&\quad = (\texttt{Lower}_3) + (\texttt{Lower}_4).
\end{aligned} \tag{3.25}$$

Before we proceed with the estimates for $(\texttt{Lower}_3)$ and $(\texttt{Lower}_4)$ we simplify the term $(\varphi_{|\nabla u_H(T_0)|})^*\big(\big|\llbracket \mathbf{A}(\nabla u_H)\cdot\mathbf{n}\rrbracket_\gamma\big|\big)$. First we show that

$$\big|\llbracket \mathbf{A}(\nabla u_H)\cdot\mathbf{n}\rrbracket_\gamma\big| \sim \varphi'_{|\nabla u_H|}\big(\big|\llbracket \nabla u_H\rrbracket_\gamma\big|\big) \sim \big|\llbracket \mathbf{A}(\nabla u_H)\rrbracket_\gamma\big|. \tag{3.26}$$

The last part of (3.26) is an immediate consequence of Lemma 3. If $\llbracket \nabla u_H\rrbracket_\gamma = 0$, then also $\llbracket \mathbf{A}(\nabla u_H)\cdot\mathbf{n}\rrbracket_\gamma = 0$ and all terms in (3.26) are zero. So we can assume $\llbracket \nabla u_H\rrbracket_\gamma \neq 0$. Since $u_H \in C(S_\gamma)$, the tangential derivatives of u_H are continuous on γ and do not jump. Hence, $|\llbracket \nabla u_H\rrbracket_\gamma| = |\llbracket \nabla u_H\rrbracket_\gamma\cdot\mathbf{n}|$ and

$$\mathbf{n} = \pm\frac{\llbracket \nabla u_H\rrbracket_\gamma}{|\llbracket \nabla u_H\rrbracket_\gamma|}.$$

This and (3.3) imply

$$|\llbracket \mathbf{A}(\nabla u_H)\cdot\mathbf{n}\rrbracket_\gamma|\cdot|\llbracket \nabla u_H\rrbracket_\gamma| = \big|\llbracket \mathbf{A}(\nabla u_H)\rrbracket_\gamma\cdot\llbracket \nabla u_H\rrbracket_\gamma\big| \sim \varphi_{|\nabla u_H|}(|\llbracket \nabla u_H\rrbracket_\gamma|).$$

Now, (2.6) proves (3.26). Lemma 22 states that $(\varphi_{|u_H|})^*$ is an N-function with $\Delta_2((\varphi_{|u_H|})^*) < \infty$ depending solely on $\Delta_2(\varphi)$ and $\Delta_2(\varphi^*)$. Therefor we get with (3.4), and (3.26), and

$$\begin{aligned}
(\varphi_{|\nabla u_H(T_0)|})^*\big(\big|\llbracket \mathbf{A}(\nabla u_H)\cdot\mathbf{n}\rrbracket_\gamma\big|\big) &\sim (\varphi_{|\nabla u_H(T_0)|})^*\big(\big|\llbracket \mathbf{A}(\nabla u_H)\rrbracket_\gamma\big|\big)\\
&\sim \big|\llbracket \mathbf{F}(\nabla u_H)\rrbracket_\gamma\big|^2.
\end{aligned} \tag{3.27}$$

Now, it becomes clear, why the preference of T_0 is not important: The expression $|\llbracket \mathbf{F}(\nabla u_H)\rrbracket_\gamma|^2$ in (3.27) is symmetrical in T_0 and T_1 and therefor independent of the choice T_0.

We proceed with the estimate for $(\texttt{Lower}_3)$. With $|S_\gamma| \sim h_\gamma\,|\gamma|$, (3.22), Young's inequality (2.3), and (3.4) we get

$$\begin{aligned}
(\texttt{Lower}_3) &\leq C \int_{S_\gamma} |\mathbf{A}(\nabla u)-\mathbf{A}(\nabla u_H)|\,|s_\gamma|\,dx\\
&\leq C_\delta\, C \int_{S_\gamma} |\mathbf{F}(\nabla u)-\mathbf{F}(\nabla u_H)|^2\,dx + \delta\, C \sum_{T\in N_\gamma} |T|\,\varphi_{|\nabla u_H(T)|}(|s_\gamma|).
\end{aligned} \tag{3.28}$$

With $|S_\gamma| \sim h_\gamma\,|\gamma|$, (3.22), and Young's inequality (2.3) we deduce

$$\begin{aligned}
&(\texttt{Lower}_4)\\
&\quad\le C \int\limits_{S_\gamma} |f|\,h_\gamma\,|s_\gamma|\,dx\\
&\quad\le C \sum_{T\in N_\gamma} \inf_{f_T\in\mathbb{R}} \int\limits_T |f-f_T|\,h_\gamma\,|s_\gamma| + |f_T|\,h_\gamma\,|s_\gamma|\,dx\\
&\quad\le C_\delta\, C \sum_{T\in N_\gamma} \inf_{f_T\in\mathbb{R}} \int\limits_T (\varphi_{|\nabla u_H|})^*(h_\gamma\,|f-f_T|)\,dx + (\varphi_{|\nabla u_H|})^*(h_\gamma|f_T|)\,dx\\
&\qquad + \delta\, C \sum_{T\in N_\gamma} \int\limits_T \varphi_{|\nabla u_H|}(|s_\gamma|)\,dx\\
&\quad\le C_\delta\, C \sum_{T\in N_\gamma} \operatorname{osc}^2(u_H,T) + C_\delta\, C \sum_{T\in N_\gamma} \eta_E^2(u_H,T) + \delta\, C \sum_{T\in N_\gamma} |T|\varphi_{|\nabla u_H(T)|}(|s_\gamma|).
\end{aligned}$$

This, (3.25), (3.28), and $|T_0|, |T_1| \le |S_\gamma|$ imply

$$\begin{aligned}
&|S_\gamma|\,(\varphi_{|\nabla u_H(T_0)|})^*\big(\big|[\![\mathbf{A}(\nabla u_H)\cdot\mathbf{n}]\!]_\gamma\big|\big) + |S_\gamma|\,\varphi_{|\nabla u_H(T_0)|}(|s_\gamma|)\\
&\quad\le C_\delta\, C \int\limits_{S_\gamma} |\mathbf{F}(\nabla u) - \mathbf{F}(\nabla u_H)|^2\,dx + C_\delta\, C \sum_{T\in N_\gamma} \operatorname{osc}^2(u_H,T)\\
&\qquad + C_\delta\, C \sum_{T\in N_\gamma} \eta_E^2(u_H,T) + \delta\, C \sum_{T\in N_\gamma} |S_\gamma|\varphi_{|\nabla u_H(T)|}(|s_\gamma|).
\end{aligned}$$

