

No. 679

January 2025

**PRESSURE ROBUST FINITE ELEMENT
DISCRETIZATIONS OF THE NONLINEAR
STOKES EQUATIONS**

L. Diening, A. Hirn, C. Kreuzer, P. Zanotti

ISSN: 2190-1767

PRESSURE ROBUST FINITE ELEMENT DISCRETIZATIONS OF THE NONLINEAR STOKES EQUATIONS

LARS DIENING, ADRIAN HIRN, CHRISTIAN KREUZER, AND PIETRO ZANOTTI

ABSTRACT. We present first-order nonconforming Crouzeix-Raviart discretizations for the nonlinear generalized Stokes equations with (r, ε) -structure. Thereby the velocity-errors are independent of the pressure-error; i.e., the method is pressure robust. This improves suboptimal rates previously experienced for non pressure robust methods.

1. INTRODUCTION

In this article, we study the finite element discretization of the nonlinear Stokes equations in a bounded domain $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$,

$$(1.1) \quad -\operatorname{div} \mathbf{A}(\mathcal{D}\mathbf{u}) + \nabla p = \mathbf{f} \quad \text{and} \quad \operatorname{div} \mathbf{u} = 0 \quad \text{in } \Omega, \quad \mathbf{u} = 0 \quad \text{on } \partial\Omega.$$

Here \mathbf{u} denotes the velocity, p the kinematic pressure, and \mathbf{f} an external body force. The symbol $\mathcal{D}\mathbf{u}$ stands either for the velocity gradient $\nabla \mathbf{u}$ or for its symmetric part $\mathcal{D}\mathbf{u} = \frac{1}{2}(\nabla \mathbf{u} + (\nabla \mathbf{u})^T)$. For given $r > 1$, $\varepsilon \geq 0$ the nonlinear function $\mathbf{A} : \mathbb{R}^{d \times d} \rightarrow \mathbb{R}^{d \times d}$ is supposed to have (r, ε) -structure (or more generally φ -structure; see Assumption 2.3). A prototypical example is

$$(1.2) \quad \mathbf{A}(\mathcal{D}\mathbf{u}) = (\varepsilon^2 + |\mathcal{D}\mathbf{u}|^2)^{\frac{r-2}{2}} \mathcal{D}\mathbf{u}.$$

For $\varepsilon = 0$ the operator corresponds to the r -Laplacian, while for $\varepsilon > 0$ the operator regularizes the degeneracy of the r -Laplacian as the modulus of $\mathcal{D}\mathbf{u}$ tends to zero. Equations with (r, ε) -structure arise in various physical applications, such as in the theory of plasticity, bimaterial problems in elastic-plastic mechanics, non-Newtonian fluid mechanics, blood rheology and glaciology; see e.g. [38, 39, 25].

The finite element (FE) approximation of equations with (r, ε) -structure has been widely studied, see e.g. [3, 15, 19, 4, 5, 9, 27]. According to [15, 9], it is relevant to bound the error between \mathbf{u} and the discrete velocity \mathbf{u}_h in the so-called natural- or \mathbf{F} -distance $\|\mathbf{F}(\mathcal{D}\mathbf{u}) - \mathbf{F}(\mathcal{D}\mathbf{v})\|_2$ where \mathbf{F} is a nonlinear vector field adapted to the problem's (r, ε) -structure; see (2.12) for the precise definition. This error notion is equivalent to the quasi-norm introduced by Barrett/Liu [3, 5].

For conforming exactly divergence-free discretizations Céa-type estimates

$$(1.3) \quad \|\mathbf{F}(\mathcal{D}\mathbf{u}) - \mathbf{F}(\mathcal{D}\mathbf{u}_h)\|_2 \lesssim \inf_{\operatorname{div} \mathbf{v}_h = 0} \|\mathbf{F}(\mathcal{D}\mathbf{u}) - \mathbf{F}(\mathcal{D}\mathbf{v}_h)\|_2$$

as for the elliptic r -Laplace problem [3, Theorem 2.1] can be obtained by straight forward application of the elliptic techniques; see e.g. [15, Lemma 5.2]. We call a method satisfying (1.3) *quasi-optimal and pressure-robust*: the velocity-error measured in the natural distance is proportional to the corresponding best error and independent of the pressure. Such discretizations have recently received much attention for linear Stokes-equations; compare with [22, 23, 42, 48]. In order to

Date: January 27, 2025.

2020 Mathematics Subject Classification. 35J62, 35J92, 65N30, 65N12, 65N15, 76A05, 76M10.

Key words and phrases. non-Newtonian Stokes, p -Stokes, finite element method, a priori error estimates.

replace the global approximation problem on the right hand side of (1.3) by local approximation problems, usually local Fortin operators are employed; compare with [20, §54.1]. Such operators are typically not perfectly local in the sense that the interpolated function on a mesh element K depends on the function values in the neighbourhood ω_K of K . This results in

$$(1.4) \quad \|\mathbf{F}(\mathcal{D}\mathbf{u}) - \mathbf{F}(\mathcal{D}\mathbf{u}_h)\|_2^2 \lesssim \sum_K \inf_{\mathbf{Q}_K \in \mathbb{R}^{d \times d}} \|\mathbf{F}(\mathcal{D}\mathbf{u}) - \mathbf{Q}_K\|_{2;\omega_K}^2.$$

Corresponding results for the nonlinear Laplacian can be found in [15]. Note that in the case of the linear Stokes problem, the neighbourhood ω_K can be avoided using techniques from [43] but for $r \neq 2$, properties of the natural distance prevent from a direct generalisation; compare with Remark 3.4. It is also noticeable that all above a priori bounds are limited to first order at most. Here again, the reason is the nonlinear structure of the natural distance, in particular, the use of Jensen's inequality while generalising stability properties to the natural distance.

