

## ANALYSIS OF AN ADAPTIVE UZAWA FINITE ELEMENT METHOD FOR THE NONLINEAR STOKES PROBLEM

CHRISTIAN KREUZER

ABSTRACT. We design and study an adaptive algorithm for the numerical solution of the stationary nonlinear Stokes problem. The algorithm can be interpreted as a disturbed steepest descent method, which generalizes Uzawa’s method to the nonlinear case. The outer iteration for the pressure is a descent method with fixed step-size. The inner iteration for the velocity consists of an approximate solution of a nonlinear Laplace equation, which is realized with adaptive linear finite elements. The descent direction is motivated by the quasi-norm which naturally arises as distance between velocities. We establish the convergence of the algorithm within the framework of descent direction methods.

### 1. INTRODUCTION

Partial differential equations like the stationary Stokes problem arise in numerous physical models, particularly in the modeling of Quasi-Newtonian fluids. For  $\Omega \subset \mathbb{R}^d$  being a bounded polyhedral domain and a given external body force  $f : \Omega \rightarrow \mathbb{R}^d$ , the velocity  $u : \Omega \rightarrow \mathbb{R}^d$  of the fluid and its pressure  $p : \Omega \rightarrow \mathbb{R}$  can be described by the stationary Stokes equations:

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

The nonlinear tensor  $\mathbf{A} : \mathbb{R}^{d \times d} \rightarrow \mathbb{R}^{d \times d}$  mimics the change of the viscosity of the fluid with respect to the shear rate  $\mathbf{E}u := \frac{1}{2}(\nabla u + \nabla u^T)$ . Typical choices in engineering are, among others, the *power law* and the *Carreau law*

$$(1.2) \quad \mathbf{A}(\mathbf{E}) = \nu_0 |\mathbf{E}|^{r-2} \mathbf{E} \quad \text{and} \quad \mathbf{A}(\mathbf{E}) = \nu_\infty + (\nu_0 - \nu_\infty)(\kappa^2 + |\mathbf{E}|^2)^{\frac{r-2}{2}} \mathbf{E},$$

$\mathbf{E} \in \mathbb{R}^{d \times d}$ , where  $r \in (1, \infty)$ ,  $\nu_0 > \nu_\infty \geq 0$  and  $\kappa \geq 0$ .

In order to treat as many models as possible, as well as the cases  $1 < r \leq 2$  and  $r > 2$  simultaneously, we formulate (1.1) in terms of so-called N-functions  $\phi$ . Based on these functions we can define function spaces  $\mathbb{V}$  for the velocity and  $\mathbb{Q}$  for the pressure in order to get an appropriate weak formulation of (1.1). The standard finite element approach is to use an adequate pair of discrete function spaces  $\mathbb{V}(\mathcal{T})$ ,  $\mathbb{Q}(\mathcal{T})$  and then compute the Ritz Galerkin approximation  $(U, P) \in \mathbb{V}(\mathcal{T}) \times \mathbb{Q}(\mathcal{T})$ .

---

Received by the editor December 16, 2009 and, in revised form, January 22, 2011.

2010 *Mathematics Subject Classification*. Primary 65N30, 65N12, 35J60.

*Key words and phrases*. Convergence, adaptive finite elements, p-Stokes, p-Laplacian, quasi norm, Uzawa algorithm, nonlinear pde.

©2011 American Mathematical Society  
 Reverts to public domain 28 years from publication

The existence and uniqueness of the weak solution and its discrete approximation is equivalent to the so-called inf-sup conditions on the space pairings  $\mathbb{V} \times \mathbb{Q}$  and  $\mathbb{V}(\mathcal{T}) \times \mathbb{Q}(\mathcal{T})$ , respectively; see also Section 2.1.

To our best knowledge, for the nonlinear Stokes equations with nonlinearities of the kind (1.2) no optimal order reliable and efficient a posteriori estimators can be found in the literature. Thus, there is no strategy on how to mark elements for refinement in order to formulate a standard adaptive finite element method. For nonoptimal estimates see e.g., [BRS04] and the references therein. However, this error estimate may be used to stop an algorithm, since there is no other a posteriori bound available. The Stokes problem leads to a saddle-point problem for its Lagrangian

$$\sup_{q \in \mathbb{Q}} \inf_{v \in \mathbb{V}} \mathcal{L}(v, q) = \mathcal{L}(u, p) = \inf_{v \in \mathbb{V}} \sup_{q \in \mathbb{Q}} \mathcal{L}(v, q);$$

see Section 2.3 i.e., the operator is neither positive definite nor coercive. Most convergence proofs for adaptive finite element methods for elliptic problems make use of the fact that the coercive differential operator defines a suitable error quantity (see e.g. [MNS00, MN05, DK08]). But they cannot be directly transferred to saddle point problems. Very recent results [MSV08, MSV07, Sie09] prove plain convergence for adaptive finite element methods for linear inf-sup problems not relying on coercivity.

Therefore, though the standard finite element method seems to work well in practice, in this article we generalize a different approach, exploiting an idea introduced in [DHU00] in the context of wavelet approximations to the Stokes problem and transferred to finite elements by [BMN02]. In these works they use an inexact Uzawa method, i.e., an Uzawa outer iteration where they substitute the iterative step by a properly wavelet respective finite element approximation. Accounting additionally for mesh refinements according to an error indicator for the pressure, [KS08] modified the inexact Uzawa method and prove optimal computational complexity. As can be observed from [Cia88] Uzawa's method can be interpreted as the method of steepest descent for the functional  $\mathcal{F} : \mathbb{Q} \rightarrow \mathbb{R}$ ,

$$\mathcal{F}(q) := - \inf_{v \in \mathbb{V}} \mathcal{L}(q, v),$$

which attains its minimum at  $p \in \mathbb{Q}$ . As in the linear case  $\mathcal{F}$  is Fréchet differentiable and its derivative in  $q \in \mathbb{Q}$  can be represented by  $\operatorname{div} u_q$ , where  $u_q \in \mathbb{V}$  is the solution to the nonlinear elliptic equation

$$(1.3) \quad -\operatorname{div} \mathbf{A}(\mathbf{E}u_q) = f - \nabla q \quad \text{in } \Omega, \quad u_q = 0 \quad \text{on } \partial\Omega.$$

In order to compute a descent direction we use the efficient adaptive finite element method (AFEM) proposed in [BDK10, Kre08] to numerically approximate (1.3). As usual the adaptive finite element method consists of the loop

$$\text{SOLVE} \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \rightarrow \text{REFINE}.$$

These kinds of methods are powerful and efficient tools for solving linear as well as nonlinear elliptic partial differential equations; see e.g. [MNS00, Vee02, MN05, DK08, CKNS08, MSV08, Ste07, Kre08, BDK10]. Due to the nonlinear nature of (1.3) our AFEM is based on the quasi-norm error concept introduced in [BL93a].

Moreover, we can utilize the quasi-norm techniques for the adaptive Uzawa algorithm (AUA) as well. As a consequence we generalize the steepest descent direction

to a so-called quasi-steepest descent direction. This leads to the following generalization of the Uzawa finite element method from [BMN02]: Starting from an initial guess  $P_0$  of the pressure  $p$ , the AUA consists of a loop

$$(AUA) \quad P_{j+1} := P_j + \mu \mathfrak{D}_j,$$

with a fixed step-size  $\mu > 0$ . The update  $\mathfrak{D}_j$  is an approximation to the quasi-steepest descent direction computed with help of the AFEM.

The main result shows convergence of the AUA with a fixed step-size  $\mu$ . We want to stress that due to the nonlinear nature of the problem, the convergence of (AUA) with a fixed step-size cannot be expected a priori. In particular, the quasi-norm approach is crucial for this result. However, there is no known equivalent to the inf-sup condition for quasi-norm terms, which is crucial for proving linear convergence of Uzawa's method as in [BMN02, NP04].

The work is organized as follows. We start from analytical fundamentals in Section 2, where we introduce basic facts about N-functions and related Orlicz and Orlicz-Sobolev spaces. These spaces are the basis for the weak formulation exhibiting existence and uniqueness of solutions. Finally, we analyze the saddle-point formulation of (1.1) as well as the equivalent minimizing problem.

Section 3 provides the AFEM to calculate a numerical solution of (1.3), which is used to approximate the quasi-steepest descent direction in Section 4. In Section 5 we introduce the AUA and prove its convergence. Finally, in Section 6 we perform numerical experiments to confirm our theoretical predictions.

## 2. THE NONLINEAR STATIONARY STOKES PROBLEM

We first introduce N-functions and related Orlicz and Orlicz-Sobolev spaces as analytical fundamentals for the weak formulation of the nonlinear steady state Stokes equations (1.1) and infer existence and uniqueness of a solution. Subsequently, we introduce a functional  $\mathcal{F}$  based on the saddle-point nature of the Lagrangian associated with (1.1). Finally, we reformulate the weak formulation to the equivalent problem of minimizing  $\mathcal{F}$  and analyze its properties.

**2.1. The nonlinear Stokes problem.** Let  $\Omega \subset \mathbb{R}^d$ ,  $d \in \mathbb{N}$ , be a bounded polyhedral domain. A convenient way of treating most of the common types of equations (1.1) 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  $\phi : \mathbb{R}^+ \mapsto \mathbb{R}^+$ , such that

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

where  $\mathbb{R}^+$  denotes the nonnegative real semi-axis. For more detailed information on N-functions and related Orlicz and Orlicz-Sobolev spaces consider e.g., [RR91, KK91, Mus83, KR61, 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  $\int_0^t \phi'(s) ds = \phi(t)$ ; see [RR91]. However, we require more regularity: Throughout the paper we fix an N-function  $\phi$  and assume that

$$(2.1) \quad \phi \in C^1([0, \infty)) \cap C^2((0, \infty)) \quad \text{and} \quad ct\phi''(t) \leq \phi'(t) \leq Ct\phi''(t), \quad t \geq 0,$$

for some constants  $C, c > 0$ . Denoting the Frobenius norm by  $|\mathbf{Q}|^2 := \sum_{i,j=1}^d Q_{ij}^2$ ,  $\mathbf{Q} = (Q_{ij}) \in \mathbb{R}^{d \times d}$  we define the nonlinear vector field

$$(2.2) \quad \mathbf{A}(\mathbf{Q}) := \phi'(|\mathbf{Q}|) \frac{\mathbf{Q}}{|\mathbf{Q}|} \quad \text{for all } \mathbf{Q} \in \mathbb{R}^{d \times d}.$$

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

*Remark 1.* For  $r \in (1, \infty)$  the functions  $\phi_1(t) := \frac{\nu_0}{r} |t|^r$  and  $\phi_2(t) := \int_0^t (\nu_\infty + (\nu_0 - \nu_\infty)(\kappa^2 + s^2)^{(r-2)/2}) s ds$  satisfy assumption (2.1). Using definition (2.2) of  $\mathbf{A}$ , we observe that the concept of N-functions covers among others the power law as well as the Carreau law; compare also with [Kre08]. Exploiting the abstract structure of N-functions enables us to treat many different viscosity laws simultaneously.

An N-function  $\phi$  is said to satisfy the  $\Delta_2$ -condition if there exists a constant  $K$  such that  $\phi(2t) \leq K\phi(t)$ . The smallest constant is then denoted by  $\Delta_2(\phi)$  and for a family of N-functions  $(\phi_a)_{a \in I}$  we define  $\Delta(\{(\phi_a)_{a \in I}\}) := \max_{a \in I} \Delta_2(\phi_a)$ . Assumption (2.1) yields  $(\ln \phi'(t))' \leq \frac{1}{ct}$  and thus

$$(2.3) \quad \ln \phi'(2t) - \ln \phi'(t) \leq \int_t^{2t} \frac{1}{cs} ds = \frac{\ln 2}{c} \quad \Rightarrow \quad \phi'(2t) \leq 2^{1/c} \phi'(t).$$

By the fundamental theorem of calculus and the transformation theorem this implies that an N-function  $\phi$  satisfies the  $\Delta_2$ -condition if (2.1) holds.

For  $\omega \subset \mathbb{R}^d$  a measurable subset, we denote the classical Orlicz and Orlicz-Sobolev spaces by  $L^\phi(\omega)$  and  $W^{1,\phi}(\omega)$  i.e.,  $f \in L^\phi(\omega)$  if  $\int_\omega \phi(|f|) dx < \infty$  and  $f \in W^{1,\phi}(\omega)$  if  $f, \nabla f \in L^\phi(\omega)$ . Equipped with the Luxemburg norm  $\|f\|_{(\phi),\omega} := \inf \{ \lambda > 0 : \int_\omega \phi(|f|/\lambda) dx \leq 1 \}$  the space  $L^\phi(\omega)$  becomes a Banach space, and  $W^{1,\phi}(\omega)$  becomes a 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  $W^{1,\phi}(\omega)$  and  $L_0^\phi(\omega) \subset L^\phi(\omega)$  is the subspace of functions with mean-value zero. For  $\omega = \Omega$  we skip the domain in the notion of the norm, e.g.,  $\|\cdot\|_{(\phi)} = \|\cdot\|_{(\phi),\Omega}$ . If vector-valued functions are considered we denote the dimension of the function-values as superscript e.g.,  $W_0^{1,\phi}(\Omega)^d$  and  $L^\phi(\Omega)^d$ . Whereas we denote the norms like the norms for scalar-valued functions, since it is always clear from the context which norm is considered.

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

$$\phi^*(t) := \max\{st - \phi(s) : s \geq 0\}.$$

It holds that  $\phi^*$  itself is an N-function and  $(\phi^*)^* = \phi$ . If  $\phi$  satisfies (2.1), then  $\phi^*$  does as well; see [DE08, Kre08]. Thanks to the  $\Delta_2$ -condition the dual  $(L^\phi(\Omega))^*$  is isomorphic to  $L^{\phi^*}(\Omega) \subset (W^{1,\phi}(\Omega))^*$  and by  $(\phi^*)^* = \phi$  we have that  $(L^\phi(\Omega))^*$  is reflexive. The space  $W_0^{1,\phi}(\Omega)$  is also reflexive and its dual space is denoted by  $W^{-1,\phi^*}(\Omega)$ . The dual space of  $L_0^\phi(\Omega)$  is  $L^{\phi^*}(\Omega)/\mathbb{R}$  with norm  $\|\cdot\|_{L^{\phi^*}/\mathbb{R}} := \inf_{c \in \mathbb{R}} \|\cdot - c\|_{(\phi^*)}$ . By  $\langle \cdot, \cdot \rangle$  we denote dual pairings regardless of the space-pairings that are considered. For the ease of exposition, in the remainder of this article, we will often use the abbreviations

$$\mathbb{V} := W_0^{1,\phi}(\Omega)^d \quad \text{and} \quad \mathbb{Q} := L^{\phi^*}(\Omega)/\mathbb{R},$$

with corresponding norms  $\|\cdot\|_{\mathbb{V}} := \|\cdot\|_{1,(\phi)}$  and  $\|\cdot\|_{\mathbb{Q}} := \|\cdot\|_{L^{\phi^*}(\Omega)/\mathbb{R}}$ .

Throughout this paper we assume that  $f \in L^{\phi^*}(\Omega)^d$ . Then the weak formulation of the nonlinear Stokes problem (1.1) reads as follows: Find  $u \in \mathbb{V}$ ,  $p \in \mathbb{Q}$ , such that

$$(2.4) \quad \int_{\Omega} \mathbf{A}(\mathbf{E}u) : \mathbf{E}v \, dx - \int_{\Omega} p \operatorname{div} v \, dx + \int_{\Omega} q \operatorname{div} u \, dx = \int_{\Omega} f \cdot v \, dx$$

for all  $(v, q) \in \mathbb{V} \times \mathbb{Q}$ , where  $\mathbf{Q} : \mathbf{P} := \sum_{i,j=1}^d Q_{ij} P_{ij}$ ,  $\mathbf{P} = (P_{ij})_{i,j=1}^d$ ,  $\mathbf{Q} = (Q_{ij})_{i,j=1}^d \in \mathbb{R}^{d \times d}$ . Thanks to the embedding  $L^{\phi^*}(\Omega)^d \subset \mathbb{V}^* := W^{-1, \phi^*}(\Omega)$  we can rewrite (2.4) as the operator equation

$$(2.5) \quad \begin{aligned} -\operatorname{div} \mathbf{A}(\mathbf{E}u) + \nabla p &= f && \text{in } \mathbb{V}^*, \\ \operatorname{div} u &= 0 && \text{in } L_0^{\phi}(\Omega), \end{aligned}$$

where  $\langle -\operatorname{div} \mathbf{A}(\mathbf{E}u), v \rangle := \int_{\Omega} \mathbf{A}(\mathbf{E}u) : \mathbf{E}v \, dx$  and  $\langle \nabla p, v \rangle := -\int_{\Omega} p \operatorname{div} v \, dx$ .

The existence and uniqueness of the function  $u \in W_0^{1, \phi}(\Omega)^d$  is ensured by the theory of monotone operators. In particular, by [DE08, Lemma 20] it holds for all  $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^{d \times d}$  that

$$(2.6) \quad \begin{aligned} (\mathbf{A}(\mathbf{P}) - \mathbf{A}(\mathbf{Q})) : (\mathbf{P} - \mathbf{Q}) &\geq \phi''(|\mathbf{P}| + |\mathbf{Q}|) |\mathbf{P} - \mathbf{Q}|^2, \\ |\mathbf{A}(\mathbf{P}) - \mathbf{A}(\mathbf{Q})| &\leq \phi''(|\mathbf{P}| + |\mathbf{Q}|) |\mathbf{P} - \mathbf{Q}|. \end{aligned}$$

This together with Korn's inequality,

$$(2.7) \quad \|v\|_{1,(\phi)} \leq \|\mathbf{E}v\|_{(\phi)} \quad \text{for all } v \in \mathbb{V}$$

(see [DRS10] and Remark 2), implies that  $-\operatorname{div} \mathbf{A}(\mathbf{E}\cdot) : \mathbb{V} \rightarrow \mathbb{V}^*$  is a coercive, monotone and continuous operator; see also [Kre08]. Hence, the theory of monotone operators shows that this operator is bijective. Since the space  $\{v \in \mathbb{V} : \operatorname{div} v = 0\}$  is closed in  $\mathbb{V}$  there exists a unique  $u \in \mathbb{V}$ ,  $\operatorname{div} u \equiv 0$ , such that

$$\int_{\Omega} \mathbf{A}(\nabla u) : \nabla v \, dx = \int_{\Omega} f \cdot v \, dx \quad \text{for all } v \in \mathbb{V} \text{ with } \operatorname{div} v \equiv 0.$$

The unique existence of a  $p \in \mathbb{Q}$  such that (2.4) holds is then guaranteed by the so-called inf-sup condition

$$(2.8) \quad \inf_{q \in \mathbb{Q}} \sup_{v \in \mathbb{V}} \frac{\int_{\Omega} q \operatorname{div} v \, dx}{\|v\|_{\mathbb{V}} \|q\|_{\mathbb{Q}}} > \beta.$$

Since  $L_0^{\phi^*}(\Omega)$  is isomorphic to  $L^{\phi^*}(\Omega)/\mathbb{R}$  (see [Kre08]), the inf-sup condition is equivalent to the solvability of the divergence equation i.e., for each  $q \in L_0^{\phi^*}(\Omega)$ ,  $\int_{\Omega} q \, dx = 0$ , there exists  $v \in W_0^{1, \phi^*}(\Omega)$  such that  $\operatorname{div} v = q$  and  $\|v\|_{1,(\phi^*)} \leq 1/\beta \|q\|_{(\phi^*)}$ . This estimate is proved in [DRS10]. In the common cases of the power law and the Carreau law we have  $\|\cdot\|_{\mathbb{Q}} = \|\cdot\|_{L^{r'}(\Omega)/\mathbb{R}}$  and  $\|\cdot\|_{\mathbb{V}} = \|\cdot\|_{W^{1, r'}(\Omega)}$  and hence the inf-sup condition can also be found in [AG94].