For small $\delta > 0$ the summand of the last term with $T = T_0$ could be absorbed on the left hand side, but the term with $T = T_1$ bothers us, since it has the wrong shift $|\nabla u_H(T_1)|$. With Corollary 26 and (3.27) we get rid of this term:

$$\begin{aligned}
\varphi_{|\nabla u_H(T_1)|}(|s_\gamma|) &\le C\,\varphi_{|\nabla u_H(T_0)|}(|s_\gamma|) + C\,\big|[\![\mathbf{F}(\nabla u_H)]\!]_\gamma\big|^2\\
&\le C\,\varphi_{|\nabla u_H(T_0)|}(|s_\gamma|) + C\,(\varphi_{|\nabla u_H(T_0)|})^*\big(\big|[\![\mathbf{A}(\nabla u_H)\cdot\mathbf{n}]\!]_\gamma\big|\big).
\end{aligned}$$

This and (3.28) gives for $\delta > 0$ small

$$\begin{aligned}
&|S_\gamma|\,(\varphi_{|\nabla u_H(T_0)|})^*\big(\big|[\![\mathbf{A}(\nabla u_H)\cdot\mathbf{n}]\!]_\gamma\big|\big)\\
&\quad\le C\,\|\mathbf{F}(\nabla u) - \mathbf{F}(\nabla u_H)\|^2_{L^2(S_\gamma)} + C \sum_{T\in N_\gamma} \operatorname{osc}^2(u_H,T) + C \sum_{T\in N_\gamma} \eta_E^2(u_H,T).
\end{aligned}$$

Now, an application of (3.27) and $|S_\gamma| \sim h_\gamma\,|\gamma|$ prove the following assertion:

Lemma 10. *For finite element solutions u_H of (3.5) and $\gamma \in \Gamma_H$ it holds*

$$\eta_J^2(u_H,\gamma) \le C\,\|\mathbf{F}(\nabla u_H) - \mathbf{F}(\nabla u)\|^2_{L^2(S_\gamma)} + C\operatorname{osc}^2(u_H,N_\gamma) + C\,\eta_E^2(u_H,N_\gamma),$$

where the constant C only depends on $\Delta_2(\varphi)$, $\Delta_2(\varphi^*)$, and the shape regularity of $\mathcal{T}_H$.

Lemma 9 and Lemma 10 can be combined:

Corollary 11. *For finite element solutions u_H of (3.5) and $T \in \mathcal{T}_H$ it holds*

$$\eta^2(u_H,T) \le C\,\|\mathbf{F}(\nabla u_H) - \mathbf{F}(\nabla u)\|^2_{L^2(S_T)} + C\operatorname{osc}^2(u_H,N_T), \tag{3.29}$$

where the constant C only depends on $\Delta_2(\varphi)$, $\Delta_2(\varphi^*)$, and the shape regularity of $\mathcal{T}_H$.

3.5. **Discrete Lower Estimates.** In the following let $\mathcal{T}_h$ be a refinement of $\mathcal{T}_H$, which is generated from $\mathcal{T}_H$ by finitely many bisections. Then $V_h := V(\mathcal{T}_h)$ and $V_H := V(\mathcal{T}_H)$ are nested, i.e. $V_H \subset V_h \subset V$. Let u_h, u_H denote the unique solution of (2.1) with respect to V_h respective V_H.

Our aim is to generalize Corollary 11 from $V_H \subset V$ to $V_H \subset V_h$. Therefore we have to ensure that V_h is a sufficient refinement of V_H. In particular, we have to ensure the existence of bump functions as required in (3.13) and (3.22). We say that $T \in \mathcal{T}_H$ is *fully refined in* $\mathcal{T}_h$ if T and each of its faces contains a node in its interior. This yields the existence of a bump function $w_T \in V_h$ on T which satisfies (3.13) and bump functions $w_\gamma \in V_h$ on S_γ for all $\gamma \in \Gamma_H \cup T$ which satisfy (3.22). Thus, to obtain the local lower bound (3.29) on S_T for a certain $T \in \mathcal{T}_H$ it suffices to assume that each $T' \in N_T$ is fully refined in $\mathcal{T}_h$. With these additional assumptions we can now transfer the estimates from Section 3.4 to $V_H \subset V_h$:

Lemma 12. *If for $T \in \mathcal{T}_H$ each $T' \in N_T$ is fully refined in $\mathcal{T}_h$ then*

$$\eta^2(u_H, T) \le C\,\|\mathbf{F}(\nabla u_H) - \mathbf{F}(\nabla u_h)\|^2_{L^2(S_T)} + C\,\mathrm{osc}^2(u_H, N_T), \tag{3.30}$$

where the constant C only depends on $\Delta_2(\varphi)$, $\Delta_2(\varphi^*)$, and the shape regularity of $\mathcal{T}_H$.

4. Algorithm and Convergence

Let us first state the adaptive algorithm which produces a sequence u_k of solutions in nested spaces $V_k := V(\mathcal{T}_k)$ over triangulations $\mathcal{T}_k$. We substitute the index H, resp. h, of the previous sections by the Index k, resp. $k+1$, to indicate the underlying mesh. Then we introduce the concept of energy reduction and thereafter we prove convergence of the algorithm.

4.1. **Adaptive Algorithm.** We propose the following adaptive algorithm to solve (1.1):

Algorithm 13 (AFEM)**.** Choose $\theta \in (0, 1)$. Construct an initial triangulation $\mathcal{T}_0$ of Ω and set $k := 0$.

(1) (*'Solve'*) Compute the solution $u_k \in V_k$ of Problem (3.5);
(2) (*'Estimate'*) Compute $\eta^2(u_k, T)$ for all $T \in \mathcal{T}_k$.
(3) If $\eta^2(u_k, \mathcal{T}_k) = 0$ then STOP;
(4) (*'Mark'*) Choose a subset $\mathcal{T}_k^m \subset \mathcal{T}_k$ such that

$$\eta^2(u_k, \mathcal{T}_k^m) \ge \theta^2 \eta^2(u_k, \mathcal{T}_k). \tag{4.1}$$

(5) (*'Refine'*) Perform a (minimal) conforming refinement of $\mathcal{T}_k$ using newest vertex bisections to obtain $\mathcal{T}_{k+1}$ such that each element of the neighbourhood $N(\mathcal{T}_k^m) := \bigcup_{T\in\mathcal{T}_k^m} N_T \subset \mathcal{T}_k$ of $\mathcal{T}_k^m$, is fully refined in $\mathcal{T}_{k+1}$. This ensures that each $T \in N(\mathcal{T}_k^m)$ as well as each of its faces contains a node of $\mathcal{T}_{k+1}$ in its interior. Increment k and go to step (1).