For not exactly divergence-free discretizations, it was proved in [9] that

$$(1.5) \quad \begin{aligned} \|\mathbf{F}(\mathcal{D}\mathbf{u}) - \mathbf{F}(\mathcal{D}\mathbf{u}_h)\|_2^2 &\lesssim \sum_K \inf_{\mathbf{Q}_K \in \mathbb{R}^{d \times d}} \|\mathbf{F}(\mathcal{D}\mathbf{u}) - \mathbf{Q}_K\|_{2;\omega_K}^2 \\ &+ \inf_{q_h \in \mathbb{Q}_h} \int_{\Omega} (|p - q_h|^2 + |\mathcal{D}\mathbf{u}|^2)^{\frac{r'-2}{2}} |p - q_h|^2 \end{aligned}$$

with r' the Hölder conjugate of r . This estimate substantially differs from (1.4) in that it is not pressure robust, i.e., the velocity-error is interfered by the pressure; compare with [28]. In particular, for $r \neq 2$, this interference is of nonlinear nature and may lead to suboptimal convergence rates even for regular solutions with $\mathbf{F}(\mathcal{D}\mathbf{u}) \in W^{1,2}(\Omega)^{d \times d}$ and $p \in W^{1,r'}(\Omega)$. Indeed, Belenki, Berselli, Diening and Růžička [9] derived from (1.5) for the MINI element the the following error estimate in terms of the mesh-size h

$$(1.6) \quad \|\mathbf{F}(\mathcal{D}\mathbf{u}) - \mathbf{F}(\mathcal{D}\mathbf{u}_h)\|_2 \lesssim h^{\min\{1, \frac{r'}{2}\}}.$$

Similar results are derived by [27] for a stabilized $\mathbb{Q}_1/\mathbb{Q}_1$ -discretization. Comparing (1.6) and (1.4), we conclude that for $r > 2$ the velocity-error estimate (1.6) is suboptimal with respect to the approximation properties of the discrete space. This is also confirmed by numerical studies in [9].

A similar reasoning (i.e. the pressure-error depends on the velocity-error) limits the theoretical convergence of the pressure error in [9] to

$$(1.7) \quad \|p - p_h\|_{r'} \lesssim h^{\min\{\frac{2}{r'}, \frac{r'}{2}\}}.$$

In this case, however, the numerical results in [9] show $\mathcal{O}(h^{\min\{\frac{2}{r'}, 1\}})$ convergence, indicating that the bound for the pressure error possibly is not sharp for $r > 2$.

In [30], Kaltenbach/Růžička obtain first order convergence for the velocity-error of a local discontinuous Galerkin method. The results, however are ε -dependent in the sense that the constants blow up fast as $\varepsilon \rightarrow 0$. Moreover, they exploit additional regularity of f and the involved jump penalizations result in non-monotone schemes, with possible non-unique solutions; cf. [29, §4].

In this paper we aim at achieving (1.4) without the above restrictions. To this end, we follow a different approach and extend the quasi-optimal and pressure-robust approach [47, 35, 34] of Kreuzer/Verfürth/Zanotti from linear to nonlinear Stokes equations (1.1) with (r, ε) -structure. In particular, the methods have the following features: First, we propose two discretizations with nonconforming Crouzeix-Raviart (CR) elements that are stabilization-free and therefore yield monotone numerical schemes. As a consequence, the numerical solution is unique. Second, a sophisticated *smoothing* operator acting on the test functions in the load term,

allows to handle load-functions $\mathbf{f} \in \mathbf{W}^{-1,r'}(\Omega)$ without requiring extra regularity of the data. Third, this operator maps element-wise divergence-free functions onto exactly divergence-free functions, so to ensure the pressure robustness. Note that the pressure robust methods of Linke et al. [37, 36] focus only on the latter property and are not considered here.

Based on these features, we prove a priori bounds of the form (1.4) (see Theorems 3.3 and 4.8). It is remarkable that the results neither resort to additional regularity of the data and are robust for the critical limit $\varepsilon \rightarrow 0$ of the regularisation parameter in (1.2).

As a consequence, we obtain for regular velocities $\mathbf{F}(\mathcal{D}\mathbf{u}) \in W^{1,2}(\Omega)^{d \times d}$ full first order convergence independent of the pressure

$$(1.8) \quad \|\mathbf{F}(\mathcal{D}\mathbf{u}) - \mathbf{F}(\mathcal{D}_h\mathbf{u}_h)\|_2 \lesssim \|h\nabla\mathbf{F}(\mathcal{D}\mathbf{u})\|_2.$$

As the pressure-error depends on the velocity-error, as a side effect, the improved a priori bounds for the velocity also improve the bounds for the pressure-error. In fact, if in addition $p \in W^{1,r'}(\Omega)$ then we obtain $\|p - p_h\|_{r'} = \mathcal{O}(h^{\min\{\frac{2}{r}, 1\}})$, which corresponds to the rates observed in the numerical examples of [9] for the MINI element using a different method.

The plan of the paper is as follows: In Section 2, we introduce the φ -structure (a generalization of the (r, ε) -structure), the nonlinear problem (1.1) as well as the finite element framework. Sections 3 and 4 each presents the a priori analysis for the velocity-error of a pressure robust method based on Crouzeix-Raviart elements for $\mathcal{D} = \nabla$ and $\mathcal{D} = \mathbf{D}$ respectively. Section 5 concerns the improved bounds for the pressure-error followed by some numerical experiments in Section 6.

2. PRELIMINARIES

To begin with, we clarify our notation and we state important properties of the nonlinear operator \mathbf{A} and of the natural distance. Further, we introduce the variational formulation of (1.1).

2.1. Basic notation and function spaces. Let \mathbb{R}^+ be the set of all positive real numbers, and $\mathbb{R}_0^+ := \mathbb{R}^+ \cup \{0\}$. The Euclidean scalar-product of two vectors $\mathbf{p}, \mathbf{q} \in \mathbb{R}^d$ is denoted by $\mathbf{p} \cdot \mathbf{q}$ and the Frobenius product of $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^{d \times d}$ is defined by $\mathbf{P} : \mathbf{Q} := \sum_{i,j=1}^d P_{ij}Q_{ij}$. We set $|\mathbf{p}| := (\mathbf{p} \cdot \mathbf{p})^{1/2}$ and $|\mathbf{Q}| := (\mathbf{Q} : \mathbf{Q})^{1/2}$.