*Remark 2.* The Korn inequality (2.7) as well as the inf-sup condition (2.8) in [DRS10] are only proved for John domains. A domain  $\omega \subset \mathbb{R}^d$  is called a John domain, if there exist some constant  $\alpha > 0$  such that any distinct points  $x, y \in \omega$

can be joint by a rectifiable path  $\gamma$  parametrized by its arc-length  $|\gamma|$ , such that

$$\text{cig}(\gamma, \alpha) := \bigcup_{t \in [0, |\gamma|]} \left\{ B\left(\gamma(t), \frac{1}{\alpha} \min\{t, |\gamma| - t\}\right) \right\} \subset \omega;$$

see [DRS10]. Thereby  $B(z, r)$  is the ball with center  $z$  and radius  $r$ . Hence, the constant  $1/\alpha$  determines the ‘angle’ of the cigar  $\text{cig}(\gamma, \alpha)$  at its end-points  $x$  and  $y$ .

It is clear that this condition can be satisfied for all pairs of points interior to the polyhedral domain  $\Omega$ . Moreover, since boundary angles of polyhedral domains are bounded it can be easily verified that  $\Omega$  is a John domain.

**2.2. Properties of N-functions.** In order to be able to continue analyzing (1.1) we have to introduce some more properties of N-functions. For proofs of the results presented in this section consider e.g., [RR91, KK91, Mus83, KR61] or the detailed overview in [Kre08].

**Proposition 3.** *Let  $\phi, \psi$  be N-functions. Then for all  $t \geq 0$ ,*

$$(2.9a) \quad \phi(\alpha t) \leq \alpha \phi(t) \quad \text{for all } \alpha \in [0, 1],$$

$$(2.9b) \quad \frac{t}{2} \phi'\left(\frac{t}{2}\right) \leq \phi(t) \leq t \phi'(t),$$

$$(2.9c) \quad t \leq (\phi^*)^{-1}(t) \phi^{-1}(t) \leq 2t,$$

$$(2.9d) \quad \phi\left(\frac{\phi^*(t)}{t}\right) \leq \phi^*(t) \leq \phi\left(2 \frac{\phi^*(t)}{t}\right),$$

$$(2.9e) \quad \phi(t) \leq \psi(t) \Rightarrow \psi^*(t) \leq \phi^*(t).$$

We recall that (2.1) implies  $\Delta_2(\{\phi, \phi^*\}) < \infty$  and hence  $\phi(t) \approx \phi(2t)$ . Therefore, (2.9b) and (2.9d) imply

$$(2.10) \quad t \phi'(t) \approx \phi(t) \quad \text{and} \quad \phi((\phi^*)'(t)) \approx \phi^*(t) \quad \text{for } t \geq 0.$$

For each constant  $C \leq 2^k$ ,  $k \in \mathbb{N}$  we have

$$(2.11) \quad \phi(Ct) \leq \Delta_2(\phi)^k \phi(t) \quad \text{for all } t \geq 0,$$

as well as by the convexity of  $\phi$ ,

$$(2.12) \quad \phi(s+t) \leq \frac{1}{2} \phi(2t) + \frac{1}{2} \phi(2s) \leq \frac{\Delta_2(\phi)}{2} (\phi(s) + \phi(t)) \quad \text{for all } s, t \geq 0.$$

Moreover,  $\Delta_2(\phi) < \infty$  is equivalent to the existence of a constant  $\nabla_2(\phi^*) > 1$  such that

$$(2.13) \quad \nabla_2(\phi^*) \phi^*(t) \leq \phi'(t) t \quad \text{for all } t \geq 0.$$

An immediate consequence of the definition of the complementary function is the Young inequality

$$(2.14) \quad ts \leq \phi(t) + \phi^*(s) \quad \text{for all } s, t \geq 0,$$

where equality is obtained if  $s = \phi'(t)$  or  $t = (\phi^*)'(s)$ . Moreover, by (2.11) it holds a scaled Young inequality 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

$$(2.15) \quad ts \leq C_\delta \phi(t) + \delta \phi^*(s) \quad \text{for all } s, t \geq 0.$$

Another important consequence of the  $\Delta_2$ -condition is that norm convergence is equivalent to mean convergence i.e., let  $(v_k)_{k \in \mathbb{N}} \subset L^\phi(\Omega)$ , then for some  $v \in L^\phi(\Omega)$ , we have

$$(2.16) \quad \lim_{k \rightarrow \infty} \|v_k - v\|_{(\phi)} = 0 \quad \text{if and only if} \quad \lim_{k \rightarrow \infty} \int_{\Omega} \phi(|v_k - v|) dx = 0.$$

Given a fixed N-function  $\phi$  that satisfies (2.1) we introduce the family of shifted N-functions  $\{\phi_a\}_{a \geq 0}$  by

$$\phi'_a(t) := \frac{\phi'(a+t)}{a+t} t \quad \text{and} \quad \phi_a(t) := \int_0^t \phi'_a(s) ds, \quad a \geq 0.$$

The following properties of this family of N-functions are crucial in the subsequent analysis and can e.g., be found in [DE08, DK08, Kre08].

**Lemma 4.** *Let  $\phi$  be an N-function such that (2.1). Then  $\Delta_2(\{\phi_a, (\phi_a)^*\}_{a \geq 0})$  is uniformly bounded depending solely on  $\Delta_2(\{\phi, \phi^*\})$  and thus on the constants in (2.1). Moreover, for all  $t \geq 0$  we have*

$$(2.17) \quad (\phi^*)_{\phi'(a)}(t) \approx (\phi_a)^*(t).$$

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

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

The relation of  $\mathbf{A}$ ,  $\mathbf{F}$ , and  $\{\phi_a\}_{a \geq 0}$  is best reflected in the following lemma from [DE08].

**Proposition 5.** *Let  $\phi$  be an N-function that satisfies assumption (2.1). Then, for all  $\mathbf{Q}, \mathbf{P} \in \mathbb{R}^{d \times d}$  it holds that*

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

and

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

This directly implies the next result.

**Corollary 6.** *Let  $\phi$  be an N-function that satisfies assumption (2.1). Then, for all  $v, w \in \mathbb{V}$  it holds that*

$$\begin{aligned} \int_{\Omega} (\mathbf{A}(\mathbf{E}v) - \mathbf{A}(\mathbf{E}w)) : (\mathbf{E}v - \mathbf{E}w) dx &\approx \int_{\Omega} \phi_{|\mathbf{E}v|}(|\mathbf{E}v - \mathbf{E}w|) dx \\ &\approx \|\mathbf{F}(\mathbf{E}v) - \mathbf{F}(\mathbf{E}w)\|_{L^2(\Omega)}^2. \end{aligned}$$

The three equivalent quantities of Corollary 6 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 11.

The next lemmas deal with the question of how the shifted N-functions of the family  $\{\phi_a\}_{a \geq 0}$  are related to each other; see [DE08, DK08].

**Lemma 7.** *Let  $\phi$  be an  $N$ -function with (2.1), then for all  $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^{d \times d}$  and  $t \geq 0$  it holds that*

$$(2.18) \quad \phi'_{|\mathbf{P}|}(t) \preceq \phi'_{|\mathbf{Q}|}(t) + \phi'_{|\mathbf{P}|}(|\mathbf{P} - \mathbf{Q}|).$$

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

$$(2.19) \quad \phi_{|\mathbf{P}|}(t) \preceq (1 + C_\delta) \phi_{|\mathbf{Q}|}(t) + \delta \phi_{|\mathbf{P}|}(|\mathbf{P} - \mathbf{Q}|)$$

and

$$(2.20) \quad (\phi_{|\mathbf{P}|})^*(t) \approx (1 + C_\delta) (\phi_{|\mathbf{Q}|})^*(t) + \delta \phi_{|\mathbf{P}|}(|\mathbf{P} - \mathbf{Q}|).$$

**Lemma 8.** *Let  $\phi$  be an  $N$ -function that satisfies assumption (2.1). We then have for all  $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^{d \times d}$*

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

As a consequence of these results we get the following characterization of convergence in  $L^\phi(\Omega)$ .

**Proposition 9.** *Let  $\phi$  be an  $N$ -function such that (2.1),  $(v_k)_{k \in \mathbb{N}} \subset L^\phi(\Omega)$ . Then for some  $v, w \in L^\phi(\Omega)$  we have*

$$\lim_{k \rightarrow \infty} \int_{\Omega} \phi_{|w|}(|v_k - v|) dx = 0 \quad \text{if and only if} \quad \lim_{k \rightarrow \infty} v_k = v \quad \text{in } L^\phi(\Omega).$$

*Proof.* Assume that  $v_k \rightarrow_{k \rightarrow \infty} v$  in  $L^\phi(\Omega)$ . Then by (2.19) we have for  $\delta > 0$ ,

$$\int_{\Omega} \phi_{|w|}(|v_k - v|) dx \preceq (1 + C_\delta) \int_{\Omega} \phi(|v_k - v|) dx + \delta \int_{\Omega} \phi_{|w|}(|w|) dx.$$

The second addend is small provided  $\delta$  is small. By (2.16) the first addend converges to zero as  $k$  converges to infinity. The converse implication can be proved with the same arguments interchanging the roles of  $\phi_{|w|}$  and  $\phi$ .  $\square$

As a consequence the space  $W_0^{1,\phi}(\Omega)$  is closed with respect to the quasi-norm.

**Corollary 10.** *Let  $\phi$  be an  $N$ -function such that (2.1) holds. Then*

$$\mathbf{d}(v, w) := \|\mathbf{F}(\mathbf{E}v) - \mathbf{F}(\mathbf{E}w)\|_{L^2(\Omega)}$$

*is a metric on  $W_0^{1,\phi}(\Omega)$ .*

*Proof.* Obviously,  $\mathbf{d}$  is positive and satisfies the triangle quality. Moreover, as a consequence of Proposition 9 (take  $w = v$ ) and Korn's inequality [DRS10] we have for any sequence  $(v_k)_{k \in \mathbb{N}} \subset W_0^{1,\phi}(\Omega)$ ,  $v \in W_0^{1,\phi}(\Omega)$  that

$$\lim_{k \rightarrow \infty} \|\mathbf{F}(\mathbf{E}v_k) - \mathbf{F}(\mathbf{E}v)\|_{L^2(\Omega)} = 0 \quad \Leftrightarrow \quad \lim_{k \rightarrow \infty} \|v_k - v\|_{1,(\phi)} = 0.$$

This proves the assertion.  $\square$

*Remark 11.* A distance measure named quasi-norm was first introduced by Barrett and Liu in [BL93a, BL93b, BL94] 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  $\phi(t) = \frac{1}{r}t^r$  with  $r \in (1, \infty)$ , they use

$$\int_{\Omega} (|\mathbf{E}u| + |\mathbf{E}u - \mathbf{E}v|)^{r-2} |\mathbf{E}u - \mathbf{E}v|^2 dx$$

which in turn is equivalent to

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

Hence our approach generalizes this concept to N-functions satisfying assumption (2.1); compare with [DR07a, DE08, DR07b, DK08, BDK10, Kre08].

**2.3. The Lagrangian functional.** The stationary Stokes equation (1.1) with the pressure as Lagrange multiplier arise from minimizing the energy functional

$$(2.21) \quad \mathcal{J}(v) := \int_{\Omega} \phi(|\mathbf{E}v|) - f \cdot v dx \rightarrow \min \quad \text{subject to} \quad \operatorname{div} v = 0.$$

For a given N-function  $\phi$ , we define the Lagrangian  $\mathcal{L} : \mathbb{V} \times \mathbb{Q} \rightarrow \mathbb{R}$  associated with (2.21) by

$$\mathcal{L}(v, q) := \int_{\Omega} \phi(|\mathbf{E}v|) - q \operatorname{div} v - f \cdot v dx.$$

Basic calculations show that the Lagrangian has a unique saddle-point that corresponds to the solution  $(u, p)$  of (2.4); see also [ET76, Kre08].

**Proposition 12.** *Let  $\phi$  be an N-function that satisfies assumption (2.1). Then the nonlinear Stokes problem (2.4) is equivalent to the saddle-point problem: Find functions  $u \in \mathbb{V}$ ,  $p \in \mathbb{Q}$ , such that*

$$(2.22) \quad \sup_{q \in \mathbb{Q}} \inf_{v \in \mathbb{V}} \mathcal{L}(v, q) = \mathcal{L}(u, p) = \inf_{v \in \mathbb{V}} \sup_{q \in \mathbb{Q}} \mathcal{L}(v, q)$$

*i.e., the unique solution  $(u, p) \in \mathbb{V} \times \mathbb{Q}$  of (2.4) is the unique saddle-point of  $\mathcal{L}$ .*

Based on the above observations we define the nonlinear functional

$$(2.23) \quad \mathcal{F}(q) := - \inf_{v \in \mathbb{V}} \mathcal{L}(v, q) \quad \text{for all } q \in \mathbb{Q}.$$

According to Proposition 12 our aim is to minimize  $\mathcal{F}$ . Note from the definition of the Lagrangian function, that evaluating  $\mathcal{F}$  at  $q \in \mathbb{Q}$  is a minimizing problem of the form

$$(2.24) \quad \mathcal{J}_q(v) := \int_{\Omega} \phi(|\mathbf{E}v|) dx - \langle f - \nabla q, v \rangle \rightarrow_{v \in \mathbb{V}} \min$$

with  $f - \nabla q \in W^{-1, \phi^*}(\Omega)$ . Direct methods of variations show, by coercivity and the strict convexity of  $\mathcal{J}_q(\cdot)$ , that there exists a unique minimizer  $u_q \in \mathbb{V}$ . Moreover,  $u_q \in \mathbb{V}$  is the critical point of  $D\mathcal{J}_q(v) = -\operatorname{div} \mathbf{A}(\mathbf{E}v) - f + \nabla q$ , i.e.,

$$(2.25) \quad -\operatorname{div} \mathbf{A}(\mathbf{E}u_q) = f - \nabla q \quad \text{in } W^{-1, \phi^*}(\Omega);$$

see [DE08]. Therefore, we have

$$(2.26) \quad \mathcal{F}(q) = -\mathcal{L}(u_q, q) = - \inf_{v \in \mathbb{V}} \mathcal{L}(v, q).$$

The following result from [DK08] shows that energy differences are closely connected to the error concept of the quasi-norm.

**Lemma 13.** *Let  $u_q \in \mathbb{V}$ ,  $q \in \mathbb{Q}$  be the minimizer of the energy functional  $\mathcal{J}_q$  defined in (2.24). Then*

$$\mathcal{J}_q(v) - \mathcal{J}_q(u_q) \approx \|\mathbf{F}(\mathbf{E}v) - \mathbf{F}(\mathbf{E}u_q)\|_2^2, \quad \text{for all } v \in \mathbb{V}.$$

*Proof.* The proof can be found in [DK08]. However, in this paper the result is formulated for  $v$  being a minimizer of  $\mathcal{J}_q$  in a closed subspace  $\tilde{\mathbb{V}} \subset \mathbb{V}$ . Since we need the result for arbitrary  $v \in \mathbb{V}$  we decided to sketch the proof in order to point out that the additional condition on  $v$  is not necessary. With the definition  $\Phi(\mathbf{Q}) := \phi(|\mathbf{Q}|)$  we have  $\mathcal{J}_q(v) = \int_{\Omega} \Phi(\mathbf{E}v) - f \cdot v - q \operatorname{div} v \, dx$ . We observe that for arbitrary  $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^{d \times d}$ ,  $\mathbf{Q} = (Q_{ij})_{i,j=1,\dots,d}$  it holds that

$$\sum_{i,j,k,l} \frac{\partial}{\partial x_{ij}} \frac{\partial}{\partial x_{kl}} \Phi(\mathbf{P}) Q_{ij} Q_{kl} = \frac{\phi'(|\mathbf{P}|)}{|\mathbf{P}|} \left( |\mathbf{Q}|^2 - \frac{|\mathbf{P} : \mathbf{Q}|}{|\mathbf{P}|^2} \right) + \phi''(|\mathbf{P}|) \frac{|\mathbf{P} : \mathbf{Q}|^2}{|\mathbf{P}|^2}.$$

Now, by (2.1) there exists constants  $C, c > 0$  such that  $c\phi'(t) \leq t\phi''(t) \leq C\phi'(t)$ . Hence, on the one hand we can estimate

$$\begin{aligned} \sum_{i,j,k,l} \frac{\partial}{\partial x_{ij}} \frac{\partial}{\partial x_{kl}} \Phi(\mathbf{P}) Q_{ij} Q_{kl} &\leq \frac{\phi'(|\mathbf{P}|)}{|\mathbf{P}|} |\mathbf{Q}|^2 + C \frac{\phi'(|\mathbf{P}|)}{|\mathbf{P}|^3} |\mathbf{P}|^2 |\mathbf{Q}|^2 \\ &= (1+C) \frac{\phi'(|\mathbf{P}|)}{|\mathbf{P}|} |\mathbf{Q}|^2, \end{aligned}$$

and on the other hand, we have

$$\begin{aligned} \sum_{i,j,k,l} \frac{\partial}{\partial x_{ij}} \frac{\partial}{\partial x_{kl}} \Phi(\mathbf{P}) Q_{ij} Q_{kl} &\geq \frac{\phi'(|\mathbf{P}|)}{|\mathbf{P}|} |\mathbf{Q}|^2 + (c-1) \frac{\phi'(|\mathbf{P}|)}{|\mathbf{P}|^3} |\mathbf{P} : \mathbf{Q}|^2 \\ &\geq c \frac{\phi'(|\mathbf{P}|)}{|\mathbf{P}|} |\mathbf{Q}|^2. \end{aligned}$$

Let  $g(t) := \mathcal{J}_q([u_q, v]_t)$  for  $t \in \mathbb{R}$ , where  $[u_q, v]_t := (1-t)u_q + tv$ . Since  $u_q$  minimizes  $\mathcal{J}_q$  we have  $g'(0) = 0$ . Combining Taylor's formula with the above estimates for  $\mathbf{P} = \mathbf{E}[u_q, v]_t$  and  $\mathbf{Q} = \mathbf{E}(u_q - v)$ , we get

$$\begin{aligned} \mathcal{J}_q(v) - \mathcal{J}_q(u_q) &= g(1) - g(0) = \frac{1}{2} \int_0^1 g''(t) (1-t) \, dt \\ (2.27) \quad &\approx \int_0^1 \int_{\Omega} \frac{\phi'(|[\mathbf{E}u_q, \mathbf{E}v]_t|)}{|[\mathbf{E}u_q, \mathbf{E}v]_t|} |\mathbf{E}v - \mathbf{E}u_q|^2 \, dx (1-t) \, dt. \end{aligned}$$

By [DE08, Lemma 19] it holds for any  $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^{d \times d}$  that

$$\int_0^1 \frac{\phi'(|[\mathbf{P}, \mathbf{Q}]_t|)}{|[\mathbf{P}, \mathbf{Q}]_t|} \, dt \approx \frac{\phi'(|\mathbf{P}| + |\mathbf{Q}|)}{|\mathbf{P}| + |\mathbf{Q}|}$$

with constants only depending on  $\Delta_2(\{\phi, \phi^*\})$  and thus on the constants in (2.1). This, (2.1), and Corollary 6 yield

$$\begin{aligned} \mathcal{J}_q(v) - \mathcal{J}_q(u_q) &\leq C \int_{\Omega} \frac{\phi'(|\mathbf{E}u_q| + |\mathbf{E}v|)}{|\mathbf{E}u_q| + |\mathbf{E}v|} |\mathbf{E}v - \mathbf{E}u_q|^2 \, dx \\ &\leq C \int_{\Omega} \phi''(|\mathbf{E}u_q| + |\mathbf{E}v|) |\mathbf{E}v - \mathbf{E}u_q|^2 \, dx \\ &\leq C \int_{\Omega} |\mathbf{F}(\mathbf{E}v) - \mathbf{F}(\mathbf{E}u_q)|^2 \, dx. \end{aligned}$$

On the other hand, we can further estimate (2.27) by (2.10), and Jensen's inequality to get

$$\begin{aligned} \mathcal{J}_q(v) - \mathcal{J}_q(u_q) &\geq \int_{\Omega} \int_0^1 \frac{\phi(|[\mathbf{E}u_q, \mathbf{E}v]_t|)}{(|\mathbf{E}u_q| + |\mathbf{E}v|)^2} (1-t) dt |\mathbf{E}v - \mathbf{E}u_q|^2 dx \\ &\geq \int_{\Omega} \frac{\phi(\int_0^1 |[\mathbf{E}u_q, \mathbf{E}v]_t| 2(1-t) dt)}{(|\mathbf{E}u_q| + |\mathbf{E}v|)^2} |\mathbf{E}v - \mathbf{E}u_q|^2 dx. \end{aligned}$$

Since  $\int_0^1 |[\mathbf{P}, \mathbf{Q}]_t| 2(1-t) dt$  as well as  $|\mathbf{P}| + |\mathbf{Q}|$  define norms on  $\mathbb{R}^{d \times d} \times \mathbb{R}^{d \times d}$  it holds that  $\int_0^1 |[\mathbf{P}, \mathbf{Q}]_t| 2(1-t) dt \approx |\mathbf{P}| + |\mathbf{Q}|$  for all  $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^{d \times d}$ . This and  $\phi''(t)t^2 \approx \phi(t)$  (see (2.1) and (2.10)) imply

$$\mathcal{J}_q(v) - \mathcal{J}_q(u_q) \geq \int_{\Omega} \phi''(|\mathbf{E}u_q| + |\mathbf{E}v|) |\mathbf{E}v - \mathbf{E}u_q|^2 dx.$$

Now, the desired assertion follows by Corollary 6.  $\square$

*Remark 14.* Obviously Lemma 13 stays valid if  $\mathbb{V}$  is substituted by a closed subspace  $\tilde{\mathbb{V}} \subset \mathbb{V}$  since all used properties of  $\mathcal{J}_q$  are inherited by any subspace of  $\mathbb{V}$ .