Remark 14. *Note that our marking strategy differs from the one proposed by Morin, Nochetto and Siebert in* [MNS00]*. They used separate marking steps for the error estimator and the data oscillation. In our setting this would correspond to the following strategy: Construct $\mathcal{T}_k^m$ as in step* 'Mark'*. Second, enlarge $\mathcal{T}_k^m$ such that for $\hat\theta > 0$ also*

$$\mathrm{osc}^2(u_k, \mathcal{T}_k^m) \ge \hat\theta^2\,\mathrm{osc}^2(u_k, \mathcal{T}_k). \tag{4.2}$$

This requires the calculation of the oscillation in step 'Estimate'*. We want to point out that by the marking strategy (*'Mark'*) our new proof of convergence overcomes the drawback of additional marking for oscillation. This reflects the practical experience that the effect of oscillation plays a minor role (see e.g.* [MNS00]*). We* *prove the success of most adaptive strategies which disregard the issue of oscillation altogether. Since* $\eta^2(u_k, T) \geq \operatorname{osc}^2(u_k, T)$*, this implies that* (4.1) *is equivalent to*

$$\eta^2(u_k, \mathcal{T}_k^m) + \operatorname{osc}^2(u_k, \mathcal{T}_k^m) \geq \overline{\theta}^2 \big(\eta^2(u_k, \mathcal{T}_k) + \operatorname{osc}^2(u_k, \mathcal{T}_k)\big) \tag{4.3}$$

with $\overline{\theta} \in (0, 1)$*. Based on this cognitions we give a new proof to show that the combination of energy difference and oscillation is reduced in each step (see Theorem 20).*

Remark 15. *Note that the condition in* 'Refine' *of fully refined* $T \in \mathcal{T}_k^m$ *can be obtained by bisecting each* $T \in \mathcal{T}_k^m$ *three times in two dimensions respective six times in three dimensions (see* [MNS00]*). With this property we have a reduction factor* $\lambda < 1$ *of element size, i.e. if* $T' \in \mathcal{T}_{k+1}$ *is obtained by refining* $T \in \mathcal{T}_k$ *it holds* $h_{T'} \leq \lambda\, h_T$*. By using the method of* newest vertex bisection *the shape regularity of* $(\mathcal{T}_k)$ *is uniformly bounded with respect to* k *depending on the shape regularity of* $\mathcal{T}_0$*.*

4.2. **Energy Reduction in Nested Spaces.** Assume as before that $V_H \subset V_h \subset V$. One main ingredient of proving lineare convergence in [MNS00] for the linear case is the error reduction property for the energy norm

$$|||u_h - u|||^2 = |||u_H - u|||^2 - |||u_H - u_h|||^2.$$

This is a consequence of the Galerkin orthogonality and the Pythagorean Theorem which is related to Hilbert spaces. We do not have this property in the general case of the φ-Laplacian. But there is another way to interprete this property. In the linear case it is equivalent to

$$\mathcal{J}(u_h) - \mathcal{J}(u) = \mathcal{J}(u_H) - \mathcal{J}(u) - \big(\mathcal{J}(u_H) - \mathcal{J}(u_h)\big). \tag{4.4}$$

Obviously, this equality holds also in our case. Since $V_H \subset V_h \subset V$ and the minimizing property of u, u_h, and u_H we have

$$\mathcal{J}(u) \leq \mathcal{J}(u_h) \leq \mathcal{J}(u_H).$$

Thus we have a reduction of energy difference. Now, it remains to find a link between the energy differences and the error. This is the content of the following Lemma. For $v, w \in V$ we define the energy difference by

$$\varepsilon(v, w) := \mathcal{J}(v) - \mathcal{J}(w).$$

Lemma 16. *Let* u_1, u_2 *be minimizers of the energy functional* $\mathcal{J}$ *with respect to the* $V_1 \subset V_2 \subset V$*. Then*

$$\mathcal{J}(u_1) - \mathcal{J}(u_2) \sim \|\mathbf{F}(\nabla u_1) - \mathbf{F}(\nabla u_2)\|_2^2,$$

where the constants only depend on $\Delta_2(\varphi)$*,* $\Delta_2(\varphi^*)$*, and* (2.7)*.*

Proof. Define $\Phi : \mathbb{R}^d \to \mathbb{R}$ by $\Phi(\mathbf{a}) := \varphi(|\mathbf{a}|)$ then $\mathcal{J}(u) = \int_\Omega \Phi(\nabla u)\, dx - \int_\Omega u\, f\, dx$. Let $g(t) := \mathcal{J}([u_2, u_1]_t)$ for $t \in \mathbb{R}$, where $[u_2, u_1] := (1-t)\, u_2 + t\, u_1$. Since u_2 is the minimizer of $\mathcal{J}$ on $V_2 \supset V_1$,we have $g'(0) = 0$. We estimate by Taylor's formula

(4.5)

$$\mathcal{J}(u_1) - \mathcal{J}(u_2) = g(1) - g(0) = \frac{1}{2} \int_0^1 g''(t)\, (1-t)\, dt$$

$$= \frac{1}{2} \sum_{k,m} \int_0^1 \int_\Omega (\partial_k \partial_m \Phi)([\nabla u_2, \nabla u_1]_t)\, (\partial_k u_1 - \partial_k u_2)(\partial_m u_1 - \partial_m u_2)\, dx\, (1-t)\, dt.$$

Note that for $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$

$$\sum_{k,m} (\partial_k \partial_m \Phi)(\mathbf{a})\, b_k\, b_m = \frac{\varphi'(|\mathbf{a}|)}{|\mathbf{a}|} \left(|\mathbf{b}|^2 - \frac{|\mathbf{a} \cdot \mathbf{b}|}{|\mathbf{a}|^2} \right) + \varphi''(|\mathbf{a}|) \frac{|\mathbf{a} \cdot \mathbf{b}|^2}{|\mathbf{a}|^2}.$$

By Assumption 1 we have $c\,\varphi'(t) \le t\,\varphi''(t) \le C\,\varphi'(t)$ uniformly in $t \ge 0$. Therefor,

$$\sum_{k,m} (\partial_k \partial_m \varPhi)(\mathbf{a})\, b_k\, b_m \le (1+C)\,\frac{\varphi'(|\mathbf{a}|)}{|\mathbf{a}|}|\mathbf{b}|^2$$

and

$$\sum_{k,m} (\partial_k \partial_m \varPhi)(\mathbf{a})\, b_k\, b_m \ge \frac{\varphi'(|\mathbf{a}|)}{|\mathbf{a}|}|\mathbf{b}|^2 + (c-1)\,\frac{\varphi'(|\mathbf{a}|)}{|\mathbf{a}|}\,\frac{|\mathbf{a}\cdot\mathbf{b}|^2}{|\mathbf{a}|^2} \ge c\,\frac{\varphi'(|\mathbf{a}|)}{|\mathbf{a}|}|\mathbf{b}|^2.$$

In other words $\sum_{k,m}(\partial_k \partial_m \varPhi)(\mathbf{a})\, b_k\, b_m \sim \frac{\varphi'(|\mathbf{a}|)}{|\mathbf{a}|}|\mathbf{b}|^2$ uniformly in $\mathbf{a},\mathbf{b} \in \mathbb{R}^d$. Combining the last estimates with (4.5) yields

$$\mathcal{J}(u_1) - \mathcal{J}(u_2) \sim \int_0^1 \int_\Omega \frac{\varphi'(|[\nabla u_2, \nabla u_1]_t|)}{|[\nabla u_2, \nabla u_1]_t|} |\nabla u_1 - \nabla u_2|^2\,dx\,(1-t)\,dt. \tag{4.6}$$