For an open set $\omega \subset \mathbb{R}^d$, and $r \in [1, \infty]$, we denote by $L^r(\omega)$ the space of scalar r -integrable functions on ω with corresponding norm $\|\cdot\|_{r;\omega}$. The space $L_0^r(\omega)$ is the closed subspace of $L^r(\omega)$ of functions with vanishing mean value. We equip the first order Sobolev space $W^{1,r}(\omega)$ with norm $\|\cdot\|_{1,r;\omega} = (\|\cdot\|_{r;\omega}^r + \|\nabla \cdot\|_{r;\omega}^r)^{1/r}$ and denote by $W_0^{1,r}(\omega)$ its closed subspace of functions with vanishing trace on $\partial\Omega$. For $r \in (1, \infty)$ and $\frac{1}{r} + \frac{1}{r'} = 1$, i.e. $r' = \frac{r}{r-1}$, we have that $W_0^{1,r}(\omega)$ is a reflexive Banach space with dual space $W^{-1,r'}(\omega) = (W_0^{1,r}(\omega))^*$ and for the dual pairing between $f \in W^{-1,r'}(\omega)$ and $v \in W_0^{1,r}(\omega)$, we use the notation $\langle f, v \rangle_\omega$. We use the convention $W^{0,r}(\omega) = W_0^{0,r}(\omega) = L^r(\omega)$. Orlicz spaces $L^\varphi(\omega)$ and Sobolev-Orlicz spaces $W^{1,\varphi}(\omega)$ are introduced later in Section 2.3.

In case of $\omega = \Omega$, we usually omit the index Ω , e.g. we write $\|\cdot\|_r$ instead of $\|\cdot\|_{r;\Omega}$. Spaces of \mathbb{R}^d -valued functions are denoted with boldface type, though no distinction is made in the notation of norms and inner products; for instance, the norm in $\mathbf{W}^{1,r}(\Omega) = (W^{1,r}(\Omega))^d$ is given by $\|\mathbf{v}\|_{1,r} = \left(\sum_{1 \leq i \leq d} \|v_i\|_{1,r}^r\right)^{1/r}$. Thanks to Poincaré and Korn inequalities, an alternative norm on $\mathbf{W}_0^{1,r}(\Omega)$ is given by $\|\mathcal{D} \cdot\|_r$ if $r \in (1, \infty)$; compare e.g. with [16, Theorem 6.10].

For a set $\omega \subset \mathbb{R}^d$ with Hausdorff dimension d or $d-1$, we denote by $|\omega|$ its d - or $(d-1)$ -dimensional Hausdorff measure and for $f \in L^r(\omega)$, we use the abbreviation

$$\langle f \rangle_\omega := \int_\omega f := \frac{1}{|\omega|} \int_\omega f$$

for its mean value in ω .

In order to simplify the notation, we shall often denote $a \lesssim b$ when $a \leq cb$ for a constant $c > 0$ only depending on fixed but maybe problem specific constants. For $a \lesssim b \lesssim a$ we shall also write $a \approx b$.

2.2. N-functions. A convenient abstract framework for problems with an (r, ε) -structure like (1.2) is based on so-called N-functions, that are standard in the theory of Orlicz spaces; cf. [40, 32]. A continuous, convex, and strictly monotone function $\psi : \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+$ is called an N-function if

$$\psi(0) = 0, \quad \lim_{t \rightarrow 0} \frac{\psi(t)}{t} = 0 \quad \text{and} \quad \lim_{t \rightarrow \infty} \frac{\psi(t)}{t} = \infty.$$

Thanks to the convexity of ψ , there exists its right derivative ψ' , which is non-decreasing and satisfies $\psi'(0) = 0$, $\psi'(t) > 0$ for $t > 0$ and $\lim_{t \rightarrow \infty} \psi'(t) = \infty$.

The conjugate of an N-function ψ is defined by

$$\psi^*(t) := \sup_{s \geq 0} (st - \psi(s)) \quad \text{for all } t \geq 0$$

and is again an N-function and we have $(\psi^*)^* = \psi$. As an example: the conjugate of $\psi(t) = \frac{1}{r}t^r$ is $\psi^*(t) = \frac{1}{r'}t^{r'}$ with $\frac{1}{r} + \frac{1}{r'} = 1$.

For a quantitative numerical approach we require more regular N-functions. We say that an N-function ψ is uniform convex, if additionally $\psi \in W_{\text{loc}}^{2,1}((0, \infty))$ (or $C^1([0, \infty))$ and piecewise C^2) and

$$(2.1) \quad 1 < r^- := \inf_{t > 0} \frac{\psi''(t)t}{\psi'(t)} + 1 \leq \sup_{t > 0} \frac{\psi''(t)t}{\psi'(t)} + 1 =: r^+ < \infty.$$

We call r^\pm the indices of uniform convexity of ψ . In the following we assume that ψ is a uniformly convex N-function. Note that uniform convexity rules out almost linear growth as well as exponential growth.

Let us recall the properties of uniformly convex N-functions from [13, Appendix B]. If ψ is uniformly convex with indices r^- and r^+ , then ψ^* is uniformly convex with indices $(r^+)'$ and $(r^-)'$ respectively. Furthermore, for all $s, t \geq 0$ there holds

$$(2.2) \quad \min\{s^{r^-}, s^{r^+}\} \psi(t) \leq \psi(st) \leq \max\{s^{r^-}, s^{r^+}\} \psi(t).$$

As a consequence ψ and ψ^* satisfy the so-called Δ_2 -condition with

$$\psi(2t) \leq 2^{r^+} \psi(t) \quad \text{and} \quad \psi^*(2t) \leq 2^{(r^-)'} \psi^*(t), \quad \text{for all } t \geq 0.$$

By convexity, we have the quasi-triangle inequality

$$(2.3) \quad \psi(s+t) \leq \frac{1}{2}\psi(2s) + \frac{1}{2}\psi(2t) \leq 2^{r^+-1}(\psi(s) + \psi(t)) \quad \forall s, t \geq 0.$$

It follows also from (2.2) and $st \leq \psi(s) + \psi^*(t)$ that

$$(2.4) \quad st \leq \delta^{1-r^+} \psi(s) + \delta \psi^*(t), \quad \text{for all } s, t \geq 0, \delta \in (0, 1].$$

Moreover, we have for all $t \geq 0$ that

$$(2.5) \quad \psi(t) \approx t\psi'(t) \quad \text{and} \quad \psi^*(\psi'(t)) \approx \psi(t)$$

with constants depending solely on r^\pm .