Analyzing the functional  $\mathcal{F}$  yields the following results.

**Proposition 15.** *Let  $\phi$  be an  $N$ -function such that (2.1) holds. Then it holds that:*

- i) *The mapping  $q \mapsto u_q$  defined in (2.26) is continuous from  $\mathbb{Q}$  to  $\mathbb{V}$ .*
- ii) *The functional  $\mathcal{F}$  from (2.23) is Fréchet differentiable in  $q \in \mathbb{Q}$  with derivative  $D\mathcal{F}(q) = \operatorname{div} u_q$  i.e.,*

$$\langle D\mathcal{F}(q), h \rangle = \int_{\Omega} h \operatorname{div} u_q dx, \quad \text{for } h \in \mathbb{Q}.$$

*Proof.* In order to prove i) let  $q, h \in \mathbb{Q}$ . Then by (2.25) we have

$$\int_{\Omega} (\mathbf{A}(\mathbf{E}u_q) - \mathbf{A}(\mathbf{E}u_{q+h})) : (\mathbf{E}u_q - \mathbf{E}u_{q+h}) dx = \int_{\Omega} h \operatorname{div}(u_q - u_{q+h}) dx.$$

Applying Young's inequality (2.15), for  $\delta > 0$ , we get

$$\begin{aligned} \int_{\Omega} (\mathbf{A}(\mathbf{E}u_q) - \mathbf{A}(\mathbf{E}u_{q+h})) : (\mathbf{E}u_q - \mathbf{E}u_{q+h}) dx \\ \leq \int_{\Omega} C_{\delta} (\phi_{|\mathbf{E}u_q|})^* (|h|) + \delta \phi_{|\mathbf{E}u_q|} (|\operatorname{div}(u_q - u_{q+h})|) dx. \end{aligned}$$

Observing  $|\operatorname{div} v| \leq \sqrt{d} |\mathbf{E}v|$ , by (2.11) and (2.17) we get

$$\begin{aligned} \int_{\Omega} (\mathbf{A}(\mathbf{E}u_q) - \mathbf{A}(\mathbf{E}u_{q+h})) : (\mathbf{E}u_q - \mathbf{E}u_{q+h}) dx \\ \leq \int_{\Omega} C_{\delta} \phi_{|\mathbf{E}u_q|}^* (|h|) + \delta \phi_{|\mathbf{E}u_q|} (|\mathbf{E}(u_q - u_{q+h})|) dx. \end{aligned}$$

By Proposition 5 the left-hand side is proportional to  $\int_{\Omega} \phi_{|\mathbf{E}u_q|} (|\mathbf{E}(u_q - u_{q+h})|) dx$  and hence, for  $\delta > 0$  small enough, we get

$$\int_{\Omega} \phi_{|\mathbf{E}u_q|} (|\mathbf{E}(u_q - u_{q+h})|) dx \leq \int_{\Omega} C_{\delta} (\phi_{|\mathbf{E}u_q|})^* (|h|).$$

Proposition 9 yields  $\mathbf{E}u_{q+h} \rightarrow \mathbf{E}u_q$  in  $\mathbb{V}$  as  $h \rightarrow 0$  in  $\mathbb{Q}$  and hence the first assertion follows by Korn's inequality.

To prove the second claim ii) we observe by (2.26) and (2.24) that

$$\mathcal{F}(q+h) - \mathcal{F}(q) - \int_{\Omega} h \operatorname{div} u_q dx = \mathcal{J}_q(u_q) - \mathcal{J}_q(u_{q+h}) - \int_{\Omega} h \operatorname{div}(u_q - u_{q+h}) dx.$$

Lemma 13 and (2.25) imply, for  $c \in \mathbb{R}$ ,

$$\begin{aligned} \mathcal{F}(q+h) - \mathcal{F}(q) - \int_{\Omega} h \operatorname{div} u_q dx & \\ & \approx \int_{\Omega} (\mathbf{A}(\mathbf{E}u_q) - \mathbf{A}(\mathbf{E}u_{q+h})) : (\mathbf{E}u_q - \mathbf{E}u_{q+h}) dx + \int_{\Omega} h \operatorname{div}(u_q - u_{q+h}) dx \\ & = 2 \int_{\Omega} (h - c) \operatorname{div}(u_q - u_{q+h}) dx, \leq 4 \|h - c\|_{(\phi^*)} \|\operatorname{div}(u_q - u_{q+h})\|_{(\phi)}, \end{aligned}$$

where we used Hölder's inequality in the last step; see [RR91]. The continuity of  $q \mapsto u_q$  yields that this expression is  $o(\|h\|_{(\phi^*)})$  and hence the assertion is proved.  $\square$

### 3. AFEM FOR THE NONLINEAR LAPLACIAN PROBLEM

As is motivated in the previous section we are interested in solving nonlinear Laplace equations of the following kind: Find  $\tilde{u} \in \mathbb{V} = W_0^{1,\phi}(\Omega)^d$  such that

$$(3.1) \quad -\operatorname{div} \mathbf{A}(\mathbf{E}\tilde{u}) = g \quad \text{in } W^{-1,\phi^*}(\Omega).$$

It turns out that for this kind of problem the usual energy norm is not the appropriate concept of distance; see for example [Vee02, Remark 3.5]. For the power law Barrett and Liu provided in [BL93a, BL94] with the so-called quasi-norm, an optimal error measure; cf. Remark 11. In this section we present a convergent AFEM, which shows quasi-optimal convergence rates; see [DK08, BDK10] and Remark 18.

However, in order to apply the results of [DK08, BDK10] we have to restrict ourselves to right-hand sides  $g \in L^{\phi^*}(\Omega)^d \subset W^{-1,\phi^*}(\Omega)^d$ . The difference of the subsequent results compared to the results presented in [DK08, BDK10] is that we consider vector-valued solutions and the symmetric gradient instead of the gradient. Since the results can be transferred straight forward (compare with [Kre08]) we omit the proofs and only state the results.

A conforming triangulation  $\mathcal{T}$  of  $\Omega$  is a finite set of closed  $d$  simplices such that  $\Omega = \bigcup_{T \in \mathcal{T}} T$  and for each pair of simplices  $T, T' \in \mathcal{T}$  the intersection  $T \cap T'$  is either empty or a common sub-simplex of  $T$  and  $T'$ . We define  $\mathbb{V}(\mathcal{T}) \subset \mathbb{V}$  to be the space of piecewise affine continuous functions over  $\mathcal{T}$  that are zero at the boundary  $\partial\Omega$ . We will further denote as  $\mathcal{S} = \mathcal{S}(\mathcal{T})$  the set of closed element-sides of  $\mathcal{T}$  interior  $\Omega$  and as  $\mathcal{N}(\mathcal{T})$  the set of element-vertices. For  $\omega \subset \Omega$  let  $\mathcal{T}(\omega) := \{T \in \mathcal{T} : T \subset \omega\}$ . For  $T \in \mathcal{T}$  we denote by  $S_T := \bigcup\{T' \in \mathcal{T} | T' \cap T \neq \emptyset\}$  the patch around  $T$  and by  $\Sigma_T = \{\sigma \in \mathcal{S} : \sigma \cap T \neq \emptyset\}$  the set of sides interior to  $S_T$ .

For the rest of this paper the constants hidden in  $\lesssim$ ,  $\gtrsim$ , and  $\approx$  may additionally depend on the shape regularity of triangulations.

We define the local error indicator for  $v \in \mathbb{V}$ ,  $W \in \mathbb{V}(\mathcal{T})$  on  $T \in \mathcal{T}$  by

$$(3.2) \quad \eta^2(v, W, T, g) := \int_T (\phi_{|\mathbf{E}v|})^* (h_T |g|) dx + \int_{\partial T \cap \Omega} h_T \|\llbracket \mathbf{F}(\mathbf{E}W) \rrbracket\|^2 d\sigma,$$

where  $h_T := |T|^{1/d}$  is the mesh-size of  $T$ . Hereafter,  $\llbracket q \rrbracket |_{\sigma}$  is the jump of  $q$  across an interior side  $\sigma \in \mathcal{S}$ . The first term in (3.2) usually is called the element-residual,

whereas the second part is called the jump-residual. Furthermore, we define for any subset  $\hat{\mathcal{T}} \subset \mathcal{T}$ ,

$$\eta^2(v, W, \hat{\mathcal{T}}, g) := \sum_{T \in \hat{\mathcal{T}}} \eta^2(v, W, T, g) \quad \text{and} \quad \eta(W, \hat{\mathcal{T}}, g) := \eta(W, W, \hat{\mathcal{T}}, g).$$

Let  $\tilde{u} \in \mathbb{V}$  be the solution of (3.1) and  $\tilde{U} \in \mathbb{V}(\mathcal{T})$  its finite element approximation i.e.,

$$(3.3) \quad \int_{\Omega} \mathbf{A}(\mathbf{E}\tilde{U}) : \mathbf{E}V \, dx = \int_{\Omega} g \cdot V \, dx \quad \text{for all } V \in \mathbb{V}(\mathcal{T}).$$

Then by [DK08, Kre08] there holds the upper bound

$$(3.4) \quad \|\mathbf{F}(\mathbf{E}\tilde{u}) - \mathbf{F}(\mathbf{E}\tilde{U})\|_{L^2(\Omega)} \leq C_1 \eta(\tilde{U}, \mathcal{T}, g)$$

as well as the lower bound

$$(3.5) \quad C_2 \eta(\tilde{U}, \mathcal{T}, g) \leq \|\mathbf{F}(\mathbf{E}\tilde{u}) - \mathbf{F}(\mathbf{E}\tilde{U})\|_{L^2(\Omega)} + \text{osc}(\tilde{U}, \mathcal{T}, g),$$

where the constants  $C_1, C_2 > 0$  depend solely on the dimension  $d$ , the shape regularity of  $\mathcal{T}$  and the constants in (2.1). The oscillation is defined by

$$\text{osc}^2(v, \mathcal{T}, g) := \sum_{T \in \mathcal{T}} \inf_{g_T \in \mathbb{R}} \int_T (\phi_{|\mathbf{E}v|})^* (h_T |g - g_T|) \, dx \quad \text{for } v \in \mathbb{V}.$$

Based on these definitions we can formulate the following adaptive finite element method: Let  $g \in L^{\phi^*}(\Omega)$ ,  $\mathcal{T}$  be a conforming triangulation of  $\Omega$ ,  $\epsilon > 0$ , and  $\theta \in (0, 1]$ .

**Algorithm 16** (AFEM( $g, \mathcal{T}, \epsilon, \theta$ )). Let  $k = 0$ ,  $\tilde{\mathcal{T}}_0 = \mathcal{T}$ ;

- (1)  $\tilde{U}_k = \text{SOLVE}(\tilde{\mathcal{T}}_k, g)$ ;
- (2)  $\{\eta(\tilde{U}_k, T, g)\}_{T \in \tilde{\mathcal{T}}_k} = \text{ESTIMATE}(\tilde{U}_k, \tilde{\mathcal{T}}_k, g)$ ;
- (3) if  $\eta(\tilde{U}_k, \tilde{\mathcal{T}}_k, g) < \epsilon$ , then  

RETURN  $(\tilde{U}_k, \tilde{\mathcal{T}}_k)$ ;
- (4)  $\mathcal{M}_k = \text{MARK}(\{\eta(\tilde{U}_k, T, g)\}_{T \in \tilde{\mathcal{T}}_k}, \tilde{\mathcal{T}}_k, \theta)$ ;
- (5)  $\tilde{\mathcal{T}}_{k+1} = \text{REFINE}(\tilde{\mathcal{T}}_k, \mathcal{M}_k)$ ; increment  $k$  and go to step (1);

We sketch the single modules of the AFEM. For any conforming triangulation  $\mathcal{T}$  of  $\Omega$  we suppose that the routine SOLVE outputs the exact Ritz-Galerkin solution  $\tilde{U} \in \mathbb{V}(\mathcal{T})$  of (3.3).

Next, the error between the discrete solution  $\tilde{U}$  and the continuous solution  $\tilde{u}$  of (3.1) is estimated by ESTIMATE i.e.,

$$\{\eta(\tilde{U}, T, g)\}_{T \in \mathcal{T}} = \text{ESTIMATE}(\tilde{U}, \mathcal{T}, g).$$

In the selection of elements for refinement we rely on Dörfler marking. For a given marking parameter  $\theta \in (0, 1]$ , we suppose that MARK outputs a subset  $\mathcal{M} \subset \mathcal{T}$  of marked elements i.e.,

$$\mathcal{M} = \text{MARK}(\{\eta(\tilde{U}, T, g)\}_{T \in \mathcal{T}}, \mathcal{T}, \theta),$$

such that  $\mathcal{M}$  satisfies the Dörfler property

$$\eta(\tilde{U}, \mathcal{M}, g) \geq \theta \eta(\tilde{U}, \mathcal{T}, g).$$

Refinement is based on shape-regular bisection of single elements. We do not go into very much detail and just assume that there exists a procedure REFINE, that

produces a conforming refinement of a given triangulation  $\mathcal{T}$  based on a certain subset  $\mathcal{M} \subset \mathcal{T}$  of marked elements. In particular, let

$$\mathcal{T}_* = \text{REFINE}(\mathcal{T}, \mathcal{M}),$$

then  $\mathcal{T}_*$  is a conforming triangulation of  $\Omega$  such that  $T \in \mathcal{M}$  is at least bisected once. Shape regularity of meshes produced by repeated applications of **REFINE** from some initial conforming triangulation  $\mathcal{T}_0$  is uniformly bounded depending on the shape regularity of  $\mathcal{T}_0$ . For the existence of such a procedure **REFINE** we refer to [Bän91, Mau95, Kos94, Mit89, SS05, Ste07, Ste08].

The following theorem shows convergence of the AFEM and can be found in [BDK10, Kre08].

**Theorem 17.** *Let  $g \in L^{\phi^*}(\Omega)^d$ ,  $\mathcal{T}$  being a conforming triangulation of the polyhedral domain  $\Omega$ ,  $\epsilon > 0$ , and  $\theta \in (0, 1]$ . Let  $\tilde{u} \in \mathbb{V}$  be the solution of (3.1). Then, for  $(\tilde{U}_k, \tilde{\mathcal{T}}_k)_{k \in \mathbb{N}}$  being the sequence of solutions and triangulations produced by AFEM( $g, \mathcal{T}, \epsilon, \theta$ ), there exists  $\alpha \in (0, 1)$  depending solely on the shape-regularity of  $\tilde{\mathcal{T}}$ ,  $\theta$ , but not on  $\epsilon$ , such that*

$$\begin{aligned} \|\mathbf{F}(\mathbf{E}\tilde{u}) - \mathbf{F}(\mathbf{E}\tilde{U}_k)\|_{L^2(\Omega)}^2 + \eta^2(\tilde{U}_k, \tilde{\mathcal{T}}_k, g) \\ \preccurlyeq \alpha^k (\|\mathbf{F}(\mathbf{E}\tilde{u}) - \mathbf{F}(\mathbf{E}\tilde{U}_0)\|_{L^2(\Omega)}^2 + \eta^2(\tilde{U}_0, \mathcal{T}, g)) \end{aligned}$$

for all  $k \geq 0$  until the algorithm stops.

This result ensures that AFEM stops and outputs a finite element solution with the corresponding triangulation

$$(3.6) \quad (\tilde{U}_*, \mathcal{T}_*) = \text{AFEM}(g, \mathcal{T}, \epsilon, \theta) \quad \text{such that} \quad \eta(\tilde{U}_*, \mathcal{T}_*, g) < \epsilon.$$

*Remark 18* (optimality). Marking in each iteration  $k$  a minimal subset  $\mathcal{M}_k \subset \mathcal{T}_k$  with Dörflers marking strategy, it can be shown as in [BDK10], that for all marking parameters  $\theta$  below a threshold  $\theta_* \in (0, 1)$  the AFEM yields quasi-optimal meshes. In particular, assume that there exist constants  $s > 0$  and  $C > 0$ , such that for all  $\epsilon > 0$  there exists a conforming refinement  $\mathcal{T}$  of  $\mathcal{T}_0$  satisfying

$$\inf_{V \in \mathbb{V}(\mathcal{T})} \{ \|\mathbf{F}(\mathbf{E}\tilde{u}) - \mathbf{F}(\mathbf{E}V)\|_{L^2(\Omega)} + \text{osc}(\tilde{u}, \mathcal{T}_*, g) \} \leq \epsilon$$

and

$$\#\mathcal{T}_* - \#\mathcal{T} \leq C \epsilon^{-1/s}.$$

Then, if  $\mathcal{T}$  satisfies condition b) of §4 in [Ste08], there exists a constant  $\theta_* \in (0, 1)$  such that for all  $\theta \in (0, \theta_*)$  the sequence  $(\tilde{U}_k, \mathcal{T}_k)_k$  of solutions and meshes produced by AFEM( $g, \mathcal{T}, \theta, \theta$ ) satisfies

$$\|\mathbf{F}(\mathbf{E}\tilde{u}) - \mathbf{F}(\mathbf{E}\tilde{U}_k)\|_{L^2(\Omega)} + \text{osc}(\tilde{u}, \tilde{\mathcal{T}}_k, g) \preccurlyeq (\#\tilde{\mathcal{T}}_k - \#\mathcal{T})^{-s},$$

where the constant hidden in  $\preccurlyeq$  solely depends on the triangulation  $\mathcal{T}$ , the N-function  $\phi$  and the marking parameter  $\theta$ . In other words, the sequence  $(\tilde{U}_k)_k$  converges to  $\tilde{u}$  with quasi-optimal cardinality. This performance is also documented by the numerical experiments in [BDK10].