Now, we cite Lemma 19 from [DE05], which states that for any $\mathbf{a},\mathbf{b} \in \mathbb{R}^d$

$$\int_0^1 \frac{\varphi'(|[\mathbf{a},\mathbf{b}]_t|)}{|[\mathbf{a},\mathbf{b}]_t|}\,dt \sim \frac{\varphi'(|\mathbf{a}|+|\mathbf{b}|)}{|\mathbf{a}|+|\mathbf{b}|}$$

with constants only depending on $\Delta_2(\varphi)$ and $\Delta_2(\varphi^*)$. In particular, this, (4.6), Assumption 1, and Lemma 3 gives

$$\begin{aligned}
\mathcal{J}(u_1) - \mathcal{J}(u_2) &\le C \int_\Omega \frac{\varphi'(|\nabla u_2| + |\nabla u_1|)}{|\nabla u_2| + |\nabla u_1|}\,|\nabla u_1 - \nabla u_2|^2\,dx \\
&\le C \int_\Omega \varphi''(|\nabla u_2| + |\nabla u_1|)|\nabla u_1 - \nabla u_2|^2\,dx \\
&\le C \int_\Omega |\mathbf{F}(\nabla u_1) - \mathbf{F}(\nabla u_2)|^2\,dx.
\end{aligned}$$

On the other hand, (4.6), $\varphi'(t)\,t \sim \varphi(t)$ by (2.6), and Jensen's inequality give

$$\begin{aligned}
\mathcal{J}(u_1) - \mathcal{J}(u_2) &\ge c \int_\Omega \int_0^1 \frac{\varphi(|[\nabla u_2, \nabla u_1]_t|)}{(|\nabla u_2| + |\nabla u_1|)^2}\,(1-t)\,dt\,|\nabla u_1 - \nabla u_2|^2\,dx. \\
&\ge c \int_\Omega \frac{\varphi(\int_0^1 |[\nabla u_2, \nabla u_1]_t|\,2\,(1-t)\,dt)}{(|\nabla u_2| + |\nabla u_1|)^2}\,|\nabla u_1 - \nabla u_2|^2\,dx.
\end{aligned}$$

Uniformly in $\mathbf{a},\mathbf{b} \in \mathbb{R}^d$ holds $\int_0^1 |[\mathbf{a},\mathbf{b}]_t|\,2(1-t)\,dt \sim |\mathbf{a}| + |\mathbf{b}|$, because both sides are a norm for the couple $(\mathbf{a},\mathbf{b})$. This and $\varphi''(t)\,t^2 \sim \varphi(t)$ imply

$$\mathcal{J}(u_1) - \mathcal{J}(u_2) \ge c \int_\Omega \varphi''(|\nabla u_2| + |\nabla u_1|)\,|\nabla u_1 - \nabla u_2|^2\,dx.$$

Now, Lemma 3 proves $\mathcal{J}(u_1) - \mathcal{J}(u_2) \ge c\,\int_\Omega |\mathbf{F}(\nabla u_1) - \mathbf{F}(\nabla u_2)|^2\,dx$. □

We recall that by Algorithm 13 for each element $T \in \mathcal{T}_k^m$ it holds $N_T \subset N(\mathcal{T}_k^m)$. Thus each of elements in N_T is marked for full refinement. Thus by Lemma 12 we have a discrete lower bound for each $T \in \mathcal{T}_k^m$. Summing over all sides in $\mathcal{T}_k^m$ yields together with Lemma 16:

Corollary 17. *For the sequence of finite element solutions produced by Algorithm 13 holds*

$$\eta^2(u_k, \mathcal{T}_k^m) \le C\,\varepsilon(u_k, u_{k+1}) + C\,\mathrm{osc}^2(u_k, N(\mathcal{T}_k^m)), \tag{4.7}$$

where the constant C only depends on $\Delta_2(\varphi)$, $\Delta_2(\varphi^*)$, (2.7), and the shape regularity of $\mathcal{T}_0$.

4.3. **Convergence.** To prove that Algorithm 13 produces a sequence (u_k), which converges to the weak solution u of (1.1) we need an auxiliary lemma which deals with oscillation.

Lemma 18. *Then for the sequence of finite element solutions produced by Algorithm 13 there exists $\rho \in (0,1)$ such that*

$$\operatorname{osc}^2(u, \mathcal{T}_{k+1}) \leq \operatorname{osc}^2(u, \mathcal{T}_k) - \rho \operatorname{osc}^2(u, N(\mathcal{T}_k^m)), \tag{4.8}$$

with $\rho = 1 - \lambda$ and λ from Remark 15.

Proof. Recall that for any $T \in \mathcal{T}_k$ and any $T' \in \mathcal{T}_{k+1}$ with $T' \subset T$ we have $h_{T'} \leq h_T$. Moreover, if T is refined in $\mathcal{T}_{k+1}$, then we even have $h_{T'} \leq \lambda\, h_T$. In particular, this holds if $T \in \mathcal{T}_k^m$ since these T are fully refined in $\mathcal{T}_{k+1}$. First for every $T \in \mathcal{T}_k$ we estimate

$$\begin{aligned}
\sum_{T' \in \mathcal{T}_{k+1}, T' \subset T} \operatorname{osc}^2(u, T') &= \sum_{T' \in \mathcal{T}_{k+1}, T' \subset T} \inf_{f_{T'} \in \mathbb{R}} \int_{T'} (\varphi_{|\nabla u|})^*(h_{T'}\, |f - f_{T'}|)\, dx \\
&\leq \sum_{T' \in \mathcal{T}_{k+1}, T' \subset T} \inf_{f_{T'} \in \mathbb{R}} \int_{T'} (\varphi_{|\nabla u|})^*(h_T\, |f - f_{T'}|)\, dx \\
&\leq \inf_{f_T \in \mathbb{R}} \int_T (\varphi_{|\nabla u|})^*(h_T\, |f - f_T|)\, dx \\
&= \operatorname{osc}^2(u, T).
\end{aligned} \tag{4.9}$$

Second, for $T \in N(\mathcal{T}_k^m)$ we have a better estimate, since all elements in $N(\mathcal{T}_k^m)$ are fully refined in $\mathcal{T}_{k+1}$. Therefor, it holds with the convexity of N-functions

$$\begin{aligned}
\sum_{T' \in \mathcal{T}_{k+1}, T' \subset T} \operatorname{osc}^2(u, T') &= \sum_{T' \in \mathcal{T}_{k+1}, T' \subset T} \inf_{f_{T'} \in \mathbb{R}} \int_{T'} (\varphi_{|\nabla u|})^*(h_{T'}\, |f - f_{T'}|)\, dx \\
&\leq \sum_{T' \in \mathcal{T}_{k+1}, T' \subset T} \inf_{f_{T'} \in \mathbb{R}} \int_{T'} (\varphi_{|\nabla u|})^*(\lambda\, h_T\, |f - f_{T'}|)\, dx \\
&\leq \inf_{f_T \in \mathbb{R}} \int_T (\varphi_{|\nabla u|})^*(\lambda\, h_T\, |f - f_T|)\, dx \\
&\leq \lambda \operatorname{osc}^2(u, T).
\end{aligned} \tag{4.10}$$