We next introduce the notion of shifted N-function first introduced in [18] and further developed for example in [15, 17, 10, 33, 13]. Here, we use the version of

[13, Appendix B], where also the self-contained proofs for the properties below can be found.

For a given uniformly convex N-function ψ with indices of uniform convexity r^\pm , the family of shifted functions $\{\psi_a\}_{a \geq 0}$ is defined by

$$(2.6) \quad \psi_a(t) := \int_0^t \psi'_a(s) \, ds \quad \text{with} \quad \psi'_a(t) := \frac{\psi'(\max\{a, t\})}{\max\{a, t\}} t.$$

In the original work [18], the maximum $\max\{a, t\}$ was replaced by $a + t$. However, though equivalent, the definition (2.6) has some technical advantages. We have $(\psi_a)^* = (\psi^*)_{\psi'_a}$ and the functions ψ_a are again uniformly convex and with indices of uniform convexity $\min\{r^-, 2\}$ and $\max\{r^+, 2\}$. Hence, ψ_a and $(\psi_a)^*$ satisfy the Δ_2 -condition with constants independent of $a \geq 0$.

A first basic property states for all $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^{d \times d}$ that

$$(2.7) \quad \psi_{|\mathbf{P}|}(|\mathbf{P} - \mathbf{Q}|) \approx \psi_{|\mathbf{Q}|}(|\mathbf{P} - \mathbf{Q}|) \quad \text{and} \quad \psi'_{|\mathbf{P}|}(|\mathbf{P} - \mathbf{Q}|) \approx \psi'_{|\mathbf{Q}|}(|\mathbf{P} - \mathbf{Q}|).$$

Moreover, for $t \geq 0$ and all $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^{d \times d}$ we have

$$(2.8) \quad |\psi'_{|\mathbf{P}|}(t) - \psi'_{|\mathbf{Q}|}(t)| \lesssim \psi'_{|\mathbf{Q}|}(||\mathbf{P}| - |\mathbf{Q}||) \leq \psi'_{|\mathbf{Q}|}(|\mathbf{P} - \mathbf{Q}|),$$

where the hidden constant depends only on r^\pm . Based on this, for any $\delta \in (0, 1]$, all $t \geq 0$ and all $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^{d \times d}$ there holds

$$(2.9a) \quad \psi_{|\mathbf{P}|}(t) \leq C \delta^{1-r^+} \psi_{|\mathbf{Q}|}(t) + \delta |\mathbf{F}(\mathbf{P}) - \mathbf{F}(\mathbf{Q})|^2,$$

$$(2.9b) \quad (\psi_{|\mathbf{P}|})^*(t) \leq C \delta^{1-(r^-)'} (\psi_{|\mathbf{Q}|})^*(t) + \delta |\mathbf{F}(\mathbf{P}) - \mathbf{F}(\mathbf{Q})|^2,$$

where C only depends on r^\pm . These estimates are often called *shift-change*.

When $\mathcal{D} = \mathbf{D}$ we shall need the following Korn inequality on bounded John domains, which can e.g. be found in [16, Theorem 6.13]. The precise definition of a John domain is not important in this context, just note that it applies to any bounded Lipschitz domain.

Proposition 2.1 ([16, Theorem 6.13]). *Let ψ be a uniformly convex N-function with indices r^\pm . Then for a bounded John domain $\omega \subset \mathbb{R}^d$ and all $\mathbf{w} \in \mathbf{W}^{1,1}(\omega)$ we have*

$$\int_\omega \psi(|\nabla \mathbf{w} - \langle \nabla \mathbf{w} \rangle_\omega|) \lesssim \int_\omega \psi(|\mathbf{D} \mathbf{w} - \langle \mathbf{D} \mathbf{w} \rangle_\omega|).$$

The hidden constant only depends on the constants r^\pm of uniform convexity of ψ and the John constant of ω . Note that if $\mathbf{w} \in \mathbf{W}_0^{1,1}(\omega)$, then $\langle \nabla \mathbf{w} \rangle_\omega = 0$ and $\langle \mathbf{D} \mathbf{w} \rangle_\omega = 0$.

2.3. Properties of the nonlinear operator. In this section we state our assumptions on the nonlinear operator \mathbf{A} from (1.1). To this end, let φ be a fixed uniformly convex N-function with indices of uniform convexity r^\pm . For simplicity, we assume that $\varphi \in C^2((0, \infty))$ (although $\varphi \in W_{\text{loc}}^{2,1}((0, \infty))$ would suffice in what follows).

For a bounded open set $\omega \subset \mathbb{R}^d$, we denote by $L^\varphi(\omega)$ and $W^{1,\varphi}(\omega)$ the usual Orlicz and Sobolev-Orlicz spaces; see [40]. In fact, $v \in L^\varphi(\omega)$ if $\int_\omega \varphi(|v|) < \infty$ and $v \in W^{1,\varphi}(\omega)$ if in addition $\int_\omega \varphi(|\nabla v|) < \infty$. The space $L^\varphi(\omega)$ is a reflexive Banach space equipped with the Luxembourg norm $\|v\|_\varphi := \inf\{\lambda: \int_\omega \varphi(|v|/\lambda) \leq 1\}$. By $L_0^\varphi(\omega)$ we denote the closed subspace of $L^\varphi(\omega)$ of functions with integral zero. The subspace $W_0^{1,\varphi}(\omega)$ of Sobolev-Orlicz functions with zero boundary values is a reflexive Banach space with norm $\|\nabla \cdot\|_\varphi$ and we denote its dual space by $W^{-1,\varphi^*}(\omega)$.