We emphasize that this result is proved only for the nonlinear Poisson problem with fixed right-hand side  $g$ . Therefore, it cannot be directly transferred to the adaptive Uzawa algorithm, since here the AFEM is used within an exterior loop

and in general starts each time from a different mesh and with different right-hand side  $g$ ; compare with Algorithm 23.

#### 4. QUASI-STEEPEST DESCENT DIRECTION

The experiences of §3 indicate that, for nonlinear problems, norms might not be the appropriate concept of distance for the pressure as well. Recalling Corollary 6, the quasi-norm is a quantity, which is equivalent to the residual tested with the error. We fix  $q \in \mathbb{Q}$ . Carrying over these ideas to the functional  $\mathcal{F}$  suggests to test the residual  $D\mathcal{F}(q) \in L_0^\phi(\Omega)$  with the error  $q - p$ , where  $p \in \mathbb{Q}$  is the minimizer of  $\mathcal{F}$  defined in (2.23). Let  $u_q$  defined as in (2.25) and recall from (2.4) that  $D\mathcal{F}(p) = \operatorname{div} u = 0$ . Hence by Corollary 6 we have

$$\begin{aligned} \langle D\mathcal{F}(q), q - p \rangle &= \int_{\Omega} (\mathbf{A}(\mathbf{E}u_q) - \mathbf{A}(\mathbf{E}u)) : (\mathbf{E}u_q - \mathbf{E}u) \, dx \\ &\approx \int_{\Omega} \phi_{|\mathbf{E}u_q|} (|\mathbf{E}u_q - \mathbf{E}u|) \, dx, \end{aligned}$$

which indicates to use  $\int_{\Omega} \phi_{|\mathbf{E}u_q|} (|\mathbf{E}u_q - \mathbf{E}u|) \, dx$  as distance measure for the pressure.

This leads to the question of what is a *steepest descent direction* in this context. For norms a steepest descent direction  $\mathfrak{d} \in \mathbb{Q}$  of  $D\mathcal{F}$  in  $q \in \mathbb{Q}$  is defined by

$$\|D\mathcal{F}(q)\|_{\mathbb{Q}^*} = \sup_{h \in \mathbb{Q}, \|h\|_{\phi^*} = 1} \langle D\mathcal{F}(q), h \rangle = - \left\langle D\mathcal{F}(q), \frac{\mathfrak{d}}{\|\mathfrak{d}\|_{\phi}} \right\rangle.$$

To generalize this principle, we have to generalize the dual or operator norm. In the case of  $\phi(t) = \frac{1}{2}t^2$ , when the two concepts coincide, we know for  $l \in L_0^2(\Omega) = (L_0^2(\Omega))^*$ , that

$$\frac{1}{2} \|l\|_{L_0^2(\Omega)^*}^2 = \sup_{h \in L_0^2(\Omega)} \left\{ \langle l, h \rangle - \frac{1}{2} \|h\|_{L^2(\Omega)}^2 \right\} = \sup_{h \in L_0^2(\Omega)} \left\{ \langle l, h \rangle - \int_{\Omega} \phi(|h|) \, dx \right\}.$$

Recall that  $(L_0^\phi(\Omega))^* = L^{\phi^*}(\Omega)/\mathbb{R}$ , then this motivates the following definition of the dual error concept. For  $l \in L_0^\phi(\Omega)$ ,  $w \in W_0^{1,\phi}(\Omega)$ , we define

$$(4.1) \quad \|l\|_{(\mathbf{E}u_q),*}^2 := \sup_{h \in L^{\phi^*}(\Omega)} \left\{ \langle l, h \rangle - \inf_{c \in \mathbb{R}} \int_{\Omega} \phi_{|\mathbf{E}u_q|}^* (|h - c|) \, dx \right\}.$$

Young's inequality (2.14) and the observation  $\phi'(l) \frac{l}{|l|} \in L^{\phi^*}(\Omega)/\mathbb{R}$  yield

$$\|l\|_{(\mathbf{E}u_q),*}^2 = \left\langle l, \phi'_{|\mathbf{E}u_q|} (|l|) \frac{l}{|l|} \right\rangle - \int_{\Omega} \phi_{|\mathbf{E}u_q|}^* (\phi'_{|\mathbf{E}u_q|} (|l|)) \, dx = \int_{\Omega} \phi_{|\mathbf{E}u_q|} (|l|) \, dx;$$

see [Kre08]. Hence  $\phi'_{|\mathbf{E}u_q|} (|l|) \frac{l}{|l|}$  seems to be the natural *quasi-steepest descent direction*.

**Definition 19** (quasi-steepest descent direction). Let  $\phi$  be an N-function such that (2.1). Then, the quasi-steepest descent direction with respect to  $\mathcal{F}$  in  $q \in \mathbb{Q}$  is defined as

$$(4.2) \quad \mathfrak{d}_q := -\phi'_{|\mathbf{E}u_q|} (|\operatorname{div} u_q|) \frac{\operatorname{div} u_q}{|\operatorname{div} u_q|}.$$

**4.1. Approximation of the quasi-steepest descent direction.** As we know from Definition 19, we have to solve the nonlinear elliptic system (2.25) for the quasi-steepest descent direction. Recall that the AFEM yields linear convergence for a right-hand side  $g \in L^{\phi^*}(\Omega)^d$ . Therefore, due to the right-hand side of (2.25) it is convenient that the gradient of the pressure is in  $L^{\phi}(\Omega)^d$ . Thus, for  $\mathcal{T}$  being a conforming triangulation of  $\Omega$ , we define

$$\mathbb{Q}(\mathcal{T}) := \{Q \in C(\Omega) : Q|_T \in \mathcal{P}^1(T) \text{ for all } T \in \mathcal{T}\};$$

hence  $\nabla Q \in L^{\phi^*}(\Omega)^d$  for  $Q \in \mathbb{Q}(\mathcal{T})$ . Note that  $\mathbb{Q}(\mathcal{T})$  is not a subspace of  $\mathbb{Q}$  but  $\mathbb{Q}(\mathcal{T})/\mathbb{R} \subset \mathbb{Q}$ . We use the functions in  $\mathbb{Q}(\mathcal{T})$  as representants of those in  $\mathbb{Q}(\mathcal{T})/\mathbb{R}$  and say that two of them are equal if they differ by a constant value.

The aim is to calculate an approximation  $\mathfrak{D}_Q \in \mathbb{Q}(\mathcal{T})$  of  $\mathfrak{d}_Q$  with the help of the AFEM. Let  $\mathbb{Q}_D(\mathcal{T})$  be the space of piecewise constants over  $\mathcal{T}$ , then  $\operatorname{div} V \in \mathbb{Q}_D(\mathcal{T})$ ,  $V \in \mathbb{V}(\mathcal{T})$ , hence for  $(U_Q, \mathcal{T}^*) = \text{AFEM}(f - \nabla Q, \mathcal{T}, \epsilon, \theta)$  we cannot simply take

$$(4.3) \quad \phi'_{|\mathbf{E}U_Q|}(|\operatorname{div} U_Q|) \frac{\operatorname{div} U_Q}{|\operatorname{div} U_Q|} \in \mathbb{Q}_D(\mathcal{T}^*) \not\subset \mathbb{Q}(\mathcal{T}^*),$$

$Q \in \mathbb{Q}(\mathcal{T})$ , as an approximation of  $\mathfrak{d}_Q$ .

To overcome this problem we use an interpolation operator  $\Pi_{\mathcal{T}}^{\mathbb{Q}} : L^1(\Omega) \rightarrow \mathbb{Q}(\mathcal{T})$ , which is closely related to the Cl ement operator [Cl 75, BMN02]:

For  $z \in \mathcal{N}(\mathcal{T})$  let  $\omega_z := \operatorname{supp} \Phi_z$  be the support of the corresponding Lagrange basis-function  $\Phi_z$  of  $\mathbb{Q}(\mathcal{T})$ . For  $q \in L^1(\Omega)$  let  $\Pi_{\mathcal{T}}^z : L^1(\Omega) \rightarrow \mathcal{P}^1(\omega_z)$  be the  $L^2$ -projection into the space of continuous piecewise linear polynomials i.e.,

$$(4.4) \quad \int_{\omega_z} (q - \Pi_{\mathcal{T}}^z q) Q \, dx = 0 \quad \text{for all } Q \in \mathcal{P}^1(\omega_z).$$

We then set  $\Pi_{\mathcal{T}}^{\mathbb{Q}} q(z) := \Pi_{\mathcal{T}}^z q(z)$ ; hence,  $\Pi_{\mathcal{T}}^{\mathbb{Q}} q = \sum_{z \in \mathcal{N}} \Pi_{\mathcal{T}}^z q(z) \Phi_z \in \mathbb{Q}(\mathcal{T})$ . Note that  $\Pi_{\mathcal{T}}^{\mathbb{Q}} : L^1(\Omega) \rightarrow \mathbb{Q}(\mathcal{T})$  is a projection; see [Cl 75].

The following estimates based on the  $L^1$ -norm use standard arguments; compare with [Cl 75, BMN02, Kre08]. Since we want to focus on the subsequent quasi-norm estimates, we decided not to prove them in detail. Observe from [Cl 75, Kre08] that to estimate  $q - \Pi_{\mathcal{T}}^{\mathbb{Q}} q$  it suffices to bound  $q - \Pi_{\mathcal{T}}^z q$  over  $\omega_z$ ,  $z \in \mathcal{N}$ . In particular, we have

$$(4.5) \quad \int_T |q - \Pi_{\mathcal{T}}^{\mathbb{Q}} q| \, dx \preccurlyeq \sum_{z \in \mathcal{N} \cap T} \int_T |q - \Pi_{\mathcal{T}}^z q| \, dx.$$

The next estimate is an adaption of the  $L^2$ -estimate from [BMN02] to the  $L^1$ -case. It makes use of the fact that the functions we focus on lie in  $\mathbb{Q}_D(\mathcal{T}) \subset L^1(\Omega)$ , which in turn is finite-dimensional. Let  $Q \in \mathbb{Q}_D(\mathcal{T}) \subset L^1(\Omega)$ , then we can estimate

$$(4.6) \quad \int_{\omega_z} |Q - \Pi_{\mathcal{T}}^z Q| \, dx \preccurlyeq \operatorname{diam}(\omega_z) \int_{\sigma_z} \| [Q] \| \, d\sigma,$$

where  $\sigma_z := \bigcup \{\sigma \in \mathcal{S}(\mathcal{T}) \mid z \cap \sigma = z\}$  is the union of sides interior  $\omega_z$ . Scaling the left-hand side to a reference situation, observing the fact that the jump of  $Q$  across inter-element sides is a norm for  $Q - \Pi_z Q$ , using the fact that norms on the finite dimensional spaces are equivalent, and scaling back to the physical element proves the result.

The goal is now to quantify the difference between  $\mathfrak{d}_Q$ ,  $Q \in \mathbb{Q}(\mathcal{T})$ , and its approximation

$$(4.7) \quad \mathfrak{D}_Q := \Pi_{\mathcal{T}}^{\mathbb{Q}} \phi'_{|\mathbf{E}U_Q|}(|\operatorname{div} U_Q|) \frac{\operatorname{div} U_Q}{|\operatorname{div} U_Q|} \in \mathbb{Q}(\mathcal{T}).$$

But before we turn to this problem, we need to prove the following lemma, which estimates the distance between arbitrary descent directions.

**Lemma 20.** *Let  $\phi$  be an N-function that satisfies assumption (2.1). For  $v, w \in \mathbb{V}$ , we set*

$$\mathfrak{d}(v) := \phi'_{|\mathbf{E}v|}(|\operatorname{div} v|) \frac{\operatorname{div} v}{|\operatorname{div} v|} \quad \text{and} \quad \mathfrak{d}(w) := \phi'_{|\mathbf{E}w|}(|\operatorname{div} w|) \frac{\operatorname{div} w}{|\operatorname{div} w|}.$$

Then, for all  $v, w \in \mathbb{V}$  it holds that

$$\int_{\Omega} (\phi_{|\mathbf{E}v|})^* (|\mathfrak{d}(v) - \mathfrak{d}(w)|) dx \preceq \|\mathbf{F}(\mathbf{E}v) - \mathbf{F}(\mathbf{E}w)\|_{L^2(\Omega)}^2.$$

*Proof.* By Lemma 7, Lemma 8, and (2.12), it holds that

$$\begin{aligned} & \int_{\Omega} (\phi_{|\mathbf{E}v|})^* (|\mathfrak{d}(v) - \mathfrak{d}(w)|) dx \\ &= \int_{\Omega} (\phi_{|\mathbf{E}v|})^* \left( \left| \phi'_{|\mathbf{E}v|}(|\operatorname{div} v|) \frac{\operatorname{div} v}{|\operatorname{div} v|} - \phi'_{|\mathbf{E}w|}(|\operatorname{div} w|) \frac{\operatorname{div} w}{|\operatorname{div} w|} \right| \right) dx \\ &\preceq \int_{\Omega} (\phi_{|\mathbf{E}v|})^* \left( \left| \phi'_{|\mathbf{E}v|}(|\operatorname{div} v|) \frac{\operatorname{div} v}{|\operatorname{div} v|} - \phi'_{|\mathbf{E}v|}(|\operatorname{div} w|) \frac{\operatorname{div} w}{|\operatorname{div} w|} \right| \right) dx \\ &\quad + \int_{\Omega} (\phi_{|\mathbf{E}v|})^* (\phi'_{|\mathbf{E}v|}(|\mathbf{E}v - \mathbf{E}w|)) dx. \end{aligned}$$

Applying Proposition 5 in 1-dimension for the N-function  $\phi_{|\mathbf{E}v|}$  to the first addend and (2.10) to the second yields

$$\begin{aligned} & \int_{\Omega} (\phi_{|\mathbf{E}v|})^* (|\mathfrak{d}(v) - \mathfrak{d}(w)|) dx \\ &\preceq \int_{\Omega} (\phi_{|\mathbf{E}v|})^* \left( (\phi'_{|\mathbf{E}v|})_{|\operatorname{div} v|} (|\operatorname{div} v - \operatorname{div} w|) \right) dx + \int_{\Omega} \phi_{|\mathbf{E}v|} (|\mathbf{E}v - \mathbf{E}w|) dx. \end{aligned}$$

Observe that  $|\operatorname{div} v| \leq \sqrt{d} |\mathbf{E}v|$  and that (2.11) holds also for  $\phi'$  instead of  $\phi$  since (2.3). Hence, by the monotonicity of  $\phi'$ ,

$$(4.8) \quad \begin{aligned} (\phi'_{|\mathbf{E}v|})_{|\operatorname{div} v|}(t) &= \phi'_{|\mathbf{E}v|+|\operatorname{div} v|}(t) = \frac{\phi'(|\mathbf{E}v| + |\operatorname{div} v| + t)}{|\mathbf{E}v| + |\operatorname{div} v| + t} t \\ &\leq \frac{\phi'((1 + \sqrt{d})(|\mathbf{E}v| + t))}{|\mathbf{E}v| + t} t \preceq \frac{\phi'(|\mathbf{E}v| + t)}{|\mathbf{E}v| + t} t = \phi'_{|\mathbf{E}v|}(t) \end{aligned}$$

for all  $t \geq 0$ . Therefore, by (2.11) and (2.10),

$$\begin{aligned} & \int_{\Omega} (\phi_{|\mathbf{E}v|})^* (|\mathfrak{d}(v) - \mathfrak{d}(w)|) dx \\ &\preceq \int_{\Omega} (\phi_{|\mathbf{E}v|})^* (\phi'_{|\mathbf{E}v|}(|\operatorname{div} v - \operatorname{div} w|)) dx + \int_{\Omega} \phi_{|\mathbf{E}v|} (|\mathbf{E}v - \mathbf{E}w|) dx \end{aligned}$$

$$\lesssim \int_{\Omega} \phi_{|\mathbf{E}v|} (|\operatorname{div} v - \operatorname{div} w|) dx + \int_{\Omega} \phi_{|\mathbf{E}v|} (|\mathbf{E}v - \mathbf{E}w|) dx.$$

Again using  $|\operatorname{div}(v - w)| \leq \sqrt{d} |\mathbf{E}(v - w)|$  and (2.11) yields

$$\lesssim \int_{\Omega} \phi_{|\mathbf{E}v|} (|\mathbf{E}v - \mathbf{E}w|) dx.$$

Applying Corollary 6 yields the assertion.  $\square$

In the next lemma we generalize (4.6) to the quasi-norm case. The result is crucial for estimating the error occurring during interpolation.