Now, (4.9) and (4.10) imply

$$\begin{aligned}
\operatorname{osc}^2(u, \mathcal{T}_{k+1}) &\leq \operatorname{osc}^2(u, \mathcal{T}_k \setminus \mathcal{T}_k^m) + \lambda \operatorname{osc}^2(u, \mathcal{T}_k^m) \\
&= \operatorname{osc}^2(u, \mathcal{T}_k) - (1 - \lambda) \operatorname{osc}^2(u, \mathcal{T}_k^m).
\end{aligned}$$

This proves the Lemma. □

Lemma 19. *For the sequence of finite element solutions produced by Algorithm 13 holds*

$$\eta^2(u_k, \mathcal{T}_k) + \operatorname{osc}^2(u_k, \mathcal{T}_k) \sim \varepsilon(u_k, u) + \operatorname{osc}^2(u_k, \mathcal{T}_k) \tag{4.11a}$$

$$\sim \varepsilon(u_k, u) + \operatorname{osc}^2(u, \mathcal{T}_k). \tag{4.11b}$$

Proof. First we prove (4.11a). From Corollary 11, Lemma 16, and Lemma 8 follows

$$\begin{aligned}
\eta^2(u_k, \mathcal{T}_k) &\leq C\, \varepsilon(u_k, u) + C \operatorname{osc}^2(u_k, \mathcal{T}_k), \\
\varepsilon(u_k, u) &\leq C\, \eta^2(u_k, \mathcal{T}_k) + C \operatorname{osc}^2(u_k, \mathcal{T}_k),
\end{aligned}$$

which immediately implies (4.11a). With the help of Corollary 28 we can change the shift in the last term, i.e.

$$
\begin{aligned}
&\operatorname{osc}^2(u_k,\mathcal{T}_k)\\
&\quad= \sum_{T\in\mathcal{T}_k}\inf_{f_T\in\mathbb{R}}\int\limits_T (\varphi_{|\nabla u_k|})^*(h_K\,|f-f_T|)\,dx\\
&\quad\le \sum_{T\in\mathcal{T}_k}\Bigg(\inf_{f_T\in\mathbb{R}} C\int\limits_T (\varphi_{|\nabla u|})^*(h_K\,|f-f_T|)\,dx + C\,|\mathbf{F}(\nabla u_k)-\mathbf{F}(\nabla u)|^2\Bigg)\\
&\quad= C\operatorname{osc}^2(u,\mathcal{T}_k) + C\,\|\mathbf{F}(\nabla u_k)-\mathbf{F}(\nabla u)\|_2^2 \le C\operatorname{osc}^2(u,\mathcal{T}_k)+C\,\varepsilon(u_k,u).
\end{aligned}
\tag{4.12}
$$

The same calculation with u_k and u exchanged proves

$$\operatorname{osc}^2(u,\mathcal{T}_k) \le C\operatorname{osc}^2(u_k,\mathcal{T}_k) + C\,\varepsilon_k(u_k,u). \tag{4.13}$$

Now, (4.12) and (4.13) proves (4.11b). □

Now we are able to prove our main result.

Theorem 20 (Energy/Oscillation Reduction)**.** *There exist a constant $\alpha\in(0,1)$ such that for the sequence (u_k) of finite elements solutions produced by Algorithm AFEM*

$$\varepsilon(u_{k+1},u)+\operatorname{osc}^2(u,\mathcal{T}_{k+1}) \le \alpha^2\,\big(\varepsilon(u_k,u)+\operatorname{osc}^2(u,\mathcal{T}_k)\big). \tag{4.14}$$

Here α only depends on the shape regularity of the sequence $(\mathcal{T}_k)$, $\Delta_2(\varphi)$, $\Delta_2(\varphi^*)$, (2.7), and θ from the marking strategy (4.1). In particular, for $k\in\mathbb{N}$

$$\varepsilon(u_k,u)+\operatorname{osc}^2(u,\mathcal{T}_k) \le \alpha^{2k}\,\big(\varepsilon(u_0,u)+\operatorname{osc}^2(u,\mathcal{T}_0)\big). \tag{4.15}$$

Proof. We start with the energy reduction (4.4)

$$\varepsilon(u_{k+1},u) = \varepsilon(u_k,u) - \varepsilon(u_k,u_{k+1}) \tag{4.16}$$

and note, that by Lemma 18 we have a similar relation for oscillation, i.e.,

$$\operatorname{osc}^2(u,\mathcal{T}_{k+1}) \le \operatorname{osc}^2(u,\mathcal{T}_k) - \rho\operatorname{osc}^2(u,N(\mathcal{T}_k^m)).$$

for some $\rho\in(0,1)$. Summing these two terms yields

$$
\begin{aligned}
&\varepsilon(u_{k+1},u)+\operatorname{osc}^2(u,\mathcal{T}_{k+1})\\
&\qquad\le \varepsilon(u_k,u)+\operatorname{osc}^2(u,\mathcal{T}_k) - \big(\varepsilon(u_k,u_{k+1})+\rho\operatorname{osc}^2\big(u,N(\mathcal{T}_k^m)\big)\big).
\end{aligned}
\tag{4.17}
$$

The main idea is to apply the discrete lower bound to $\varepsilon(u_k,u_{k+1})$, whereas the oscillation on the right hand side of (3.30) shall be absorbed at the oscillation term in the above inequality. Since in (3.30) the oscillation appears with shift u_k, we use Corollary 28 (change of shift) as in (4.12) to obtain

$$\operatorname{osc}^2\big(u_k,N(\mathcal{T}_k^m)\big) \le (1+C_\delta)C\operatorname{osc}^2\big(u,N(\mathcal{T}_k^m)\big) + \delta\,C\,\varepsilon_k(u_k,u)$$

for $\delta>0$ and a corresponding $C_\delta\ge 0$. We set $D=(1+C_\delta)>1$ and obtain equivalently,

$$\frac{1}{C\,D}\operatorname{osc}^2\big(u_k,N(\mathcal{T}_k^m)\big) - \frac{\delta}{D}\,\varepsilon(u_k,u) \le \operatorname{osc}^2\big(u,N(\mathcal{T}_k^m)\big).$$

This, (4.17), and (4.7) imply

$$
\begin{aligned}
&\varepsilon(u_{k+1},u)+\operatorname{osc}^2\big(u,\mathcal{T}_{k+1}\big)\\
&\quad\le \varepsilon(u_k,u)+\operatorname{osc}^2(u,\mathcal{T}_k) - \left(\varepsilon(u_k,u_{k+1}) + \frac{\rho}{CD}\operatorname{osc}^2\big(u_k,N(\mathcal{T}_k^m)\big) - \frac{\rho\,\delta}{D}\,\varepsilon(u_k,u)\right)\\
&\quad\le \left(1+\frac{\rho\,\delta}{D}\right)\varepsilon(u_k,u)+\operatorname{osc}^2(u,\mathcal{T}_k) - \left(\varepsilon(u_k,u_{k+1}) + \frac{\rho}{CD}\operatorname{osc}^2\big(u_k,N(\mathcal{T}_k^m)\big)\right).
\end{aligned}
$$