The most prominent example for φ is

$$(2.10) \quad \varphi(t) := \int_0^t \varphi'(s) \, ds \quad \text{with} \quad \varphi'(t) := (\varepsilon + t)^{r-2} t$$

with $1 < r < \infty$ and $\varepsilon \geq 0$ and we say that φ has (r, ε) -structure. In this case, we have $L^\varphi(\omega) = L^r(\omega)$ and $W^{1,\varphi}(\omega) = W^{1,r}(\omega)$. If $\varepsilon = 0$, then $r^\pm = r$ and if $\varepsilon > 0$, then $r^- = \min\{r, 2\}$ and $r^+ = \max\{r, 2\}$. This follows from

$$\frac{\varphi''(t)t}{\varphi'(t)} + 1 = \frac{rt + 2\varepsilon}{t + \varepsilon}.$$

In particular, the indices of uniform convexity of φ are bounded independently of $\varepsilon \geq 0$. As a consequence, all estimates that only depend on the indices of uniform convexity are independent of $\varepsilon \geq 0$.

Remark 2.2. *Other examples of uniformly convex N -functions are*

- (1) $\varphi(t) = t^{r_1} + t^{r_2}$ for $1 < r_1 \leq r_2 < \infty$.
- (2) $\varphi(t) = t^r \log(e + t)$ for $1 < r < \infty$.

Assumption 2.3 (Nonlinear operator). *We assume that the nonlinear operator $\mathbf{A} : \mathbb{R}^{d \times d} \rightarrow \mathbb{R}^{d \times d}$ belongs to $C^0(\mathbb{R}^{d \times d}, \mathbb{R}^{d \times d}) \cap C^1(\mathbb{R}^{d \times d} \setminus \{\mathbf{0}\}, \mathbb{R}^{d \times d})$ and satisfies $\mathbf{A}(\mathbf{0}) = \mathbf{0}$. Furthermore, we assume that \mathbf{A} has φ -structure, i.e., there exist constants $C_0, C_1 > 0$, such that¹*

$$(2.11a) \quad \sum_{i,j,k,l=1}^d \partial_{kl} A_{ij}(\mathbf{Q}) P_{ij} P_{kl} \geq C_0 \varphi''(|\mathbf{Q}|) |\mathbf{P}|^2,$$

$$(2.11b) \quad |\partial_{kl} A_{ij}(\mathbf{Q})| \leq C_1 \varphi''(|\mathbf{Q}|)$$

holds for all $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^{d \times d}$ with $\mathbf{Q} \neq \mathbf{0}$ and all $i, j, k, l \in \{1, \dots, d\}$.

Remark 2.4. *If φ is a uniformly convex N -function, then $\mathbf{A} : \mathbb{R}^{d \times d} \rightarrow \mathbb{R}^{d \times d}$ defined by*

$$\mathbf{A}(\mathbf{Q}) = \varphi'(|\mathbf{Q}|) \frac{\mathbf{Q}}{|\mathbf{Q}|}$$

has automatically φ -structure.

Remark 2.5. *If φ has (r, ε) -structure, then*

$$\sum_{i,j,k,l=1}^d \partial_{kl} A_{ij}(\mathbf{Q}) P_{ij} P_{kl} \geq \tilde{C}_0 (\varepsilon + |\mathbf{Q}|)^{r-2} |\mathbf{P}|^2, \quad |\partial_{kl} A_{ij}(\mathbf{Q})| \leq \tilde{C}_1 (\varepsilon + |\mathbf{Q}|)^{r-2}$$

and

$$\varphi_a(t) \simeq (\varepsilon + a + t)^{r-2} t^2 \quad \text{as well as} \quad (\varphi_a)^* \simeq ((\varepsilon + a)^{r-1} + t)^{r'-2} t^2$$

uniformly in $t, a \geq 0$ with constants only depending on r (but not on ε).

Several studies on the finite element analysis of the r -Laplace equation (e.g., [3, 4, 5, 15, 9]) indicate that, for quasi-optimal error bounds, a certain error measurement adapted to the structure of φ is necessary. In order to introduce this so-called natural distance, we define

$$(2.12) \quad \mathbf{F} : \mathbb{R}^{d \times d} \rightarrow \mathbb{R}^{d \times d} \quad \text{by} \quad \mathbf{F}(\mathbf{Q}) := \sqrt{\varphi'(|\mathbf{Q}|) |\mathbf{Q}|} \frac{\mathbf{Q}}{|\mathbf{Q}|}.$$

If φ has (r, ε) -structure, then $\mathbf{F}(\mathbf{Q}) = (\varepsilon + |\mathbf{Q}|)^{\frac{r-2}{2}} \mathbf{Q}$.

The functions \mathbf{A} and \mathbf{F} are closely related as depicted in following Lemma from [18, 17] and [13, Appendix B].

¹For functions $g : \mathbb{R}^{d \times d} \rightarrow \mathbb{R}$ we use the notation $\partial_{kl} g(\mathbf{Q}) := \frac{\partial g(\mathbf{Q})}{\partial Q_{kl}}$.

Lemma 2.6. *Let \mathbf{A} satisfy Assumption 2.3 and let \mathbf{F} be defined by (2.12). Then for all $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^{d \times d}$, we have*

$$\begin{aligned} (\mathbf{A}(\mathbf{P}) - \mathbf{A}(\mathbf{Q})) : (\mathbf{P} - \mathbf{Q}) &\approx (\varphi_{|\mathbf{P}|})^* (|\mathbf{A}(\mathbf{P}) - \mathbf{A}(\mathbf{Q})|) \\ &\approx |\mathbf{F}(\mathbf{P}) - \mathbf{F}(\mathbf{Q})|^2 \\ &\approx \varphi_{|\mathbf{P}|}(|\mathbf{P} - \mathbf{Q}|) \end{aligned}$$

and

$$|\mathbf{A}(\mathbf{P}) - \mathbf{A}(\mathbf{Q})| \approx \varphi'_{|\mathbf{P}|}(|\mathbf{P} - \mathbf{Q}|),$$

where the hidden constants only depend on r^\pm .