**Lemma 21.** *Let  $\mathcal{T}$  be a conforming triangulation of  $\Omega$  and  $\phi$  an  $N$ -function such that we get (2.1). For  $V \in \mathbb{V}(\mathcal{T})$  let  $\mathfrak{d} := \phi_{|\mathbf{E}V|} (|\operatorname{div} V|) \frac{\operatorname{div} V}{|\operatorname{div} V|}$ . Then,*

$$\int_{\Omega} (\phi_{|\mathbf{E}V|})^* (|\mathfrak{d} - \Pi_{\mathcal{T}}^{\mathbb{Q}} \mathfrak{d}|) dx \lesssim \sum_{T \in \mathcal{T}} \int_{\partial T \cap \Omega} h_T \|\mathbf{F}(\mathbf{E}V)\|^2 d\sigma.$$

*Proof.* Let  $T \in \mathcal{T}$ . We observe that  $\mathfrak{d} \in \mathbb{Q}_D(\mathcal{T})$ . Therefore, scaling  $\mathfrak{d}$  to the reference element  $\hat{T}$ , applying equivalence of norms on finite dimensional spaces, and scaling back to the physical element  $T$ , we obtain

$$\sup_T |\mathfrak{d} - \Pi_{\mathcal{T}}^{\mathbb{Q}} \mathfrak{d}| = \sup_{\hat{T}} |\hat{\mathfrak{d}} - \widehat{\Pi_{\mathcal{T}}^{\mathbb{Q}} \mathfrak{d}}| \lesssim \int_{\hat{T}} |\hat{\mathfrak{d}} - \widehat{\Pi_{\mathcal{T}}^{\mathbb{Q}} \mathfrak{d}}| dx \lesssim \frac{1}{|T|} \int_T |\mathfrak{d} - \Pi_{\mathcal{T}}^{\mathbb{Q}} \mathfrak{d}| dx.$$

Thus, we can apply (4.5) and (4.6) to get

$$\begin{aligned} \sup_T |\mathfrak{d} - \Pi_{\mathcal{T}}^{\mathbb{Q}} \mathfrak{d}| &\lesssim \frac{1}{|T|} \sum_{z \in \mathcal{N} \cap T} \int_T |\mathfrak{d} - \Pi_{\mathcal{T}}^z \mathfrak{d}| dx \leq \frac{1}{|T|} \sum_{z \in \mathcal{N} \cap T} \int_{\omega_z} |\mathfrak{d} - \Pi_{\mathcal{T}}^z \mathfrak{d}| dx \\ &\lesssim \frac{1}{|T|} \sum_{z \in \mathcal{N} \cap T} \operatorname{diam}(\omega_z) \int_{\sigma_z} \|\mathfrak{d}\| d\sigma \lesssim \sum_{z \in \mathcal{N} \cap T} \frac{1}{|\sigma_z|} \int_{\sigma_z} \|\mathfrak{d}\| d\sigma, \end{aligned}$$

where we have used  $\frac{\operatorname{diam}(\omega_z)}{|T|} \approx \frac{1}{|\sigma_z|}$  depending on the shape-regularity of  $\mathcal{T}$ . Since  $\#(\mathcal{N} \cap T)$  is bounded by  $d + 1$ , this estimate yields with (2.11) and (2.12) that

$$(\phi_{|\mathbf{E}V|})^* \left( \sup_T |\mathfrak{d} - \Pi_{\mathcal{T}}^{\mathbb{Q}} \mathfrak{d}| \right) \lesssim \sum_{z \in \mathcal{N} \cap T} (\phi_{|\mathbf{E}V|})^* \left( \frac{1}{|\sigma_z|} \int_{\sigma_z} \|\mathfrak{d}\| d\sigma \right).$$

Jensen's inequality and integration over  $T$  imply for the fixed shift  $|\mathbf{E}V|_T$ ,

$$\begin{aligned} \int_T (\phi_{|\mathbf{E}V|})^* (|\mathfrak{d} - \Pi_{\mathcal{T}}^{\mathbb{Q}} \mathfrak{d}|) dx &\lesssim \int_T \sum_{z \in \mathcal{N} \cap T} \frac{1}{|\sigma_z|} \int_{\sigma_z} (\phi_{|\mathbf{E}V|_T})^* (\|\mathfrak{d}\|) d\sigma dx \\ &\lesssim \sum_{z \in \mathcal{N} \cap T} \int_{\sigma_z} h_T (\phi_{|\mathbf{E}V|_T})^* (\|\mathfrak{d}\|) d\sigma \lesssim \sum_{T' \subset S_T} \int_{\partial T' \cap \Omega} h_T (\phi_{|\mathbf{E}V|_T})^* (\|\mathfrak{d}\|) d\sigma, \end{aligned}$$

where we used that  $h_T \approx \frac{|T|}{|\sigma_z|}$  depending on the shape regularity of  $\mathcal{T}$ . We notice that the shift  $|\mathbf{E}V|_T$  is constant on  $S_T$ . Hence, in order to get the shift compatible to each  $T' \in S_T$  we change it according to (2.20) and use Proposition 5 to obtain

$$\begin{aligned} \int_T (\phi_{|\mathbf{E}V|})^* (|\mathfrak{d} - \Pi_{\mathcal{T}}^{\mathbb{Q}} \mathfrak{d}|) dx &\lesssim \sum_{T' \subset S_T} \int_{\partial T' \cap \Omega} h_T (\phi_{|\mathbf{E}V|})^* (\|\mathfrak{d}\|) d\sigma \\ &\quad + \sum_{T' \subset S_T} h_T \int_{\partial T' \cap \Omega} |\mathbf{F}(\mathbf{E}V|_T) - \mathbf{F}(\mathbf{E}V|_{T'})|^2 d\sigma. \end{aligned}$$

We observe that for  $T' \subset S_T$ , there exists a pass from  $T$  to  $T'$  in  $S_T$  passing through a finite number of faces, bounded by the shape-regularity of  $\mathcal{T}$ ; see also [DK08, FV06]. In particular, there exist  $T_1, \dots, T_N \in \mathcal{T}$ , such that  $T \cap T_1 = \sigma_0, \dots, T_i \cap T_{i+1} = \sigma_i, \dots, T_N \cap T' = \sigma_N, \sigma_0, \dots, \sigma_N \in \mathcal{S}$ . We set  $T_0 := T$  and  $T_{N+1} := T'$ . Then, by the triangle inequality

$$(4.9) \quad \begin{aligned} |\mathbf{F}(\mathbf{E}V|_T) - \mathbf{F}(\mathbf{E}V|_{T'})| &\leq \sum_{i=0}^N |\mathbf{F}(\mathbf{E}V|_{T_i}) - \mathbf{F}(\mathbf{E}V|_{T_{i+1}})| \\ &\leq \sum_{\sigma \in \Sigma_T} \|[\mathbf{F}(\mathbf{E}V)]_\sigma\|, \end{aligned}$$

hence

$$\sum_{T' \subset S_T} |\mathbf{F}(\mathbf{E}V|_T) - \mathbf{F}(\mathbf{E}V|_{T'})|^2 \preccurlyeq \sum_{T' \subset S_T} \sum_{\sigma \in \Sigma_T} \|[\mathbf{F}(\mathbf{E}V)]_\sigma\|^2.$$

Note that the addends of the right-hand side are independent of  $T' \subset S_T$ . Observe further that the number of elements in  $S_T$  and hence the number of sides in  $\Sigma_T$  are bounded with respect to the shape-regularity of  $\mathcal{T}$ . Hence we obtain

$$(4.10) \quad \begin{aligned} \int_T (\phi_{|\mathbf{E}V|})^* (|\mathfrak{d} - \Pi_{\mathcal{T}}^{\mathbb{Q}} \mathfrak{d}|) dx &\preccurlyeq \sum_{T' \subset S_T} \int_{\partial T' \cap \Omega} h_T (\phi_{|\mathbf{E}V|})^* (|\llbracket \mathfrak{d} \rrbracket|) d\sigma \\ &+ \sum_{T' \subset S_T} \int_{\partial T' \cap \Omega} h_{T'} \|[\mathbf{F}(\mathbf{E}V)]\|^2 d\sigma. \end{aligned}$$

It remains to estimate the first term of the right-hand side of (4.10). For  $\sigma \in \mathcal{S}$ , let  $T_1, T_2 \in \mathcal{T}$  be its adjacent simplices i.e.,  $\sigma = T_1 \cap T_2$ . Applying the definition of  $\mathfrak{d} = \phi'_{|\mathbf{E}V|}(|\operatorname{div} V|) \frac{\operatorname{div} V}{|\operatorname{div} V|}$ , then Lemma 7 implies

$$(4.11) \quad \begin{aligned} \|[\mathfrak{d}]_\sigma\| &= \left| \phi'_{|\mathbf{E}V|}(|\operatorname{div} V|) \frac{\operatorname{div} V}{|\operatorname{div} V|} \Big|_{T_1} - \phi'_{|\mathbf{E}V|}(|\operatorname{div} V|) \frac{\operatorname{div} V}{|\operatorname{div} V|} \Big|_{T_2} \right| \\ &\preccurlyeq \left| \phi'_{|\mathbf{E}V|_{T_1}}(|\operatorname{div} V|) \frac{\operatorname{div} V}{|\operatorname{div} V|} \Big|_{T_1} - \phi'_{|\mathbf{E}V|_{T_1}}(|\operatorname{div} V|) \frac{\operatorname{div} V}{|\operatorname{div} V|} \Big|_{T_2} \right| \\ &+ \left| \phi'_{|\mathbf{E}V|_{T_1}}(\|[\mathbf{E}V]_\sigma\|) \right|. \end{aligned}$$

Now, we can estimate the first addend with the help of Proposition 5 (applied with  $\phi_{|\mathbf{E}V|}$  instead of  $\phi$ ) by

$$(4.12) \quad \begin{aligned} \left| \phi'_{|\mathbf{E}V|_{T_1}}(|\operatorname{div} V|) \frac{\operatorname{div} V}{|\operatorname{div} V|} \Big|_{T_1} - \phi'_{|\mathbf{E}V|_{T_1}}(|\operatorname{div} V|) \frac{\operatorname{div} V}{|\operatorname{div} V|} \Big|_{T_2} \right| \\ \approx \left( \phi'_{|\mathbf{E}V|_{T_1}} \right)_{|\operatorname{div} V|_{T_1}} (\|[\mathbf{E}V]_\sigma\|), \end{aligned}$$

where the constants hidden in  $\approx$  depend only on  $\Delta_2(\{\phi_{|\mathbf{E}V|_{T_1}}, (\phi_{|\mathbf{E}V|_{T_1}})^*\})$  and thus on  $\Delta_2(\{\phi, \phi^*\})$  i.e., on the constants in (2.1). We have as in (4.8) that

$$\left( \phi'_{|\mathbf{E}V|_{T_1}} \right)_{|\operatorname{div} V|_{T_1}}(t) = \phi'_{|\mathbf{E}V|_{T_1} + |\operatorname{div} V|_{T_1}}(t) \preccurlyeq \phi'_{|\mathbf{E}V|_{T_1}}(t)$$

for all  $t \geq 0$ , where the last inequality follows from (2.11). Applying this to (4.12) gives

$$\left| \phi'_{|\mathbf{E}V|_{T_1}}(|\operatorname{div} V|) \frac{\operatorname{div} V}{|\operatorname{div} V|} \Big|_{T_1} - \phi'_{|\mathbf{E}V|_{T_1}}(|\operatorname{div} V|) \frac{\operatorname{div} V}{|\operatorname{div} V|} \Big|_{T_2} \right| \preccurlyeq \phi'_{|\mathbf{E}V|_{T_1}}(|[\mathbf{E}V]_\sigma|).$$

Inserting in (4.11) implies

$$|[\mathfrak{d}]_\sigma| \preccurlyeq \phi'_{|\mathbf{E}V|_{T_1}}(|[\mathbf{E}V]_\sigma|).$$

Choosing  $T_1 = T'$  for every addend of the right-hand side of (4.10), we have

$$\begin{aligned} \int_T (\phi_{|\mathbf{E}V|})^* (|\mathfrak{d} - \Pi_{\mathcal{T}}^{\mathbb{Q}} \mathfrak{d}|) dx &\preccurlyeq \sum_{T' \subset S_T} \int_{\partial T'} h_T (\phi_{|\mathbf{E}V|})^* (\phi'_{|\mathbf{E}V|}(|[\mathbf{E}V]|)) d\sigma \\ &+ \sum_{T' \subset S_T} \int_{\partial T'} h_{T'} \|[\mathbf{F}(\mathbf{E}V)]\|^2 d\sigma. \end{aligned}$$

Now, (2.10) and Proposition 5 imply

$$\begin{aligned} \int_T (\phi_{|\mathbf{E}V|})^* (|\mathfrak{d} - \Pi_{\mathcal{T}}^{\mathbb{Q}} \mathfrak{d}|) dx &\preccurlyeq \sum_{T' \subset S_T} \int_{\partial T'} h_T \phi_{|\mathbf{E}V|}(|[\mathbf{E}V]|) d\sigma \\ &+ \sum_{T' \subset S_T} \int_{\partial T'} h_{T'} \|[\mathbf{F}(\mathbf{E}V)]\|^2 d\sigma \\ &\preccurlyeq \sum_{T' \subset S_T} \int_{\partial T'} h_{T'} \|[\mathbf{F}(\mathbf{E}V)]\|^2 d\sigma, \end{aligned}$$

where we additionally used that  $h_T \approx h_{T'}$  for all  $T' \subset S_T$  depending on the shape-regularity of  $\mathcal{T}$ . Summing over all  $T \in \mathcal{T}$ , we obtain the asserted estimate.  $\square$

The next corollary combines the above results in the particular case of the finite element approximation of the quasi-steepest descent direction. It estimates the error between  $\mathfrak{d}_Q$  and  $\mathfrak{D}_Q$  by the quantity that is controlled by the AFEM, namely by the estimator of the error between  $u_Q$  and  $U_Q$ .

**Corollary 22.** *Let  $\phi$  be an  $N$ -function such that (2.1) holds and let  $\mathcal{T}$  be a conforming triangulation of the domain  $\Omega \subset \mathbb{R}^d$ ,  $Q \in \mathbb{Q}(\mathcal{T})$ ,  $\epsilon > 0$ , and  $\theta \in (0, 1)$ . Then, with  $(U_Q, \mathcal{T}_*) := \text{AFEM}(f - \nabla Q, \mathcal{T}, \epsilon, \theta)$ , it holds that*

$$\int_{\Omega} (\phi_{|\mathbf{E}U_Q|})^* (|\mathfrak{d}_Q - \mathfrak{D}_Q|) \preccurlyeq \eta^2(U_Q, \mathcal{T}_*, f - \nabla Q) \leq \epsilon,$$

where  $\eta$  denotes the error estimator defined in (3.2).

*Proof.* We start with inequality (2.12) to obtain

$$\begin{aligned} (4.13) \quad \int_{\Omega} (\phi_{|\mathbf{E}U_Q|})^* (|\mathfrak{d}_Q - \mathfrak{D}_Q|) &\preccurlyeq \int_{\Omega} (\phi_{|\mathbf{E}U_Q|})^* \left( \left| \mathfrak{d}_Q - \phi'_{|\mathbf{E}U_Q|}(|\operatorname{div} U_Q|) \frac{\operatorname{div} U_Q}{|\operatorname{div} U_Q|} \right| \right) \\ &+ (\phi_{|\mathbf{E}U_Q|})^* \left( \left| \phi'_{|\mathbf{E}U_Q|}(|\operatorname{div} U_Q|) \frac{\operatorname{div} U_Q}{|\operatorname{div} U_Q|} - \mathfrak{D}_Q \right| \right) dx, \end{aligned}$$

where we used that the  $\Delta_2$ -constant of  $(\phi_{|\mathbf{E}U_Q|})^*$  depends only on  $\Delta_2(\{\phi, \phi^*\})$ ; see Lemma 4. Now, the first term can be estimated by Lemma 20, i.e., we have

$$\int_{\Omega} (\phi_{|\mathbf{E}U_Q|})^* \left( \left| \mathfrak{D}_Q - \phi'_{|\mathbf{E}U_Q|}(|\operatorname{div} U_Q|) \frac{\operatorname{div} U_Q}{|\operatorname{div} U_Q|} \right| \right) dx \preccurlyeq \|\mathbf{F}(\mathbf{E}u_Q) - \mathbf{F}(\mathbf{E}U_Q)\|_{L^2(\Omega)}^2.$$

This term can be further estimated via the upper bound (3.4). Furthermore, for the second addend in (4.13), by Lemma 21 we then have

$$\int_{\Omega} (\phi_{|\mathbf{E}U_Q|})^* \left( \left| \phi'_{|\mathbf{E}U_Q|}(|\operatorname{div} U_Q|) \frac{\operatorname{div} U_Q}{|\operatorname{div} U_Q|} - \mathfrak{D}_Q \right| \right) dx \preccurlyeq \sum_{T \in \mathcal{T}} \int_{\partial T} h_T \|\mathbf{F}(\mathbf{E}U_Q)\|^2 d\sigma.$$

Recalling (3.2), this is the jump residual and thus is bounded by  $\eta^2(U_Q, \mathcal{T}_*, f - \nabla Q)$ . This proves the assertion.  $\square$

## 5. CONVERGENT ADAPTIVE UZAWA ALGORITHM (AUA)

Thanks to the above results on the approximated steepest descent direction, we are now able to state the adaptive finite element algorithm for the stationary Stokes problem. We suppose that  $\phi$  is an N-function that satisfies assumption (2.1).

**Algorithm 23** (AUA). Let  $\mathcal{T}_0$  be a conforming initial triangulation of  $\Omega$  and let  $P_0 \in \mathbb{Q}(\mathcal{T}_0)$  be an initial guess for  $p \in \mathbb{Q}$ . Fix  $\theta, \rho \in (0, 1)$ , initial tolerance  $\tau > 0$  and step-size  $\mu > 0$  and let  $j = 0$ ;

(1) (APPROXIMATED DERIVATIVE)

$$(U_j, \mathcal{T}_{j+1}) := \text{AFEM}(f - \nabla P_j, \mathcal{T}_j, \rho^j \tau, \theta);$$

(2) (APPROXIMATED QUASI-STEEPEST DESCENT DIRECTION)

$$\mathfrak{D}_j := \Pi_{\mathcal{T}_{j+1}}^{\mathbb{Q}} \phi'_{|\mathbf{E}U_j|}(|\operatorname{div} U_j|) \frac{\operatorname{div} U_j}{|\operatorname{div} U_j|};$$

(3) (UPDATE)

$$P_{j+1} := P_j + \mu \mathfrak{D}_j;$$

increment  $j$  and go to step (1);

We observe that by means of the procedure **REFINE** in the AFEM the sequence of triangulations  $\{\mathcal{T}_j\}_{j \in \mathbb{N}}$  produced by AUA is shape regular depending on the shape regularity of  $\mathcal{T}_0$ . The next theorem is the main result of this work. It states the convergence of the AUA for some fixed step-size  $\mu$ .

**Theorem 24.** *Let  $\phi$  be an N-function such that (2.1) holds. Then there exists  $\mu_0 > 0$  depending only on  $\Delta_2(\{\phi, \phi^*\})$  and  $d$ , such that for all step-sizes  $\mu \in (0, \mu_0)$ , it holds for the sequence  $(P_j)_{j \in \mathbb{N}} \subset \mathbb{Q}$  produced by Algorithm 23 (AUA) that*

$$P_j \rightarrow p \quad \text{in } \mathbb{Q}, \quad \text{as } j \rightarrow \infty,$$

where  $p \in \mathbb{Q}$  is the pressure of the nonlinear Stokes problem (2.4).

In order to prove Theorem 24 we need to know what it means to  $(P_j)_{j \in \mathbb{N}} \subset \mathbb{Q}$  if the sequence of derivatives  $(D\mathcal{F}(P_j)) \subset L_0^\phi(\Omega)$  vanishes.

**Lemma 25.** *Assume the conditions of Theorem 24. Let  $(u_j)_{j \in \mathbb{N}} \subset \mathbb{V}$  be the sequence defined by  $u_j := u_{P_j}$  as in (2.26). Then*

$$\operatorname{div} u_j \rightarrow_{j \rightarrow \infty} 0 \quad \text{in } L_0^\phi(\Omega) \quad \text{implies} \quad P_j \rightarrow_{j \rightarrow \infty} p \quad \text{in } \mathbb{Q},$$

where  $p$  is the unique minimizer of  $\mathcal{F}$ .