Since $D \geq 1$ and $\rho \in (0,1)$ we have $\frac{\rho}{2\,C\,D} \leq 1$. Thus

$$\begin{aligned}
&\varepsilon(u_{k+1}, u) + \operatorname{osc}^2(u, \mathcal{T}_{k+1}) \\
&\quad \leq \left(1 + \frac{\rho\,\delta}{D}\right) \varepsilon(u_k, u) + \operatorname{osc}^2(u, \mathcal{T}_k) - \frac{\rho}{2\,CD}\left(\varepsilon(u_k, u_{k+1}) + 2\operatorname{osc}^2\big(u_k, N(\mathcal{T}_k^m)\big)\right)
\end{aligned}$$

the discrete lower estimate (4.7), and $\mathcal{T}_k^m \subset N(\mathcal{T}_k^m)$ yield

$$\begin{aligned}
&\leq \left(1 + \frac{\rho\,\delta}{D}\right) \varepsilon(u_k, u) + \operatorname{osc}^2(u, \mathcal{T}_k) - \frac{\rho}{2\,CD}\left(\frac{1}{C}\,\eta^2(u_k, \mathcal{T}_k^m) + \operatorname{osc}^2\big(u_k, N(\mathcal{T}_k^m)\big)\right) \\
&\leq \left(1 + \frac{\rho\,\delta}{D}\right)\varepsilon(u_k, u) + \operatorname{osc}^2(u, \mathcal{T}_k) - \frac{\rho}{2\,CD}\left(\eta^2(u_k, \mathcal{T}_k^m) + \operatorname{osc}^2(u_k, \mathcal{T}_k^m)\right).
\end{aligned}$$

Now it remains to utilize the marking strategy. By the equivalent form (4.3) of our marking strategy *'Mark'* and (4.11) we get

$$C\left(\varepsilon(u_k, u) + \operatorname{osc}_k^2(u, \mathcal{T}_k)\right) \leq \overline{\theta}^2\left(\eta^2(u_k, \mathcal{T}_k^m) + \operatorname{osc}_k^2(u_k, \mathcal{T}_k^m)\right). \tag{4.18}$$

This observation and Lemma 8 finally imply

$$\begin{aligned}
&\varepsilon(u_{k+1}, u) + \operatorname{osc}^2(u, \mathcal{T}_{k+1}) \\
&\quad \leq \left(1 + \frac{\rho\,\delta}{D}\right) \varepsilon(u_k, u) + \operatorname{osc}^2(u, \mathcal{T}_k) - \frac{\rho\,\overline{\theta}^2}{2\,CD}\left(\eta^2(u_k, \mathcal{T}_k) + \operatorname{osc}^2(u, \mathcal{T}_k)\right) \\
&\quad \leq \left(1 + \frac{\rho\,\delta}{D}\right) \varepsilon(u_k, u) + \operatorname{osc}^2(u, \mathcal{T}_k) - \frac{\rho\,\overline{\theta}^2}{2\,CD}\left(\frac{1}{C}\,\varepsilon(u_k, u) + \operatorname{osc}^2(u, \mathcal{T}_k)\right) \\
&\quad \leq \left(1 + \frac{\rho\,\delta}{D} - \frac{\rho\,\overline{\theta}^2}{2\,CD}\right) \varepsilon(u_k, u) + \left(1 - \frac{\rho\,\overline{\theta}^2}{2\,CD}\right) \operatorname{osc}^2(u, \mathcal{T}_k).
\end{aligned}$$

Substituting D by its definition, we set

$$\alpha^2 := \min\left\{1 + \frac{\rho\,\delta}{(1 + C_\delta)} - \frac{\rho\,\overline{\theta}^2}{2\,C\,(1 + C_\delta)}, 1 - \frac{\rho\,\overline{\theta}^2}{2\,C\,(1 + C_\delta)}\right\}$$

and obtain the proposition for $\delta > 0$ small enough. ☐

From the energy/oscillation reduction of Theorem 20 it follows with the help of Lemma 16:

Corollary 21 (Energy/Oscillation Reduction)**.** *For the sequence (u_k) of finite elements solutions produced by Algorithm AFEM there exists constant $\alpha \in (0,1)$ such that for all $k \in \mathbb{N}$*

$$\begin{aligned}
&\|\mathbf{F}(\nabla u_k) - \mathbf{F}(\nabla u)\|_{L^2(\Omega)}^2 + \operatorname{osc}^2(u, \mathcal{T}_k) \\
&\qquad \leq C\,\alpha^{2k}\left(\|\mathbf{F}(\nabla u_0) - \mathbf{F}(\nabla u)\|_{L^2(\Omega)}^2 + \operatorname{osc}^2(u, \mathcal{T}_0)\right).
\end{aligned} \tag{4.19}$$

Here α and the constants only depend on the shape regularity of $\mathcal{T}_0$, $\Delta_2(\varphi)$, $\Delta_2(\varphi^*)$, (2.7), and θ from the marking strategy (4.1).

5. Appendix

In this section we summarize the properties of the *shifted* N-functions φ_a. Recall that for given N-function φ with $\Delta_2(\varphi), \Delta_2(\varphi^*) < \infty$ we define φ_a as in (3.2) and $\mathbf{F}$ as in (3.1). The following results are from [DE05] and we present them here without proof. We use $C \geq 1$ as generic constant that does solely depend on the Δ_2-constants of φ respective φ^* and may change from line to line.

Lemma 22. *Let φ be an N-function with $\Delta_2(\varphi), \Delta_2(\varphi^*) < \infty$. Then for all $a \geq 0$ the functions φ_a and $(\varphi_a)^*$ are N-function. Moreover, the families φ_a and $(\varphi_a)^*$ satisfy the Δ_2 condition uniformly in $\alpha \geq 0$, i.e. $c_0 := \sup_{a \geq 0}(\Delta_2(\varphi_a), \Delta_2((\varphi_a)^*)) < \infty$. The constant c_0 depends on φ only by $\Delta_2(\varphi)$ and $\Delta_2(\varphi^*)$. Moreover,*

$$(\varphi_a)^*(t) \sim (\varphi^*)_{\varphi'(a)}(t) \tag{5.1}$$

uniformly in $a, t \geq 0$, where the constants only depend on $\Delta_2(\varphi)$ and $\Delta_2(\varphi^*)$.