According to Lemma 2.6, for $v, u \in W^{1,\varphi}(\Omega)$, we have the equivalence

$$\int_{\Omega} (\mathbf{A}(\mathcal{D}v) - \mathbf{A}(\mathcal{D}u)) : (\mathcal{D}v - \mathcal{D}u) \approx \|\mathbf{F}(\mathcal{D}v) - \mathbf{F}(\mathcal{D}u)\|_2^2,$$

where the constants only depend on r^\pm . We call the square root of the right-hand side the natural distance or \mathbf{F} -distance.

Remark 2.7 (Quasi-norm and natural distance). *The natural distance is an important concept in the regularity theory as well as the numerical analysis of nonlinear PDEs. It was introduced under the name quasi-norm for (r, ε) -structure by Barrett and Liu in the breakthrough paper [5]. In particular, they used the quantity*

$$|w|_{(r,v)}^2 := \int_{\Omega} (\varepsilon + |\nabla v| + |\nabla w|)^{r-2} |\nabla w|^2, \quad v, w \in W^{1,r}(\Omega).$$

Indeed, in this case, we have for all $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^{d \times d}$ that

$$\varphi_{|\mathbf{P}|}(|\mathbf{P} - \mathbf{Q}|) \approx (\varepsilon + |\mathbf{P}| + |\mathbf{Q}|)^{r-2} |\mathbf{P} - \mathbf{Q}|^2$$

which implies the equivalence to the natural distance, i.e.

$$(2.13) \quad \int_{\Omega} (\mathbf{A}(\nabla v) - \mathbf{A}(\nabla u)) : (\nabla v - \nabla u) \approx \|\mathbf{F}(\nabla v) - \mathbf{F}(\nabla u)\|_2^2 \approx |v - u|_{(r,v)}^2.$$

For the quasi-norm error, Barrett and Liu proved quasi-optimality for conforming finite element Galerkin approximation of the p -Laplace problem. This was extended in [15] to more general problems with uniformly convex N -functions.

We conclude this section stating equivalences of different integral averages in the natural distance.

Lemma 2.8 ([14, Lemma A.2]). *Let $\omega \subset \mathbb{R}^d$ be a bounded open set and let $\mathbf{Q} \in L^\varphi(\omega)^{d \times d}$. Further we define $\langle \mathbf{Q} \rangle_{\omega}^{\mathbf{A}} := \mathbf{A}^{-1}(\langle \mathbf{A}(\mathbf{Q}) \rangle_{\omega})$. Then we have*

$$\int_{\omega} |\mathbf{F}(\mathbf{Q}) - \langle \mathbf{F}(\mathbf{Q}) \rangle_{\omega}|^2 \approx \int_{\omega} |\mathbf{F}(\mathbf{Q}) - \mathbf{F}(\langle \mathbf{Q} \rangle_{\omega}^{\mathbf{A}})|^2 \approx \int_{\omega} |\mathbf{F}(\mathbf{Q}) - \mathbf{F}(\langle \mathbf{Q} \rangle_{\omega})|^2$$

and the hidden constants only depend on r^\pm .

2.4. Variational formulation of the nonlinear Stokes equations. Let φ be a uniformly convex N -function with indices r^\pm of uniform convexity and assume that \mathbf{A} has φ -structure. We define the velocity and pressure spaces by

$$(2.14) \quad \mathbf{W}_0^{1,\varphi}(\Omega) \quad \text{and} \quad L_0^{\varphi^*}(\Omega).$$

The variational formulation of (1.1) reads as: for $\mathbf{f} \in \mathbf{W}^{-1,\varphi^*}(\Omega)$ find $\mathbf{u} \in \mathbf{W}_0^{1,\varphi}(\Omega)$ and $p \in L_0^{\varphi^*}(\Omega)$ with

$$(2.15a) \quad \forall \mathbf{v} \in \mathbf{W}_0^{1,\varphi}(\Omega) \quad \int_{\Omega} \mathbf{A}(\mathcal{D}\mathbf{u}) : \mathcal{D}\mathbf{v} - \int_{\Omega} p \operatorname{div} \mathbf{v} = \langle \mathbf{f}, \mathbf{v} \rangle,$$

$$(2.15b) \quad \forall q \in L_0^{\varphi^*}(\Omega) \quad \int_{\Omega} q \operatorname{div} \mathbf{u} = 0.$$

Equation (2.15b) shows that the velocity is actually in the space of divergence-free vector fields

$$(2.16) \quad \mathbf{Z} := \{\mathbf{v} \in \mathbf{W}_0^{1,\varphi}(\Omega) \mid \operatorname{div} \mathbf{v} = 0\} \subset \mathbf{W}_0^{1,\varphi}(\Omega).$$

Consequently, \mathbf{u} solves the reduced problem

$$(2.17) \quad \forall \mathbf{v} \in \mathbf{Z} \quad \int_{\Omega} \mathbf{A}(\mathcal{D}\mathbf{u}) : \mathcal{D}\mathbf{v} = \langle \mathbf{f}, \mathbf{v} \rangle.$$

According to the strict monotonicity of \mathbf{A} (cf. Lemma 2.6), existence and uniqueness of $\mathbf{u} \in \mathbf{Z}$ follows by standard monotone operator theory and Korn's inequality in Proposition 2.1. The inf-sup condition [9, Lemma 4.3] for Orlicz spaces reads

$$(2.18) \quad \inf_{q \in L_0^{\varphi^*}(\Omega)} \sup_{\mathbf{v} \in \mathbf{W}_0^{1,\varphi}(\Omega)} \frac{\int_{\Omega} q \operatorname{div} \mathbf{v}}{\|\nabla \mathbf{v}\|_{\varphi} \|q\|_{\varphi^*}} =: \beta > 0,$$

where $\beta > 0$ only depends on r^{\pm} and Ω via its John constant. This inf-sup condition guarantees the existence of a unique pressure $p \in L_0^{\varphi^*}(\Omega)$ such that $(\mathbf{u}, p) \in \mathbf{W}_0^{1,\varphi}(\Omega) \times L_0^{\varphi^*}(\Omega)$ solves (2.15).