*Proof.* We assume the contrary. In particular, w.l.o.g., there exists a constant  $c > 0$  such that  $\|p - P_j\|_{\mathbb{Q}} > c$ ; otherwise we pass to a sub-sequence. By the inf-sup condition (2.8) and Korn's inequality (2.7) there exists a  $\tilde{\beta} > 0$  such that

$$\begin{aligned} \tilde{\beta} \|p - P_j\|_{\mathbb{Q}} &\leq \sup_{v \in W_0^{1,\phi}(\Omega)} \frac{\int_{\Omega} (p - P_j) \operatorname{div} v \, dx}{\|\mathbf{E}v\|_{(\phi)}} \\ &= \sup_{v \in W_0^{1,\phi}(\Omega)} \frac{\int_{\Omega} (\mathbf{A}(\mathbf{E}u) - \mathbf{A}(\mathbf{E}u_j)) : \mathbf{E}v \, dx}{\|\mathbf{E}v\|_{(\phi)}}, \end{aligned}$$

where we applied (2.25). By means of Young's inequality, it follows that for  $\delta > 0$ ,

$$\begin{aligned} \tilde{\beta} \|p - P_j\|_{\mathbb{Q}} &\leq C_\delta \int_{\Omega} (\phi_{|\mathbf{E}u|})^* (|\mathbf{A}(\mathbf{E}u) - \mathbf{A}(\mathbf{E}u_j)|) \, dx \\ &\quad + \delta \sup_{v \in W_0^{1,\phi}(\Omega)} \int_{\Omega} \phi_{|\mathbf{E}u|} \left( \frac{|\mathbf{E}v|}{\|\mathbf{E}v\|_{(\phi)}} \right) \, dx, \end{aligned}$$

where the constant  $C_\delta$  depends on  $\delta$  and  $\Delta_2(\{\phi_a\}_{a \geq 0})$ , and thus on  $\Delta_2(\{\phi, \phi^*\})$ ; see Lemma 4. The second term is bounded according to

$$\int_{\Omega} \phi_{|\mathbf{E}u|} \left( \frac{|\mathbf{E}v|}{\|\mathbf{E}v\|_{(\phi)}} \right) \, dx \leq \int_{\Omega} \phi \left( \frac{|\mathbf{E}v|}{\|\mathbf{E}v\|_{(\phi)}} \right) + \phi(|\mathbf{E}u|) \, dx \leq 1 + \int_{\Omega} \phi(|\mathbf{E}u|) \, dx.$$

Hence, for  $\delta$  small enough, we have by the assumption  $0 < c < \|p - P_j\|_{\mathbb{Q}}$  that

$$\tilde{\beta} \|p - P_j\|_{\mathbb{Q}} \leq \int_{\Omega} (\phi_{|\mathbf{E}u|})^* (|\mathbf{A}(\mathbf{E}u) - \mathbf{A}(\mathbf{E}u_j)|) \, dx.$$

Furthermore, Proposition 5, (2.10), and a Hölder inequality (see [RR91]) imply

$$\begin{aligned} \tilde{\beta} \|p - P_j\|_{\mathbb{Q}} &\leq \int_{\Omega} (\mathbf{A}(\mathbf{E}u) - \mathbf{A}(\mathbf{E}u_j)) : (\mathbf{E}u - \mathbf{E}u_j) \, dx \\ &= \int_{\Omega} (p - P_j) \operatorname{div}(u - u_j) \, dx = C \int_{\Omega} (p - P_j) \operatorname{div} u_j \, dx \\ &\leq 2 \|p - P_j\|_{\mathbb{Q}} \|\operatorname{div} u_j\|_{(\phi)}, \end{aligned}$$

where we used (2.25) and the fact that  $\operatorname{div} u = 0$ ; see (2.4). Since  $\|\operatorname{div} u_j\|_{(\phi)} \rightarrow 0$  as  $j \rightarrow \infty$ , this is a contradiction and hence  $P_j \rightarrow p$  in  $\mathbb{Q}$  as  $j \rightarrow \infty$ .  $\square$

*Proof of Theorem 24.* For convenience, we use the abbreviations

$$\mathfrak{d}_j := \mathfrak{d}_{P_j} = -\phi'_{|\mathbf{E}u_j|} (|\operatorname{div} u_j|) \frac{\operatorname{div} u_j}{|\operatorname{div} u_j|} \quad \text{and} \quad u_j := u_{P_j};$$

see also (2.25). For  $P_j \in \mathbb{Q}(\mathcal{T}_j)$ ,  $j \in \mathbb{N}$ , let

$$\mathcal{H}_j(\mu) := \mathcal{F}(P_j) - \mathcal{F}(P_j + \mu \mathfrak{D}_j).$$

By means of the mean-value theorem and Proposition 15, for  $\mu > 0$ , there exists  $\zeta \in (0, \mu)$ , such that

$$\begin{aligned}
(5.1) \quad \mathcal{H}_j(\mu) &= \mu \mathcal{H}'_j(\zeta) = -\mu \langle D\mathcal{F}(P_j + \zeta \mathfrak{D}_j), \mathfrak{D}_j \rangle \\
&= -\mu \langle D\mathcal{F}(P_j), \mathfrak{d}_j \rangle + \mu \langle D\mathcal{F}(P_j), \mathfrak{d}_j - \mathfrak{D}_j \rangle \\
&\quad - \frac{\mu}{\zeta} \langle D\mathcal{F}(P_j + \zeta \mathfrak{D}_j) - D\mathcal{F}(P_j), \zeta \mathfrak{D}_j \rangle.
\end{aligned}$$

We handle the terms at the right hand side separately. First, we have from (2.9b)

$$(5.2) \quad -\langle D\mathcal{F}(P_j), \mathfrak{d}_j \rangle = \int_{\Omega} \phi'_{|\mathbf{E}u_j|}(|\operatorname{div} u_j|) |\operatorname{div} u_j| \, dx \geq \int_{\Omega} \phi_{|\mathbf{E}u_j|}(|\operatorname{div} u_j|) \, dx.$$

The next term can be estimated with the help of Young's inequality (2.15). For  $\delta > 0$  it holds that

$$\begin{aligned}
|\langle D\mathcal{F}(P_j), \mathfrak{d}_j - \mathfrak{D}_j \rangle| &\leq \int_{\Omega} |(\mathfrak{d}_j - \mathfrak{D}_j) \operatorname{div} u_j| \, dx \\
&\leq \delta \int_{\Omega} \phi_{|\mathbf{E}u_j|}(|\operatorname{div} u_j|) \, dx + C_{\delta} \int_{\Omega} (\phi_{|\mathbf{E}u_j|})^*(|\mathfrak{d}_j - \mathfrak{D}_j|) \, dx.
\end{aligned}$$

The constant  $C_{\delta}$  depends only on  $\Delta_2(\{\phi_a\}_{a \geq 0})$  and thus on  $\Delta_2(\{\phi, \phi^*\})$ ; see Lemma 4. Now, applying Corollary 22, there exists a constant  $\hat{C} > 0$  depending only on  $\Delta_2(\{\phi, \phi^*\})$  and  $d$ , such that

$$(5.3) \quad |\langle D\mathcal{F}(P_j), \mathfrak{d}_j - \mathfrak{D}_j \rangle| \leq \delta \int_{\Omega} \phi_{|\mathbf{E}u_j|}(|\operatorname{div} u_j|) \, dx + C_{\delta} \hat{C} \eta^2(U_j, \mathcal{T}_j, f - \nabla P_j).$$

Taking  $u_{\zeta} := u_{P_j + \zeta \mathfrak{D}_j}$ , we have for the last term in (5.1) by Proposition 15, (2.25), and Young's inequality for  $\delta > 0$  that

$$\begin{aligned}
\langle D\mathcal{F}(P_j + \zeta \mathfrak{D}_j) - D\mathcal{F}(P_j), \zeta \mathfrak{D}_j \rangle \\
\leq \int_{\Omega} \delta \phi_{|\mathbf{E}u_j|}(|\operatorname{div}(u_{\zeta} - u_j)|) + C_{\delta} (\phi_{|\mathbf{E}u_j|})^*(|\zeta \mathfrak{D}_j|) \, dx.
\end{aligned}$$

On the other hand, we have by Corollary 6 that

$$\begin{aligned}
\langle D\mathcal{F}(P_j + \zeta \mathfrak{D}_j) - D\mathcal{F}(P_j), \zeta \mathfrak{D}_j \rangle &= \int_{\Omega} (\mathbf{A}(\mathbf{E}u_{\zeta}) - \mathbf{A}(\mathbf{E}u_j)) : (\mathbf{E}u_{\zeta} - \mathbf{E}u_j) \, dx \\
&\approx \int_{\Omega} \phi_{|\mathbf{E}u_j|}(|\mathbf{E}u_{\zeta} - \mathbf{E}u_j|) \, dx.
\end{aligned}$$

Hence, recalling that  $|\operatorname{div}(u_{\zeta} - u_j)| \leq \sqrt{d} |\mathbf{E}u_{\zeta} - \mathbf{E}u_j|$  yields for  $\delta > 0$  small enough

$$\langle D\mathcal{F}(P_j + \zeta \mathfrak{D}_j) - D\mathcal{F}(P_j), \zeta \mathfrak{D}_j \rangle \preceq \int_{\Omega} (\phi_{|\mathbf{E}u_j|})^*(|\zeta \mathfrak{D}_j|) \, dx.$$

We assume for the ease of simplicity that  $\mu_0 \leq 2$ . Applying (2.12) as well as (2.11), we obtain

$$\begin{aligned}
(5.4) \quad \langle D\mathcal{F}(P_j + \zeta \mathfrak{D}_j) - D\mathcal{F}(P_j), \zeta \mathfrak{D}_j \rangle \\
\preceq \int_{\Omega} (\phi_{|\mathbf{E}u_j|})^*(|\zeta \mathfrak{d}_j|) \, dx + \int_{\Omega} (\phi_{|\mathbf{E}u_j|})^*(|\zeta (\mathfrak{d}_j - \mathfrak{D}_j)|) \, dx \\
\preceq \int_{\Omega} (\phi_{|\mathbf{E}u_j|})^*(|\zeta \mathfrak{d}_j|) \, dx + \int_{\Omega} (\phi_{|\mathbf{E}u_j|})^*(|\mathfrak{d}_j - \mathfrak{D}_j|) \, dx.
\end{aligned}$$

We have for all  $\zeta \in (0, \mu_0)$  by  $|\operatorname{div} v| \leq \sqrt{d} |\mathbf{E}v|$ ,  $v \in \mathbb{V}$ , the definition of shifted N-functions, and the monotonicity of  $\phi'$ , that

$$\begin{aligned} |\zeta \mathfrak{D}_j| \leq 2 |\mathfrak{D}_j| &= 2 \left| \phi'_{|\mathbf{E}u_j|}(|\operatorname{div} u_j|) \right| = 2 \left| \frac{\phi'(|\operatorname{div} u_j| + |\mathbf{E}u_j|)}{|\operatorname{div} u_j| + |\mathbf{E}u_j|} |\operatorname{div} u_j| \right| \\ &\preccurlyeq \left| \frac{\phi'(|\mathbf{E}u_j|)}{|\mathbf{E}u_j|} |\mathbf{E}u_j| \right| \end{aligned}$$

i.e.,  $|\zeta \mathfrak{D}_j| \preccurlyeq \phi'(|\mathbf{E}u_j|)$ . Thus by (2.10), (2.17), and Proposition 5, for all  $\zeta \in (0, \mu_0)$  we have

$$\begin{aligned} (\phi_{|\mathbf{E}u_j|})^*(|\zeta \mathfrak{D}_j|) &\approx (\phi_{\phi'(|\mathbf{E}u_j|)})^*(|\zeta \mathfrak{D}_j|) \approx (\phi^*)''(\phi'(|\mathbf{E}u_j|) + |\zeta \mathfrak{D}_j|) |\zeta \mathfrak{D}_j|^2 \\ &\approx (\phi^*)''(\phi'(|\mathbf{E}u_j|) + |\mathfrak{D}_j|) |\zeta \mathfrak{D}_j|^2 \\ &\approx \frac{(\phi^*)''(\phi'(|\mathbf{E}u_j|) + |\mathfrak{D}_j|)}{\phi'(|\mathbf{E}u_j|) + |\mathfrak{D}_j|} (\phi'(|\mathbf{E}u_j|) + |\mathfrak{D}_j|) |\zeta \mathfrak{D}_j|^2 \\ &\approx \zeta^2 \frac{(\phi^*)'(\phi'(|\mathbf{E}u_j|) + |\mathfrak{D}_j|)}{\phi'(|\mathbf{E}u_j|) + |\mathfrak{D}_j|} |\mathfrak{D}_j|^2 = \zeta^2 (\phi^*)'_{\phi'(|\mathbf{E}u_j|)}(|\mathfrak{D}_j|) |\mathfrak{D}_j| \\ &\approx \zeta^2 (\phi_{|\mathbf{E}u_j|})^*(|\mathfrak{D}_j|) \approx \zeta^2 \phi_{|\mathbf{E}u_j|}(|\operatorname{div} u_j|). \end{aligned}$$

Applying this to the first addend of (5.4) and changing the shift of the second addend with the help of Lemma 7 to  $|\mathbf{E}U_j|$  yields

$$\begin{aligned} \langle D\mathcal{F}(P_j + \zeta \mathfrak{D}_j) - D\mathcal{F}(P_j), \zeta \mathfrak{D}_j \rangle &\preccurlyeq \int_{\Omega} \zeta^2 \phi_{|\mathbf{E}u_j|}(|\operatorname{div} u_j|) dx \\ &\quad + \int_{\Omega} (\phi_{|\mathbf{E}U_j|})^*(|\mathfrak{D}_j - \mathfrak{D}_j|) dx + \|\mathbf{F}(\mathbf{E}u_j) - \mathbf{F}(\mathbf{E}U_j)\|_{L^2(\Omega)}^2. \end{aligned}$$

Now, applying Corollary 22 as well as the upper bound (3.4), there exists a constant  $\tilde{C} > 0$  solely depending on  $\Delta_2(\{\phi, \phi^*\})$  and  $d$ , such that

$$\begin{aligned} \langle D\mathcal{F}(P_j + \zeta \mathfrak{D}_j) - D\mathcal{F}(P_j), \zeta \mathfrak{D}_j \rangle &\leq \tilde{C} \int_{\Omega} \zeta^2 \phi_{|\mathbf{E}u_j|}(|\operatorname{div} u_j|) dx \\ &\quad + \tilde{C} \eta^2(U_j, \mathcal{T}_j, f - \nabla P_j). \end{aligned}$$

This, (5.2), and (5.3) applied to (5.1) yields

$$\begin{aligned} \mathcal{H}_j(\mu) &= \mathcal{F}(P_j) - \mathcal{F}(P_j + \mu \mathfrak{D}_j) \\ &\geq \mu \int_{\Omega} \phi_{|\mathbf{E}u_j|}(|\operatorname{div} u_j|) dx \\ &\quad - \mu \left\{ \delta \int_{\Omega} \phi_{|\mathbf{E}u_j|}(|\operatorname{div} u_j|) dx + C_\delta \hat{C} \eta^2(U_j, \mathcal{T}_j, f - \nabla P_j) \right\} \\ &\quad - \frac{\mu}{\zeta} \left\{ \tilde{C} \int_{\Omega} \zeta^2 \phi_{|\mathbf{E}U_j|}(|\operatorname{div} \mathfrak{D}_j|) dx - \tilde{C} \eta^2(U_j, \mathcal{T}_j, f - \nabla P_j) \right\} \\ &= \mu(1 - \delta - \tilde{C} \zeta) \int_{\Omega} \phi_{|\mathbf{E}u_j|}(|\operatorname{div} u_j|) dx \\ &\quad - (\mu C_\delta \hat{C} + \frac{\mu}{\zeta} \tilde{C}) \eta^2(U_j, \mathcal{T}_j, f - \nabla P_j). \end{aligned}$$

Recall that  $\zeta \leq \mu$ , hence

$$\mathcal{H}_j(\mu) \geq \mu(1 - \delta - \tilde{C}\mu) \int_{\Omega} \phi_{|\mathbf{E}u_j|}(|\operatorname{div} u_j|) dx - (\mu C_\delta \hat{C} + \tilde{C}) \eta^2(U_j, \mathcal{T}_j, f - \nabla P_j).$$

Observe that for  $\mu_0 \in (0, 1/\tilde{C})$ ,  $\delta := (1 - \tilde{C}\mu)/2 > 0$ , we have for all  $\mu \in (0, \mu_0)$  that  $c_\mu := \mu(1 - \delta - \tilde{C}\mu) > 0$ . Take  $C_\mu := (\mu C_\delta \hat{C} + \tilde{C})$ , then

$$(5.5) \quad \mathcal{H}_j(\mu) \geq c_\mu \int_{\Omega} \phi_{|\mathbf{E}u_j|}(|\operatorname{div} u_j|) dx - C_\mu \eta^2(U_j, \mathcal{T}_j, f - \nabla P_j).$$

The constants  $c_\mu, C_\mu > 0$  depend only on  $\Delta_2(\{\phi, \phi^*\})$ , the step-size  $\mu$  and  $d$ . Note that due to Algorithm 23 (AUA)—step 1 (APPROXIMATED DERIVATIVE)—it holds that

$$(5.6) \quad \eta(U_j, \mathcal{T}_j, f - \nabla P_j) \leq \rho^j \tau.$$

Therefore, we have

$$\mathcal{H}_j(\mu) = \mathcal{F}(P_j) - \mathcal{F}(P_j + \mu \mathfrak{D}_j) \geq c_\mu \int_{\Omega} \phi_{|\mathbf{E}u_j|}(|\operatorname{div} u_j|) dx - C_\mu \rho^{2j} \tau^2.$$

Recalling that  $P_{j+1} = P_j + \mu \mathfrak{D}_j$ , we have for all  $J \in \mathbb{N}$  the telescopic sum

$$\begin{aligned} \mathcal{F}(P_0) - \mathcal{F}(P_J) &= \sum_{j=0}^{J-1} \mathcal{F}(P_j) - \mathcal{F}(P_{j+1}) \\ &\geq c_\mu \sum_{j=0}^{J-1} \int_{\Omega} \phi_{|\mathbf{E}u_j|}(|\operatorname{div} u_j|) dx - C_\mu \sum_{j=0}^{J-1} \rho^{2j} \tau^2. \end{aligned}$$

The last term can be estimated by a geometric series and thus by  $\tau^2/(1 - \rho^2)$ . On the other hand, we can estimate  $\mathcal{F}(P_0) - \mathcal{F}(p) \geq \mathcal{F}(P_0) - \mathcal{F}(P_J)$ , since  $p \in \mathbb{Q}$  is the minimizer of  $\mathcal{F}$ . Therefore,

$$(5.7) \quad \begin{aligned} \mathcal{F}(P_0) - \mathcal{F}(p) &\geq \mathcal{F}(P_0) - \mathcal{F}(P_J) \\ &\geq c_\mu \sum_{j=0}^{J-1} \int_{\Omega} \phi_{|\mathbf{E}u_j|}(|\operatorname{div} u_j|) dx - C_\mu \frac{\tau}{1 - \rho^2} \end{aligned}$$

for all  $J \in \mathbb{N}$ . In other words, the series  $\sum_{j=0}^{J-1} \int_{\Omega} \phi_{|\mathbf{E}u_j|}(|\operatorname{div} u_j|) dx$  is bounded independent on  $J$ . Since all its addends are positive, we get that

$$\int_{\Omega} \phi_{|\mathbf{E}u_j|}(|\operatorname{div} u_j|) dx \rightarrow 0, \quad \text{as } j \rightarrow \infty.$$

It remains to show that this implies  $\operatorname{div} u_j \rightarrow 0$  in  $\mathbb{Q}$  as  $j \rightarrow \infty$ . Then, the assertion follows by Lemma 25. In particular, we obtain by (5.7) that

$$\mathcal{F}(P_0) + C_\mu \frac{\tau}{1 - \rho^2} \geq \mathcal{F}(P_j)$$

for all  $j \in \mathbb{N}$  i.e.,  $(\mathcal{F}(P_j))_{j \in \mathbb{N}}$  is bounded. Combining (2.26) with (2.25) gives

$$\begin{aligned} \mathcal{F}(P_0) + C_\mu \frac{\tau}{1 - \rho^2} &\geq \mathcal{F}(P_j) = -\mathcal{L}(u_j, P_j) \\ &= \int_{\Omega} -\phi(|\mathbf{E}u_j|) + P_j \operatorname{div} u_j + f u_j dx = \int_{\Omega} -\phi(|\mathbf{E}u_j|) + \mathbf{A}(\mathbf{E}u_j) : \mathbf{E}u_j dx \\ &= \int_{\Omega} -\phi(|\mathbf{E}u_j|) + \phi'(|\mathbf{E}u_j|) |\mathbf{E}u_j| dx \geq (\nabla_2(\phi) - 1) \int_{\Omega} \phi(|\mathbf{E}u_j|) dx \geq 0, \end{aligned}$$

where the constant  $\nabla_2(\phi) > 1$  depends only on  $\Delta_2(\phi^*)$ ; see (2.13). Therefore, the sequence  $(\int_{\Omega} \phi(|\mathbf{E}u_j|) dx)_{j \in \mathbb{N}} \subset \mathbb{R}$  is bounded. Assume that  $(\operatorname{div} u_j)_{j \in \mathbb{N}}$  does not converge to zero in  $\mathbb{Q}$ . Then, w.l.o.g. there exists  $c > 0$  such that

$$0 < c < \int_{\Omega} \phi(|\operatorname{div} u_j|) dx \quad \text{for all } j \in \mathbb{N};$$

otherwise we pass to a sub-sequence. Hence, we get by (2.19) for  $\delta > 0$ ,

$$c < \int_{\Omega} \phi(|\operatorname{div} u_j|) dx \preceq (1 + C_{\delta}) \int_{\Omega} \phi_{|\mathbf{E}u_j|}(|\operatorname{div} u_j|) dx + \delta \int_{\Omega} \phi(|\mathbf{E}u_j|) dx$$

for all  $j \in \mathbb{N}$ . Since  $(\int_{\Omega} \phi(|\mathbf{E}u_j|) dx)_{j \in \mathbb{N}}$  is bounded, we can choose  $\delta > 0$  small enough to obtain

$$0 < c \preceq C \int_{\Omega} \phi_{|\mathbf{E}u_j|}(|\operatorname{div} u_j|) dx,$$

with a constant  $C > 0$  not depending on  $j \in \mathbb{N}$ . This is a contradiction. Thus,  $\operatorname{div} u_j \rightarrow 0$  in  $\mathbb{Q}$ , as  $j \rightarrow \infty$  and the assertion follows with Lemma 25.  $\square$

The next corollary states that the convergence of the pressure implies the convergence of the velocity.