Lemma 23. *Let φ be a N-function with $\Delta_2(\varphi), \Delta_2(\varphi^*) < \infty$. Then uniformly in $\mathbf{a} \in \mathbb{R}^d$*

$$\begin{aligned} \varphi''(|\mathbf{a}| + |\mathbf{b}|)\, |\mathbf{a} - \mathbf{b}| &\sim \varphi'_{|\mathbf{a}|}(|\mathbf{a} - \mathbf{b}|) \sim \varphi'_{|\mathbf{b}|}(|\mathbf{a} - \mathbf{b}|), \\ \varphi''(|\mathbf{a}| + |\mathbf{b}|)\, |\mathbf{a} - \mathbf{b}|^2 &\sim \varphi_{|\mathbf{a}|}(|\mathbf{a} - \mathbf{b}|) \sim \varphi_{|\mathbf{b}|}(|\mathbf{a} - \mathbf{b}|), \end{aligned} \tag{5.2}$$

with constants only depending on $\Delta_2(\varphi)$ and $\Delta_2(\varphi^*)$.

Lemma 24. *Let φ be as in Assumption 1. Then φ^* satisfies Assumption 1, too. If we define the N-function ψ by*

$$\frac{\psi'(t)}{t} := \left(\frac{\varphi'(t)}{t}\right)^{\frac{1}{2}}$$

then ψ and ψ^* satisfy Assumption 1. Moreover, $\psi''(t) \sim \sqrt{\varphi''(t)}$ uniformly in $t > 0$. The Δ_2-constants of ψ and ψ^* and the constants of $\psi''(t) \sim \sqrt{\varphi''(t)}$ only depend on $\Delta_2(\varphi)$.

The following lemma is proven already in [DR06], but we give a shorter proof:

Lemma 25. *Let φ be a N-function with $\Delta_2(\varphi), \Delta_2(\varphi^*) < \infty$. Then there exists $C \geq 1$, which only depends on $\Delta_2(\varphi)$ such that for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$ and $t \geq 0$*

$$\varphi'_{|\mathbf{a}|}(t) \leq C\, \varphi'_{|\mathbf{b}|}(t) + C\, \varphi'_{|\mathbf{a}|}(|\mathbf{a} - \mathbf{b}|). \tag{5.3}$$

Proof. Since $\varphi'_{|\mathbf{a}|}(t) \sim \varphi_{|\mathbf{a}|}(t)/t$ and $\varphi_{|\mathbf{a}|}(2t) \sim \varphi_{|\mathbf{a}|}(t)$, we have $\varphi'_{|\mathbf{a}|}(2t) \sim \varphi'_{|\mathbf{a}|}(t)$. In particular, $\varphi'_{|\mathbf{a}|}(2t) \leq C\, \varphi'_{|\mathbf{a}|}(t)$ for some $C > 0$ uniformly in $t \geq 0$. All constants depend only on $\Delta_2(\varphi_{|\mathbf{a}|})$, so by Lemma 22 the constants depend only on $\Delta_2(\varphi)$.

Case $|\mathbf{a} - \mathbf{b}| \leq \frac{1}{2}\, t$: From $|\mathbf{a} - \mathbf{b}| \leq \frac{1}{2}\, t$ follows $0 \leq \frac{1}{2}(|\mathbf{b}| + t) \leq |\mathbf{a}| + t \leq 2(|\mathbf{b}| + t)$. Hence,

$$\varphi'_{|\mathbf{a}|}(t) = \frac{\varphi'(|\mathbf{a}| + t)}{|\mathbf{a}| + t}\, t \leq \frac{\varphi'(2\,(|\mathbf{b}| + t))}{\frac{1}{2}(|\mathbf{b}| + t)}\, t \leq 2\, C\, \frac{\varphi'(|\mathbf{b}| + t)}{|\mathbf{b}| + t}\, t = 2\, C\, \varphi'_{|\mathbf{b}|}(t).$$

Case $|\mathbf{a} - \mathbf{b}| \geq \frac{1}{2}\, t$: We estimate

$$\varphi'_{|\mathbf{a}|}(t) \leq \varphi'_{|\mathbf{a}|}(2\, |\mathbf{a} - \mathbf{b}|) \leq C\, \varphi'_{|\mathbf{a}|}(|\mathbf{a} - \mathbf{b}|).$$

Combining the two cases proves the lemma. ∎

Corollary 26 (Change of Shift)**.** *Let φ be an N-function with $\Delta_2(\varphi), \Delta_2(\varphi^*) < \infty$. Then for any $\delta > 0$ there exists $C_\delta > 0$, which only depends on δ and $\Delta_2(\varphi)$ such that for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$ and $t \geq 0$*

$$\varphi_{|\mathbf{a}|}(t) \leq (1 + C_\delta) C\, \varphi_{|\mathbf{b}|}(t) + \delta\, \varphi_{|\mathbf{a}|}(|\mathbf{a} - \mathbf{b}|). \tag{5.4}$$

If φ satisfies Assumption 1, then for any $\delta > 0$ there exists $C_\delta > 0$, which only depends on δ, $\Delta_2(\varphi)$, $\Delta_2(\varphi^*)$, and (2.7) such that for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$ and $t \geq 0$

$$\varphi_{|\mathbf{a}|}(t) \leq (1 + C_\delta) C\, \varphi_{|\mathbf{b}|}(t) + \delta\, C \left|\mathbf{F}(\mathbf{a}) - \mathbf{F}(\mathbf{b})\right|^2. \tag{5.5}$$

Proof. Due to (2.7) holds $\varphi_{|\mathbf{a}|}(t) \sim \varphi'_{|\mathbf{a}|}(t)\,t$. Now inequality (5.4) follows by (5.3), an application of Young's inequality (2.3), (5.2), and (2.5)

$$\begin{aligned}\varphi_{|\mathbf{a}|}(t) &\leq C\,\varphi'_{|\mathbf{a}|}(t)\,t \leq C\,\varphi'_{|\mathbf{b}|}(t)\,t + C\,\varphi'_{|\mathbf{a}|}(|\mathbf{a}-\mathbf{b}|)\,t\\ &\leq C\,\varphi_{|\mathbf{b}|}(t) + C\,\delta\,\varphi^*_{|b|}(\varphi'_{|\mathbf{a}|}(|\mathbf{a}-\mathbf{b}|)) + C\,C_\delta\,\varphi_{|\mathbf{b}|}(t).\end{aligned}$$

On the other hand (5.5) follows from (5.4) with the help of Lemma 3. □

The following lemma is new. It generalizes the *change of shift* to complementary functions.