2.5. Mesh and finite element spaces. At this point we introduce our notation for the finite element spaces of the following sections. Let \mathcal{T} be a shape regular, face-to-face decomposition of Ω consisting of d -simplices $K \in \mathcal{T}$ such that $\bar{\Omega} = \bigcup_{K \in \mathcal{T}} K$. The faces of the elements in \mathcal{T} are denoted by \mathcal{F} . The subset of faces in the interior of Ω is denoted by \mathcal{F}_i and its complement of boundary faces by \mathcal{F}_{∂} . The skeleton of \mathcal{T} is $\Sigma := \bigcup_{F \in \mathcal{F}} F$.

For $K \in \mathcal{T}$ and $F \in \mathcal{F}$ we denote by h_K respectively h_F its diameter and we introduce, with some ambiguity of notation, the mesh-size function $h : \Omega \rightarrow \mathbb{R}^+$ by

$$h|_K := h_K, \quad \text{in } K \in \mathcal{T}, \quad \text{and} \quad h|_F := h_F, \quad \text{on } F \in \mathcal{F}$$

and let $h_{\max} = \|h\|_{L^{\infty}(\Omega)}$. Denoting by ρ_K , $K \in \mathcal{T}$, the supremum of the diameters of all balls inscribed in the respective K , the shape constant of \mathcal{T} is given by

$$(2.19) \quad \max_{K \in \mathcal{T}} \frac{h_K}{\rho_K} < \infty.$$

We emphasise that, in the following, constants hidden in \lesssim or \gtrsim may depend on the shape constant of \mathcal{T} but not on the mesh-size function.

Each face $F \in \mathcal{F}_i$ in the interior of Ω is oriented by a unit normal vector \mathbf{n}_F . We indicate by $\llbracket v \rrbracket_F$ and $\{\!\!\{ v \}\!\!\}_F$ the jump respectively average of v on the face $F \in \mathcal{F}_i$ in direction of \mathbf{n}_F . In particular, for $K_1, K_2 \in \mathcal{T}$ with $F = K_1 \cap K_2$ we define for $\mathbf{x} \in F$

$$\llbracket v \rrbracket_F(\mathbf{x}) = v|_{K_1}(\mathbf{x}) - v|_{K_2}(\mathbf{x}) \quad \text{and} \quad \{\!\!\{ v \}\!\!\}_F(\mathbf{x}) = \frac{v|_{K_1}(\mathbf{x}) + v|_{K_2}(\mathbf{x})}{2},$$

where \mathbf{n}_F points outside K_1 . Of course, the sign of $\llbracket v \rrbracket_F$ depends on the orientation of \mathbf{n}_F , which however is not relevant to our presentation. For boundary faces $F \in \mathcal{F}_{\partial}$, \mathbf{n}_F is oriented outside Ω and $\llbracket v \rrbracket_F$ equals the trace on F , i.e. for $K \in \mathcal{T}$ with $F = K \cap \partial\Omega$ we define $\llbracket v \rrbracket_F(\mathbf{x}) = v|_K(\mathbf{x})$ for $\mathbf{x} \in F$. We let $\mathbf{n} \in L^{\infty}(\Sigma)$ with $\mathbf{n}|_F = \mathbf{n}_F$, $F \in \mathcal{F}$.

For a differential operator \mathcal{D} , the notation \mathcal{D}_h is used for the broken version of \mathcal{D} , i.e.

$$(2.20) \quad (\mathcal{D}_h v)|_K := \mathcal{D}(v|_K) \quad \text{in } K, \quad K \in \mathcal{T}$$

for piecewise smooth v . E.g., the broken gradient of a piecewise $W^{1,r}$ -function v is given by $(\nabla_h v)|_K = \nabla(v|_K)$ for all $K \in \mathcal{T}$. For an element $K \in \mathcal{T}$ we define the set of neighbours \mathcal{N}_K and the neighbourhood ω_K by

$$(2.21a) \quad \mathcal{N}_K := \{K' \in \mathcal{T} \mid K' \cap K \neq \emptyset\}, \quad \omega_K := \text{interior} \left(\bigcup_{K' \in \mathcal{N}_K} K' \right).$$

The sets ω_K are open, bounded and connected. The shape constant (2.19) of \mathcal{T} implies the following properties of \mathcal{T} :

$$(2.21b) \quad |\omega_K| \approx |K| \approx h_K^d \quad \text{for all } K \in \mathcal{T} \text{ and } \#\mathcal{N}_K \leq m_0 \text{ for some } m_0 \in \mathbb{N}.$$

For $\ell \geq 0$ let $\mathbb{P}_\ell(K)$ and $\mathbb{P}_\ell(F)$ be the set of polynomials of degree at most ℓ on a d -simplex $K \in \mathcal{T}$ and a face $F \in \mathcal{F}$, respectively. For $k \in \mathbb{N}_0$, the space of $W^{k,1}$ -conforming element-wise polynomials of order $\ell \in \mathbb{N}_0$ on \mathcal{T} is defined by

$$(2.22) \quad S_\ell^k := S_\ell^k(\mathcal{T}) := \{v \in W^{k,1}(\Omega) \mid \forall K \in \mathcal{T} \ v|_K \in \mathbb{P}_\ell(K)\}.$$

The lowest-order Crouzeix-Raviart finite element is given as the subspace of functions in S_1^0 , for which the face mean values of jumps vanish, i.e.

$$(2.23) \quad \text{CR} := \text{CR}(\mathcal{T}) := \left\{ v \in S_1^0 \mid \forall F \in \mathcal{F} \int_F \llbracket v \rrbracket = 0 \right\}.$$

Note that $\int_F \llbracket v \rrbracket_F = \int_F v$ for boundary edges $F \in \mathcal{F}_\partial$. In other words, $\text{CR}(\mathcal{T})$ consists of all functions in $S_1^0(\mathcal{T})$ that are continuous at the barycenters of interior faces and vanish at the barycenters of boundary faces.

It is well known that broken Sobolev norms are definite on CR and they even dominate jumps across edges. As we could find this statement only for Hilbert norms, we provide a short proof in the required $W^{1,1}$ setting.