**Corollary 26.** *Assume the conditions of Theorem 24. Let  $(U_j)_{j \in \mathbb{N}} \subset \mathbb{V}$  be the sequence produced by the AUA. Then*

$$U_j \rightarrow u \quad \text{in } \mathbb{V} \text{ as } j \rightarrow \infty,$$

where  $u$  denotes the velocity of the nonlinear Stokes problem (2.4).

*Proof.* We use the notation of the proof of Theorem 24. Then by Corollary 6 we have

$$\begin{aligned} \|\mathbf{F}(\mathbf{E}u) - \mathbf{F}(\mathbf{E}u_j)\|_{L^2(\Omega)}^2 &\approx \int_{\Omega} (\mathbf{A}(\mathbf{E}u) - \mathbf{A}(\mathbf{E}u_j)) : (\mathbf{E}u - \mathbf{E}u_j) dx \\ &= - \int_{\Omega} (p - P_j) \operatorname{div} u_j dx, \end{aligned}$$

where we used that  $\operatorname{div} u = 0$ . In the proof of Theorem 24 we have shown that  $\operatorname{div} u_j \rightarrow 0$  in  $L_0^{\phi}(\Omega)$  and  $P_j \rightarrow p$  in  $\mathbb{Q}$  as  $j \rightarrow \infty$ . This and observation (5.6) in combination with (3.4) prove the assertion.  $\square$

We finish with some remarks on the AUA.

*Remark 27.* It turns out that Theorem 24 is a generalization of the convergence of the Uzawa algorithm in the linear case in [NP04, BMN02]. There it is proved, that for  $\mu \in (0, 2)$  the Uzawa algorithm converges linearly in terms of the outer iteration counter, which is a consequence of the inf-sup condition (2.8). In the nonlinear case a similar result can be proved provided that there holds the following inf-sup condition related to quasi-norm expressions: there exists  $\beta > 0$  such that for all  $q \in \mathbb{Q}$ ,

$$(5.8) \quad \begin{aligned} \|\nabla(q - p)\|_{(\mathbf{E}u),*}^2 &:= \sup_{v \in \mathbb{V}} \left\{ \int_{\Omega} (q - p) \operatorname{div} v dx - \int_{\Omega} \phi_{|\mathbf{E}u|}(|\mathbf{E}v|) dx \right\} \\ &\geq \beta \inf_{c \in \mathbb{R}} \int_{\Omega} (\phi_{|\mathbf{E}u|})^*(|q - p - c|) dx. \end{aligned}$$

Then, for sufficiently small step-size  $\mu$ , the AUA produces a sequence  $(U_j, P_j)_j$  of approximations to the solution  $(u, p)$  of (2.4), such that

$$\|\mathbf{F}(\mathbf{E}u) - \mathbf{F}(\mathbf{E}U_j)\|_{L^2(\Omega)}^2 \preccurlyeq \lambda^j, \quad j \geq 0,$$

for some  $\lambda \in (0, 1)$ ; compare with [Kre08, Remark 141] and [BMN02]. Moreover, the equivalence of error quantities

$$\int_{\Omega} \phi_{|\mathbf{E}u|}(|\operatorname{div} u_q|) dx \approx \|\mathbf{F}(\mathbf{E}u) - \mathbf{F}(\mathbf{E}u_q)\|_{L^2(\Omega)}^2 \approx \inf_{c \in \mathbb{R}} \int_{\Omega} (\phi_{|\mathbf{E}u|})^*(|p - q - c|) dx$$

for all  $q \in \mathbb{Q}$  iff (5.8) holds; for more details see [Kre08, Remarks 141-143].

We further remark, that linear convergence is crucial to prove quasi optimal complexity for a modification of the AUA as in [KS08]. Unfortunately, to our best knowledge, a generalized inf-sup condition (5.8) is not known.

*Remark 28* (Pressure with mean-value zero). For the reason of numerical cancellations it may be convenient to try to avoid extreme values of  $P_j$ . For this purpose one may consider functions with mean-value zero since the pressure is only determined up to a constant value. Hence, starting Algorithm 23 (AUA) with an initial guess  $P_0 \in \mathbb{Q}(\mathcal{T}_0)$ , which has mean-value zero we can substitute step 3 (UPDATE) of (AUA) by

3'. (UPDATE')

$$P_{j+1} := P_j + \mu \mathfrak{D}_j - \frac{1}{|\Omega|} \int_{\Omega} \mu \mathfrak{D}_j dx;$$

increment  $j$  and go to step (1).

Therefore, by induction  $(P_j)_{j \in \mathbb{N}} \subset L_0^{\phi^*}(\Omega)$ . Note that the modifications do not affect the theoretical behavior of (AUA), since the pressure is only defined up to a constant;  $\mathbb{Q} = L^{\phi^*}(\Omega)/\mathbb{R}$ . Hence, subtracting the mean-value has no theoretical effect.

*Remark 29* (Coarsening). Since the right-hand side  $f - \nabla P_j$  of (3.1) in the (AUA) is changing in each iteration, it might be reasonable to apply a coarsening step in order to obtain optimal meshes. Recall, that for the proof of the convergence of AUA we only used that  $\eta(U_j, \mathcal{T}_j, f - \nabla P_j) \leq \rho^j \tau$ . In fact, the AFEM can be substituted by any procedure that approximates  $u_{P_j}$  such that the estimator is up to this accuracy. Hence, it is possible to apply a coarsening routine, e.g., after step (3) (UPDATE) of the AUA. Note, that  $P_j$  is defined on the common refinement of all triangulations  $\mathcal{T}_i$   $i = 1, \dots, k$ . Therefore, it may be necessary to handle two grids, namely one grid for calculating  $U_j$  in step (1) and then the common refinement of all triangulations  $\mathcal{T}_i$ ,  $i = 1, \dots, k$ , in order to store  $P_j$ .

*Remark 30* (Stopping criterion). Finding a stopping criterion for Algorithm 23 (AUA) for an adequate distance quantity turns out to be no easy task. In fact, proving reasonable a posteriori estimates usually requires a continuous inf-sup condition; see [AO00, Section 9.2]. To have a reasonable estimator for a quasi-norm error notion, we need a inf-sup condition, which is somehow related to the quasi-norm; see (27). Since such a condition is not available so far, we have to settle for nonoptimal estimates as in [BRS04, BS08]; compare also [Kre08].

*Remark 31.* Note that the spaces  $\mathbb{V}(\mathcal{T})$  and  $\mathbb{Q}(\mathcal{T})$  are not stable in the sense that they satisfy a discrete inf-sup condition

$$\inf_{Q \in \mathbb{Q}(\mathcal{T})} \sup_{V \in \mathbb{V}(\mathcal{T})} \frac{\int_{\Omega} Q \operatorname{div} v \, dx}{\|Q\|_{\mathbb{Q}} \|V\|_{\mathbb{V}}} \geq \beta_{\mathcal{T}} > 0,$$

with  $\beta_{\mathcal{T}}$  independent of the triangulation  $\mathcal{T}$ ; for pairs of stable function spaces; cf. e.g. [BL93b, BF91, GR86, For81]. However, Algorithm 23 (AUA) is a generalized inexact Uzawa iteration at an infinite dimensional level. The convergence of our algorithm does not require a discrete inf-sup condition but rather the continuous inf-sup condition (2.8).

## 6. NUMERICAL EXPERIMENTS

We conclude this article with numerical experiments. We want to focus on two main aspects. From an analytical point of view, it is interesting to see if there holds a generalization (5.8) of the inf-sup condition related to quasi-norm expressions. For this reason we study the three quantities,

$$\begin{aligned} \mathcal{E}^2(\operatorname{div} U_j) &:= \int_{\Omega} \phi_{|\mathbf{E}u|} (|\operatorname{div} U_j|) \, dx, \\ \mathcal{E}^2(U_j) &:= \|\mathbf{F}(\mathbf{E}u) - \mathbf{F}(\mathbf{E}U_j)\|_{L^2(\Omega)}^2, \\ \mathcal{E}^2(P_j) &:= \inf_{c \in \mathbb{R}} \int_{\Omega} (\phi_{|\mathbf{E}u|})^* (|p - P_j - c|) \, dx, \end{aligned} \tag{6.1}$$

relative to the number  $j$  of outer iterations of the AUA; compare with Remark 27.

On the other hand, we are interested in the performance of the AUA, i.e., the error decay with respect to the degrees of freedom (DOFs). Note that the AUA is only proved to convergence without a rate; see Theorem 24. Nevertheless, thanks to piecewise linear ansatz functions for the velocity, we expect a relation of the form

$$\mathcal{E}(\operatorname{div} U_j) \approx \mathcal{E}(U_j) \approx \mathcal{E}(P_j) \approx \operatorname{DOF}_j^{-1/d}. \tag{6.2}$$

In all examples the nonlinear vector-field of the Stokes problem (1.1) is given by

$$\mathbf{A}(\mathbf{E}) := (\kappa + |\mathbf{E}|^2)^{\frac{r-2}{2}} \mathbf{E}, \quad \mathbf{E} \in \mathbb{R}^{d \times d},$$

for some  $r \in (1, \infty)$  and  $\kappa = 10^{-6}$ . The N-function is then given by  $\phi(t) := \frac{1}{r}(\kappa + t^2)^{r/2} - \frac{1}{r}\kappa^{r/2}$ . We consider three experiments in two dimensions and one in three dimensions. In particular, we consider two smooth experiments in two dimensions with  $r = 3$  and  $r = 1.5$ , one singular experiment in two dimensions with  $r = 3$ , and a regular experiment in three dimensions. For the computations we used the the finite element toolbox ALBERTA [SS05]. Some pictures (Figures 4 and 6) were produced by the graphics package ParaView [Squ08].

We run all experiments with Dörfler marking parameter  $\theta = \frac{1}{3}$ , reduction factor  $\rho = 0.97$  for the elliptic errors and initial guess  $P_0 \equiv 0$ . As in Remark 28 we update the pressure, such that its zero mean-value is preserved.

It is convenient to choose different initial tolerances  $\tau$  for different problems. In fact, if  $\tau$  is chosen too small, many refinements of the initial triangulation are necessary until the AFEM reaches the tolerance and the pressure is updated the first time. This costs a lot of computational power and makes it hard to get the asymptotic behavior of the error. On the other hand, choosing  $\tau$  we get large results on several pressure updates on the same grid. This somehow corresponds

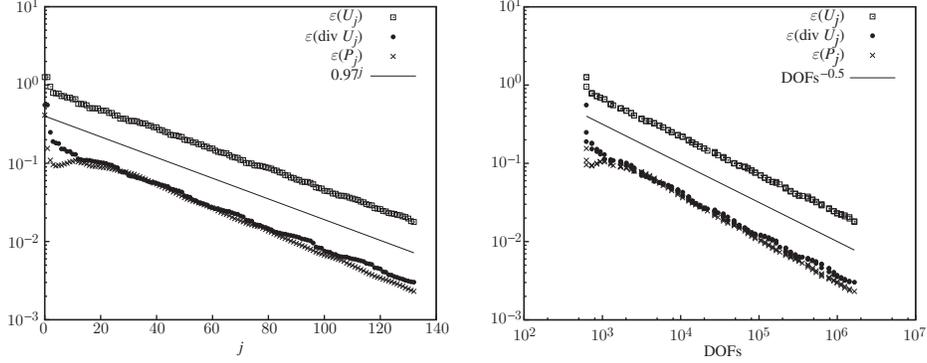


FIGURE 1. Example 6.1 with  $r = 3$ : All error quantities show linear convergence with respect to iterations  $j$ . The reduction factor is about  $\rho = 0.97$  (left). They show the same rate with respect to DOFs as the solid line having the slope  $1/d = 1/2$  (right).

to starting the AUA with less tolerance and a different initial guess on the initial grid. It turns out that  $\tau = 3$  for the two-dimensional examples and  $\tau = 30$  for the three-dimensional one are good choices.

The choice of the step-size  $\mu$  is even more delicate. In contrast to [BMN02, NP04] for the linear Stokes problem, we could not prove an optimal value nor a range for  $\mu$ , which ensures convergence of the AUA. However, computations suggest that  $\mu = 0.4$  and  $\mu = 1.1$  are good choices for  $r = 3$  and  $r = 1.5$ , respectively. Interestingly, there seem to be different ‘good’ choices for different values of  $r$ .

**6.1. Example: Regular problem in two dimensions.** We consider the domain  $\Omega := (-1, 1) \times (-1, 1)$ . The velocity  $u$  and the pressure  $p$  are given by

$$u(x, y) = \begin{bmatrix} 2y \cos(x^2 + y^2) \\ -2x \cos(x^2 + y^2) \end{bmatrix}, \quad p(x, y) = e^{-10(x^2 + y^2)} - \bar{p},$$

where  $\bar{p} \in \mathbb{R}$ , such that  $p$  has zero mean. The right-hand side is computed as  $f = -\operatorname{div} \mathbf{A}(\mathbf{E}u) + \nabla p$ . We run the experiment with two different powers  $r = 3$  and  $r = 1.5$ . The error decays are depicted in Figures 1 and 2.

**6.2. Example: Singular problem in two dimensions.** Let

$$\Omega := ((-1, 1) \times (-1, 1)) \setminus ([0, 1] \times [0, 1])$$

be the so-called L-shaped domain with reentering corner  $\omega = 3\pi/2$ . We take the power  $r = 3$  and  $r' = 3/2$ , i.e.,  $\frac{1}{r} + \frac{1}{r'} = 1$ . To our best knowledge there is no explicitly known singular solution of the nonlinear stationary Stokes problem with the right-hand side in  $L^{r'}(\Omega)^d$ . Therefore, we decided to just take the right-hand side to be zero and impose boundary values of a singular solution of the linear Stokes problem; compare with [BMN02, Ver89, Dau89]. More precisely, we seek a solution  $(u, p) \in W^{1,r}(\Omega) \times L^{r'}(\Omega)/\mathbb{R}$  such that

$$\begin{aligned} -\operatorname{div} \mathbf{A}(\mathbf{E}u) + \nabla p &= 0 && \text{in } \Omega, \\ u = v(\rho, \chi) &:= \rho^\alpha \begin{bmatrix} \cos(\chi)\psi'(\chi) + (1 + \alpha)\sin(\chi)\psi(\chi) \\ \sin(\chi)\psi'(\chi) - (1 + \alpha)\cos(\chi)\psi(\chi) \end{bmatrix} && \text{on } \partial\Omega, \end{aligned}$$

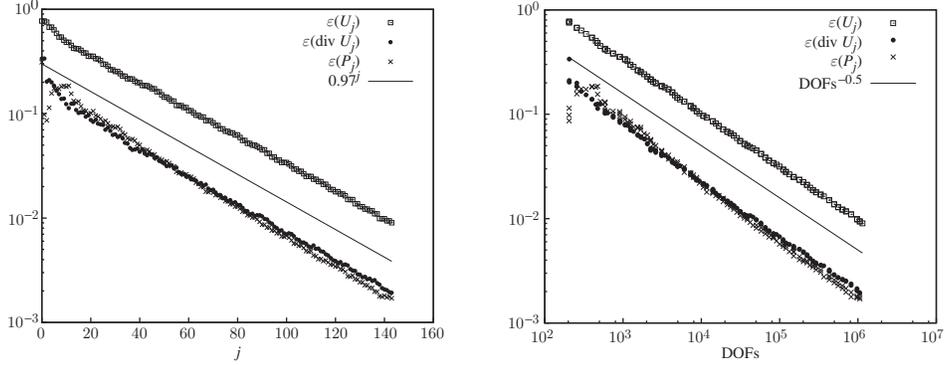


FIGURE 2. Example 6.1 with  $r = 1.5$ : All error quantities show linear convergence with respect to iterations  $j$ . The reduction factor is about  $\rho = 0.97$  (left). They show the same rate with respect to DOFs as the solid line having the slope  $1/d = 1/2$  (right).

where  $(\rho, \chi)$  are the polar coordinates on  $\Omega$  and the function  $\psi$  is given by

$$\begin{aligned} \psi(\chi) := & \frac{\sin((1 + \alpha)\chi) \cos(\alpha\omega)}{1 + \alpha} - \cos((1 + \alpha)\chi) \\ & + \frac{\sin((1 - \alpha)\chi) \cos(\alpha\omega)}{1 - \alpha} + \cos((1 - \alpha)\chi). \end{aligned}$$

Thereby  $\alpha \sim 0.544484$  is the smallest root of the equation

$$\frac{\sin^2(\alpha\omega) - \alpha^2 \sin^2(\omega)}{\alpha^2} = 0.$$

Note that  $v \in W^{1,3}(\Omega)$  and  $\operatorname{div} v = 0$ , which implies unique solvability of the problem.