Lemma 27. *If φ satisfies Assumption 1, then for any $\delta > 0$ there exists $C_\delta > 0$, which only depends on δ, $\Delta_2(\varphi)$, and $\Delta_2(\varphi^*)$, such that for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$ and $t \geq 0$*

$$\big((\varphi_{|\mathbf{a}|})^*\big)'(t) \leq C\,\big((\varphi_{|\mathbf{b}|})^*\big)'(t) + C\,|\mathbf{a}-\mathbf{b}|. \tag{5.6}$$

Proof. With Lemma 22, $|\mathbf{A}(\mathbf{a})| = \varphi'(|\mathbf{a}|)$, Lemma 25, Lemma 3, and $\Delta_2(\varphi) < \infty$ we estimate

$$\begin{aligned}\big((\varphi_{|\mathbf{a}|})^*\big)'(t) &\leq C\,\big((\varphi^*)_{|\mathbf{A}(\mathbf{a})|}\big)'(t)\\ &\leq C\,\big((\varphi^*)_{|\mathbf{A}(\mathbf{a})|}\big)'(t) + C\,\big((\varphi^*)_{|\mathbf{A}(\mathbf{a})|}\big)'(|\mathbf{A}(\mathbf{a}) - \mathbf{A}(\mathbf{b})|)\\ &\leq C\,\big((\varphi_{|\mathbf{a}|})^*\big)'(t) + C\,\big((\varphi_{|\mathbf{a}|})^*\big)'(C\,\varphi'_{|\mathbf{a}|}(|\mathbf{a}-\mathbf{b}|))\\ &\leq C\,\big((\varphi_{|\mathbf{a}|})^*\big)'(t) + C\,\big((\varphi_{|\mathbf{a}|})^*\big)'(\varphi'_{|\mathbf{a}|}(|\mathbf{a}-\mathbf{b}|))\\ &= C\,\big((\varphi_{|\mathbf{a}|})^*\big)'(t) + C\,|\mathbf{a}-\mathbf{b}|.\end{aligned}$$

This proves the Lemma. □

Corollary 28 (Change of Shift $*$)**.** *If φ satisfies Assumption 1, then for any $\delta > 0$ there exists $C_\delta > 0$, which only depends on δ, $\Delta_2(\varphi)$, and $\Delta_2(\varphi^*)$, such that for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$ and $t \geq 0$*

$$(\varphi_{|\mathbf{a}|})^*(t) \leq (1 + C_\delta)C\,(\varphi_{|\mathbf{b}|})^*(t) + \delta\,C\,\varphi_{|\mathbf{a}|}(|\mathbf{a}-\mathbf{b}|). \tag{5.7}$$

If φ satisfies Assumption 1, then for any $\delta > 0$ there exists $C_\delta > 0$, which only depends on δ, $\Delta_2(\varphi)$, $\Delta_2(\varphi^*)$, and (2.7) such that for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$ and $t \geq 0$

$$(\varphi_{|\mathbf{a}|})^*(t) \leq (1 + C_\delta)C\,(\varphi_{|\mathbf{b}|})^*(t) + \delta\,C\,|\mathbf{F}(\mathbf{a}) - \mathbf{F}(\mathbf{b})|^2. \tag{5.8}$$

Proof. Due to (2.7) holds $(\varphi_{|a|})^*(t) \sim ((\varphi_{|a|})^*)'(t)\,t$. Now, inequality (5.7) follows by (5.6), an application of Young's inequality (2.3) (5.2), and (2.5) similarly to the proof of Corollary 26. On the other hand (5.8) follows from (5.7) with the help of Lemma 3. □

References

[BL93a] J. W. Barrett and W. B. Liu, *Finite element approximation of the p-Laplacian*, Math. Comp. **61** (1993), no. 204, 523–537.

[BL94a] J. W. Barrett and W. B. Liu, *Finite element approximation of degenerate quasilinear elliptic and parabolic problems*, Numerical analysis 1993 (Dundee, 1993), Pitman Res. Notes Math. Ser., vol. 303, Longman Sci. Tech., Harlow, 1994, pp. 1–16.

[BL94b] J.W. Barrett and W.B. Liu, *Finite element approximation of the parabolic p-Laplacian*, Siam J. Numer. Anal. **31** (1994), 413–428.

[CLY06] C. Carstensen and W. B. Liu and N. Yan, *A posteriori FE error control for p-Laplacian by gradient recovery in quasi-norm*, Math. Comput. **75** (1994), no. 256, 1599–1616.

[CK06] C. Carstensen and R. Klose, *A posteriori finite element error control for the p-Laplace problem*, SIAM J. Numer. Anal. **25** (2003), no. 3, 792-814.

[DE05] L. Diening and F. Ettwein, *Fractional estimates for non-differentiable elliptic systems with general growth*, Forum Mathematicum, accepted (2006).

[DR06] L. Diening and M. Růžička, *Error estimates for interpolation operators in Orlicz-Sobolev spaces and quasi norms*, Preprint, Mathematische Fakultät, Albert-Ludwigs-Universität Freiburg (2006).

[Dör96] W. Dörfler, *A convergent adaptive algorithm for Poisson's equation*, SIAM J. Numer. Anal. **33** (1996), no. 3, 1106–1124.

[Ebm05] C. Ebmeyer, *Gobal regularity in Sobolev spaces for elliptic problems with p-structure on bounded domains*, Trends in partial differential equations of mathematical physics, Progr. Nonlinear Differential Equations Appl., vol. 61, Birkhäuser, Basel, 2005, pp. 81–89.

[EL05] C. Ebmeyer and WB. 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.

[ELS05] C. Ebmeyer, W. B. Liu, and M. Steinhauer, *Global regularity in fractional order Sobolev spaces for the p-Laplace equation on polyhedral domains*, Z. Anal. Anwendungen **24** (2005), no. 2, 353–374.

[Giu03] E. Giusti, *Direct methods in the calculus of variations*, World Scientific Publishing Co. Inc., River Edge, NJ, 2003.

[GM75] R. Glowinski and A. Marocco, *Sur l'approximation par éléments finis d'ordre un, et la résolution, par pénalisation-dualité, d'une classe de problèmes de Dirichlet non lineaires*, Rev. Française Automat. Informat. Recherche Opérationelle Ser. Rouge Anal. Numér. **2** (1975), 41–76.

[LY01] W. Liu and N. Yan, *Quasi-norm local error estimators for p-Laplacian*, SIAM J. Numer. Anal. **39** (2001), no. 1, 100–127 (electronic).

[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 (electronic).

[MN05] K. Mekchay and R. H. Nochetto, *Convergence of adaptive finite element methods for general second order linear elliptic PDEs*, SIAM J. Numer. Anal. **43** (2005), no. 5, 1803–1827 (electronic).

[MNS00] P. Morin, R. Nochetto, and K. Siebert, *Data oscillation and convergence of adaptive FEM*, Siam J. Numer. Anal. **38** (2000), 466–488.

[RR91] M. M. Rao and Z. D. Ren, *Theory of Orlicz spaces*, Monographs and Textbooks in Pure and Applied Mathematics, vol. 146, Marcel Dekker Inc., New York, 1991.

[SZ90] L. Ridgway Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp. **54** (1990), no. 190, 483–493.

[Ve02] A.Veeser, *Convergent adaptive finite elements for the nonlinear Laplacian*, Numer. Math. **92** (2002), no. 4, 743–770.

Lars Diening, Abteilung für angewandte Mathematik, Eckerstr. 1, 79104 Freiburg i. Br., Germany.

URL: `http://www.mathematik.uni-freiburg.de/IAM/homepages/diening`

E-mail address: `diening@mathematik.uni-freiburg.de`

Christian Kreuzer, , Institut für Mathematik, Universitätsrasse 14, 86159 Augsburg, Germany.

URL: `http://scicomp.math.uni-augsburg.de/~kreuzer/`

E-mail address: `kreuzer@math.uni-augsburg.de`