Since we do not know the true solution  $(u, p)$  of this problem, we cannot compute any of the error quantities in (6.1). However, we can compute

$$\mathcal{E}_j^2(\operatorname{div} U_j) := \int_{\Omega} \phi_{|\mathbf{E}U_j|} (|\operatorname{div} U_j|) dx.$$

We expect that this quantity is vanishing with the same rate as the error quantities in (6.1). Using the inequalities of Lemma 7 (in particular (2.19)), we have for each  $\delta > 0$  that

$$\begin{aligned} \mathcal{E}_j^2(\operatorname{div} U_j) & \leq (1 + C_\delta) \mathcal{E}^2(\operatorname{div} U_j) + \delta \int_{\Omega} \phi_{|\mathbf{E}u|} (|\mathbf{E}u - \mathbf{E}U_j|) dx \\ & \leq (1 + C_\delta) \mathcal{E}^2(\operatorname{div} U_j) + \delta \mathcal{E}^2(U_j), \end{aligned}$$

where we used Corollary 6 in the last step. As is corroborated by the other experiments we assume  $\mathcal{E}(U_j) \approx \mathcal{E}(\operatorname{div} U_j)$ , which then implies  $\mathcal{E}_j(\operatorname{div} U_j) \leq \mathcal{E}(\operatorname{div} U_j)$ . Under the same assumption it can be similarly shown that  $\mathcal{E}(\operatorname{div} U_j) \leq \mathcal{E}_j(\operatorname{div} U_j)$  and thus  $\mathcal{E}_j(\operatorname{div} U_j) \approx \mathcal{E}(\operatorname{div} U_j)$ . This indicates that  $\mathcal{E}_j(\operatorname{div} U_j)$  may be a reasonable error quantity.

Another drawback of not knowing the true solution is, that we cannot analyse its singularity. We do not even know if the solution is singular at all. For this reason we additionally run the AUA using uniform refinement in each loop of the AFEM.

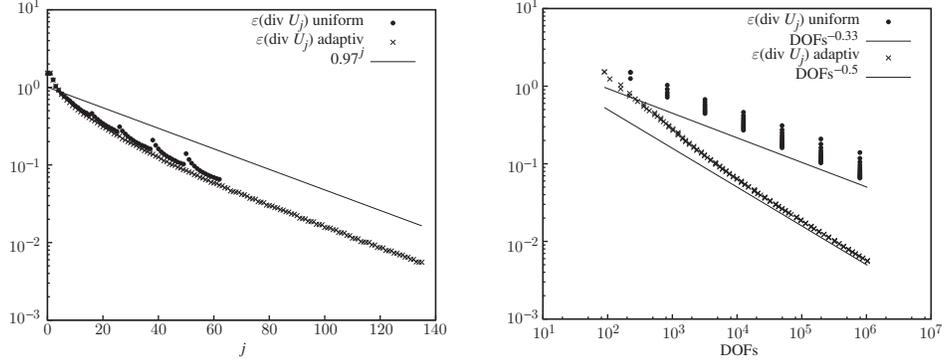


FIGURE 3. Example 6.2: In both cases the error quantity is asymptotically reduced by  $\rho = 0.97$  with respect to iterations  $j$ . Note that for uniform meshes  $j \leq 62$  (left). The AUA shows decay rate  $1/d = 1/2$  on adaptive meshes and  $\sim 1/3$  on uniform grids (right).

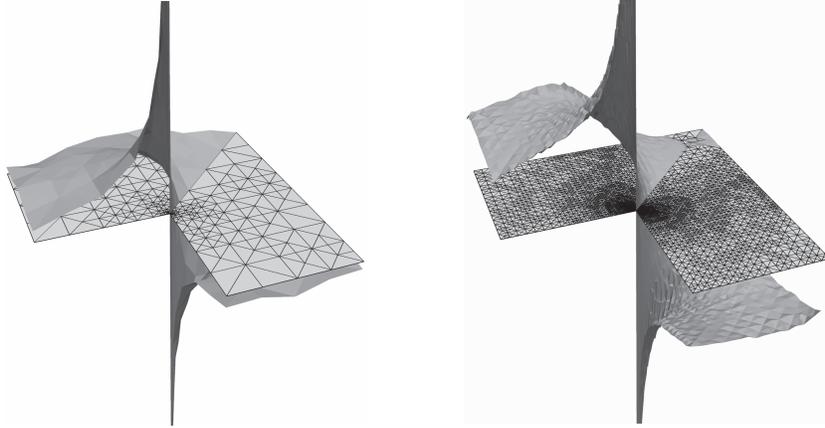


FIGURE 4. Example 6.2: Mesh and pressure of the AUA at outer iteration  $j = 6$ , DOFs = 218 (left) and  $j = 45$ , DOFs = 4667 (right).

Figure 3 shows the convergence rates using adaptive and uniform refinements. The decay rate for the uniform refining version of the AUA indicates that the true solution is indeed singular. This is also suggested by Figure 4, which shows the pressure and the adapted grid at the outer iterations  $j = 6$  and  $j = 45$ .

**6.3. Example: Regular problem in three dimensions.** We consider the exact velocity  $u$  and pressure  $p$ ,

$$u(x, y, z) := \begin{bmatrix} 2(y-z)\cos(x^2+y^2+z^2) \\ 2(z-x)\cos(x^2+y^2+z^2) \\ 2(x-y)\cos(x^2+y^2+z^2) \end{bmatrix}, \quad p(x, y, z) := e^{-10(x^2+y^2+z^2)} - \bar{p}$$

on the cube  $\Omega := (-1, 1)^3$ , where  $\bar{p} \in \mathbb{R}$  such that  $\int_{\Omega} p = 0$ . The right-hand side is computed as  $f = -\operatorname{div} \mathbf{A}(\mathbf{E}u) + \nabla p$ . The computed error decays are presented

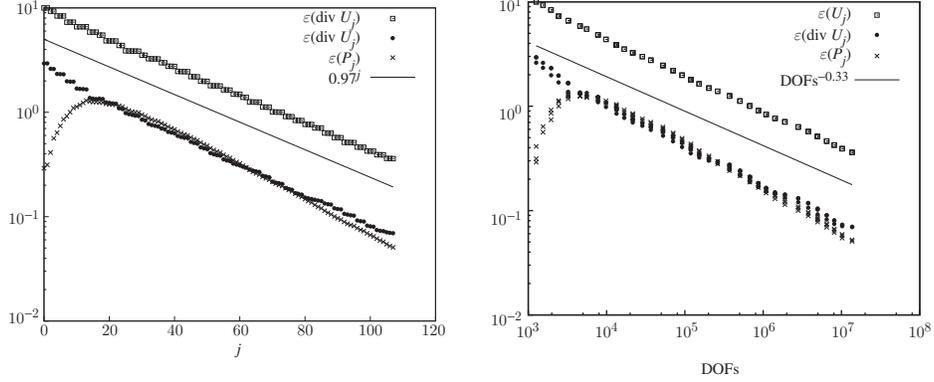


FIGURE 5. Example 6.3 with  $\sigma = 10$ : The error quantities are asymptotically reduced by  $\rho = 0.97$  with respect to iterations  $j$  (left). They show the same decay rate as the solid line having slope  $1/d = 1/3$  (right).

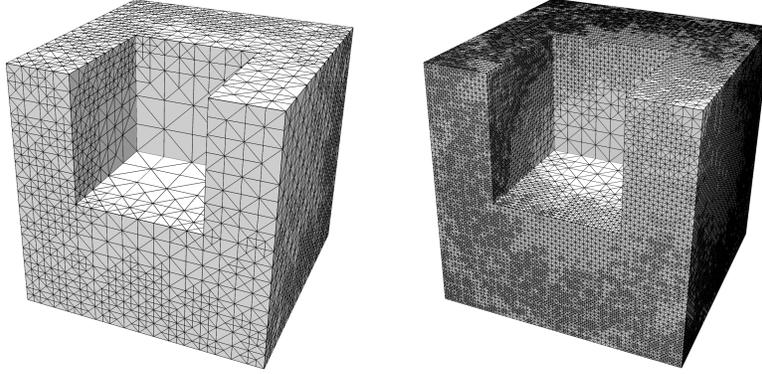


FIGURE 6. Example 6.3: Mesh produced by the AUA at outer iteration  $j = 30$ , DOFs = 21581 (left) and  $j = 75$ , DOFs = 866811 (right). We removed a cube with edge length 1 for visualization purposes.

in Figure 5. Note that the error quantity  $\mathcal{E}(P_j)$  is increasing in the first iterations. This effect can be observed already for the linear Stokes problem (see [BMN02, Figure 3.6 (left)]). We suppose that this behavior is due to the resolution of the data and indeed, starting the AUA from a very fine mesh  $\mathcal{T}_0$  (DOF  $\sim 10^6$ ) and with small  $\tau$ , we observed that  $\mathcal{E}(P_j)$  is decreasing at the first iterations.

The adapted grid is shown in Figure 6 for the iterations  $j = 30$  and  $j = 75$ . Even though the solution is regular the AUA produces a strongly adapted grid.

#### 6.4. Conclusions.

- We comment on the experiments of sections 6.1–6.3.
- The choice  $\rho = 0.97$  is rather conservative and thus may lead to unnecessary many iterations of the AUA. On the other hand, taking  $\rho$  to small may lead to suboptimal meshes. In the linear case a ‘good’ choice of  $\rho$  depends on the norm of the Schur complement operator and is thus connected with the

linear convergence of the AUA in terms of outer iterations  $j$ ; compare with [BMN02, Remark 5.6]. In the nonlinear case this question is open. Since we are interested in convergence rates we decided for the conservative choice.

- Figures 1–3 and 5 (left) show that the error is asymptotically reduced. This raises the hope that a generalized inf-sup condition (5.8) holds. As for the linear Stokes problem in [BMN02], the error decay in each iteration corresponds to the tolerance reduction factor  $\rho = 0.97$ .
- For all error quantities, the AUA shows the expected convergence rates (6.2) with respect to degrees of freedom; see Figures 1–3 and 5 (right).
- The AUA shows a strong adaptation of the grids in Figures 4 and 6. This may lead to a benefit, even for regular solutions.

#### ACKNOWLEDGEMENTS

The author would like to thank Kunibert G. Siebert for providing the ALBERTA solver for the nonlinear  $p$ -Laplace problem. The author thanks the editor and the anonymous referees for their constructive comments that helped to improve the quality and the presentation of this work.

#### REFERENCES

- [AG94] C. Amrouche and V. Girault, *Decomposition of vector spaces and application to the stokes problem in arbitrary dimension.*, Czech. Math. J. **44** (1994), no. 1, 109–140. MR1257940 (95c:35190)
- [AO00] M. Ainsworth and J. T. Oden, *A posteriori error estimation in finite element analysis.*, Pure and Applied Mathematics. A Wiley-Interscience Series of Texts, Monographs, and Tracts. Chichester: Wiley, 2000. MR1885308 (2003b:65001)
- [Bän91] E. Bänsch, *Local mesh refinement in 2 and 3 dimensions.*, IMPACT Comput. Sci. Eng. **3** (1991), no. 3, 181–191. MR1141298 (92h:65150)
- [BMN02] E. Bänsch, P. Morin, and R. H. Nochetto, *An adaptive Uzawa fem for the Stokes problem: Convergence without the inf-sup condition.*, SIAM J. Numer. Anal. **40** (2002), no. 4, 1207–1229. MR1951892 (2004e:65127)
- [BL93a] J. W. Barrett and W. B. Liu, *Finite element approximation of the  $p$ -Laplacian.*, Math. Comput. **61** (1993), no. 204, 523–537. MR1192966 (94c:65129)
- [BL93b] ———, *Finite element error analysis of a quasi-newtonian flow obeying the Carreau or power law.*, Numer. Math. **64** (1993), no. 4, 433–453. MR1213411 (94b:65144)
- [BL94] ———, *Quasi-norm error bounds for the finite element approximation of a non-newtonian flow.*, Numer. Math. **68** (1994), no. 4, 437–456. MR1301740 (95h:65078)
- [BRS04] J. W. Barrett, J. A. Robson, and E. Süli, *A posteriori error analysis of mixed finite element approximations to quasi-newtonian incompressible flows.*, Research Reports from the Numerical Analysis Group of the computing laboratory at Oxford University, UK **NA-04/13** (2004), 1–16.
- [BDK10] L. Belenki, L. Diening, and C. Kreuzer, *Optimality of an adaptive finite element method for the nonlinear  $p$ -Laplacian equation*, Preprint Universität Duisburg-Essen **718** (2010), to appear IMA Journal of Numerical Analysis.
- [BS08] Stefano Berrone and Endre Süli, *Two-sided a posteriori error bounds for incompressible quasi-Newtonian flows.*, IMA J. Numer. Anal. **28** (2008), no. 2, 382–421. MR2401203 (2009b:76091)
- [BF91] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods.*, Springer Series in Computational Mathematics, 15. New York, Springer-Verlag, 1991. MR1115205 (92d:65187)
- [CKNS08] M. Cascon, C. Kreuzer, R. H. Nochetto, and K. G. Siebert, *Quasi-optimal convergence rate for an adaptive finite element method*, SIAM J. Numer. Anal. **46** (2008), no. 5, 2524–2550. MR2421046 (2009h:65174)

- [Cia88] P. G. Ciarlet, *Introduction to numerical linear algebra and optimization*, with the assistance of Bernadette Miara and Jean-Marie Thomas for the exercises. Transl. by A. Buttigieg., Cambridge Texts in Applied Mathematics. Cambridge, Cambridge University Press, 1988. MR1015713 (90k:65001)
- [Clé75] Ph. Clément, *Approximation by finite element functions using local regularization.*, RAIRO Anal. Numér. **9** (1975), no. R-2, 77–84. MR0400739 (53:4569)
- [Dau89] M. Dauge, *Stationary Stokes and Navier-Stokes systems on two- or three-dimensional domains with corners. I: Linearized equations.*, SIAM J. Math. Anal. **20** (1989), no. 2, 74–97. MR977489 (90b:35191)
- [DE08] L. Diening and F. Ettwein, *Fractional estimates for non-differentiable elliptic systems with general growth*, Forum Mathematicum **3** (2008), 523–556. MR2418205 (2009h:35101)
- [DHU00] S. Dahlke, R. Hochmuth, and Karsten Urban, *Adaptive wavelet methods for saddle point problems.*, M2AN, Math. Model. Numer. Anal. **34** (2000), 1003–1022. MR1837765 (2002e:65078)
- [DK08] L. Diening and C. Kreuzer, *Linear convergence of an adaptive finite element method for the  $p$ -Laplacian equation*, SIAM J. Numer. Anal. **46** (2008), no. 2, 614–638. MR2383205 (2009a:65313)
- [DR07a] L. Diening and M. Růžička, *Interpolation operators in Orlicz-Sobolev spaces.*, Numer. Math. **107** (2007), no. 1, 107–129. MR2317830 (2010g:65208)
- [DR07b] ———, *Non-Newtonian fluids and function spaces.*, Nonlinear Analysis, Function Spaces and Applications **8** (2007), 94–143. MR2657118
- [DRS10] L. Diening, M. Růžička, and K. Schumacher, *A decomposition technique for John domains*, Ann. Acad. Sci. Fenn. Ser. A. I. Math. **35** (2010), 87–114. MR2643399
- [ET76] I. Ekeland and R. Temam, *Convex analysis and variational problems.*, Studies in Mathematics and its Applications. Vol. 1. Amsterdam - Oxford: North-Holland Publishing Company; New York: American Elsevier Publishing Company, Inc., 1976. MR0463994 (57:3931b)
- [FV06] F. Fierro and A. Veeger, *A posteriori error estimators, gradient recovery by averaging, and superconvergence.*, Numer. Math. **103** (2006), no. 2, 267–298. MR2222811 (2007a:65178)
- [For81] M. Fortin, *Old and new finite elements for incompressible flows.*, Int. J. Numer. Methods Fluids **1** (1981), 347–364. MR633812 (83a:76015)
- [GR86] V. Girault and P.-A. Raviart, *Finite element methods for Navier-Stokes equations. theory and algorithms. (extended version of the 1979 publ.)*, Springer Series in Computational Mathematics, 5, Berlin, Springer-Verlag, 1986. MR851383 (88b:65129)
- [KK91] V. Kokilashvili and M. Krbeč, *Weighted inequalities in Lorentz and Orlicz spaces.*, Singapore, World Scientific Publishing Co. Pte. Ltd., 1991. MR1156767 (93g:42013)
- [KS08] Y. Kondratyuk and R. Stevenson, *An optimal adaptive finite element method for the Stokes problem.*, SIAM J. Numer. Anal. **46** (2008), no. 2, 747–775. MR2383210 (2009b:65312)
- [Kos94] I. Kossaczky, *A recursive approach to local mesh refinement in two and three dimensions.*, J. Comput. Appl. Math. **55** (1994), no. 3, 275–288. MR1329875 (95m:65207)
- [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.
- [Mau95] J. M. Maubach, *Local bisection refinement for  $n$ -simplicial grids generated by reflection.*, SIAM J. Sci. Comput. **16** (1995), no. 1, 210–227. MR1311687 (95i:65128)
- [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. MR2192319 (2006i:65201)
- [Mit89] W. F. Mitchell, *A comparison of adaptive refinement techniques for elliptic problems.*, ACM Trans. Math. Softw. **15** (1989), no. 4, 326–347. MR1062496
- [MNS00] P. Morin, R. H. Nochetto, and K. G. Siebert, *Data oscillation and convergence of adaptive fem.*, SIAM J. Numer. Anal. **38** (2000), no. 2, 466–488. MR1770058 (2001g:65157)

- [MSV07] P. Morin, K. G. Siebert, and A. Veerer, *Convergence of finite elements adapted for weaker norms*, V. Cutello, G. Fotia, and L. Puccio (Eds.): Applied and Industrial Mathematics in Italy - II, Selected Contributions from the 8th SIMAI Conference **08** (2007), 468–479. MR2367592
- [MSV08] ———, *A basic convergence result for conforming adaptive finite elements*, Math. Models Methods Applications **18** (2008), no. 5, 707–737. MR2413035 (2009d:65178)
- [Mus83] J. Musielak, *Orlicz spaces and modular spaces.*, Lecture Notes in Mathematics. 1034. Berlin, Springer-Verlag, 1983. MR724434 (85m:46028)
- [NP04] R. H. Nochetto and J.-H. Pyo, *Optimal relaxation parameter for the Uzawa method.*, Numer. Math. **98** (2004), no. 4, 695–702. MR2099317 (2005h:65052)
- [RR91] M. M. Rao and Z. D. Ren, *Theory of Orlicz spaces.*, Pure and Applied Mathematics, 146. New York, Marcel Dekker, Inc., 1991. MR1113700 (92e:46059)
- [SS05] A. Schmidt and K. G. Siebert, *Design of adaptive finite element software. The finite element toolbox ALBERTA*, Lecture Notes in Computational Science and Engineering 42. Berlin: Springer, 2005. MR2127659 (2005i:65003)
- [Sie09] K. G. Siebert, *A convergence proof for adaptive finite elements without lower bound*, IMA Journal of Numerical Analysis, doi:10.1093/imanum/drq001, May 2010.
- [Squ08] A. Squillacote, *Paraview Guide*, Third edition, Kitware, Inc., 2008.
- [Ste07] R. Stevenson, *Optimality of a standard adaptive finite element method.*, Found. Comput. Math. **7** (2007), no. 2, 245–269. MR2324418 (2008i:65272)
- [Ste08] ———, *The completion of locally refined simplicial partitions created by bisection.*, Math. Comput. **77** (2008), no. 261, 227–241. MR2353951 (2008j:65219)
- [Vee02] A. Veerer, *Convergent adaptive finite elements for the nonlinear Laplacian.*, Numer. Math. **92** (2002), no. 4, 743–770. MR1935808 (2003j:65121)
- [Ver89] R. Verfürth, *A posteriori error estimators for the Stokes equations.*, Numer. Math. **55** (1989), no. 3, 309–325. MR993474 (90d:65187)

FAKULTÄT FÜR MATHEMATIK, UNIVERSITÄT DUISBURG-ESSEN, FORSTHAUSWEG 2, DUISBURG, GERMANY 47057

URL: <http://www.uni-due.de/mathematik/numa/mitarbeiter-kreuzer>

E-mail address: christian.kreuzer@uni-due.de