Lemma 2.9. *Let $v \in \text{CR}(\mathcal{T}) + W_0^{1,1}(\Omega)$, then for all $F \in \mathcal{F}$ we have*

$$\int_F h_F^{-1} |\llbracket v \rrbracket_F| \lesssim \sum_{\substack{K \in \mathcal{T} \\ F \subset K}} \int_K |\nabla v|.$$

The hidden constant depends only on the shape constant of \mathcal{T} .

Proof. Since $v \in \text{CR}(\mathcal{T}) + W_0^{1,1}(\Omega)$, the face mean value $\langle v \rangle = \int_F v$ takes the same value on both sides of the face is therefore well defined. If $F \subset \partial\Omega$, then $\langle v \rangle_F = 0$.

We first consider interior faces $F \in \mathcal{F}_i$, i.e. there exist $K_1, K_2 \in \mathcal{T}$ with $F = K_1 \cap K_2$. Then the embedding $W^{1,1}(K_i) \hookrightarrow L^1(F)$, $i = 1, 2$, and a Poincaré inequality imply

$$(2.24) \quad \int_F |\llbracket v \rrbracket_F| \leq \sum_{i=1}^2 \int_F |v|_{K_i} - \langle v \rangle_F \leq 2 \sum_{i=1}^2 \int_F |v|_{K_i} - \langle v \rangle_{K_i} \lesssim \sum_{i=1}^2 \int_{K_i} |\nabla v|.$$

The case $F \subset \partial\Omega$ is similar using $\langle v \rangle_F = 0$. \square

The *Crouzeix-Raviart* interpolation operator $\mathcal{I}^{\text{cr}} : W^{1,1}(\Omega) \rightarrow \text{CR}$ is defined by

$$\int_F \mathcal{I}^{\text{cr}} v = \int_F v \quad \forall F \in \mathcal{F}^i.$$

Element-wise integration by parts yields for all $K \in \mathcal{T}$ and $v \in W^{1,1}(\Omega)$

$$(2.25) \quad \nabla_h(\mathcal{I}^{\text{cr}} v)|_K = \int_K \nabla_h(\mathcal{I}^{\text{cr}} v) = \int_K \nabla_h v$$

i.e., \mathcal{I}^{cr} preserves averages of first derivatives.

Using Jensen's inequality and the fact that $\mathcal{D}_h(\mathcal{I}^{\text{cr}}\mathbf{v})$ is piecewise constant, we conclude from (2.25) that for any N-function ψ we have

$$(2.26) \quad \int_K \psi(|\mathcal{D}_h(\mathcal{I}^{\text{cr}}\mathbf{v})|) = \int_K \psi\left(\left|\int_K \mathcal{D}\mathbf{v}(\mathbf{y}) \, d\mathbf{y}\right|\right) \leq \int_K \psi(|\mathcal{D}\mathbf{v}|).$$

Thanks to the definition of the Luxembourg norm, this readily implies $\|\mathcal{D}_h\mathcal{I}^{\text{cr}}\mathbf{v}\|_\psi \leq \|\mathcal{D}\mathbf{v}\|_\psi$, i.e. \mathcal{I}^{cr} is $\mathbf{W}^{1,\psi}$ -stable with constant one. Moreover, we even have that $\mathcal{I}^{\text{cr}}\mathbf{v}$ is a quasi-optimal approximation of \mathbf{v} in the natural distance.

Corollary 2.10. *For any $\mathbf{v} \in \mathbf{W}_0^{1,\varphi}(\Omega)$ and $K \in \mathcal{T}$, we have*

$$\|\mathbf{F}(\mathcal{D}\mathbf{v}) - \mathbf{F}(\mathcal{D}_h\mathcal{I}^{\text{cr}}\mathbf{v})\|_{2;K} \lesssim \inf_{\mathbf{Q}_K \in \mathbb{R}^{d \times d}} \|\mathbf{F}(\mathcal{D}\mathbf{v}) - \mathbf{Q}_K\|_{2;K}.$$

The hidden constant depends only on r^\pm .

Proof. Note that $\mathcal{D}_h(\mathcal{I}^{\text{cr}}\mathbf{v})|_K = \int_K \mathcal{D}\mathbf{v}$ due to (2.25) and use Lemma 2.8. \square

With some ambiguity of notation, we extend all the terms introduced above to vector-valued functions by applying them component by component and possibly adding them up, e.g. in the case of norms.

Following [11], we discretize (2.15) with nonconforming Crouzeix-Raviart functions for the velocity and piecewise constant functions for the pressure

$$(2.27) \quad \mathbf{CR} := (\text{CR})^d \quad \text{and} \quad \widehat{S}_0^0 := S_0^0(\mathcal{T}) \cap L_0^{\varphi^*}(\Omega).$$

This pair is inf-sup stable. Indeed, the interpolant \mathcal{I}^{cr} is a bounded Fortin operator, according to (2.25) and (2.26). This observation and (2.18) imply

$$(2.28) \quad \inf_{q_h \in \widehat{S}_0^0} \sup_{\mathbf{v}_h \in \mathbf{CR}} \frac{\int_\Omega q_h \operatorname{div}_h \mathbf{v}_h}{\|\nabla_h \mathbf{v}_h\|_\varphi \|q_h\|_{\varphi^*}} \geq \beta > 0.$$

2.6. Smoothing operator. Note that the Crouzeix-Raviart space is nonconforming, i.e. $\mathbf{CR} \not\subset \mathbf{W}_0^{1,\varphi}(\Omega)$ in view of the lack of global continuity and the violation of the boundary conditions. Consequently the datum $\mathbf{f} \in \mathbf{W}^{-1,\varphi^*}(\Omega)$ in (2.15) cannot be directly applied on test-functions from \mathbf{CR} . Still, it is known from [44] that an error estimate like (1.4), not involving extra regularity of \mathbf{u} , can hold true only for discretizations defined for all possible data \mathbf{f} . Therefore, we resort to a so-called *smoothing operator* $\mathbf{E} : \mathbf{CR} \rightarrow \mathbf{W}_0^{1,\infty}(\Omega) \subset \mathbf{W}_0^{1,\varphi}(\Omega)$, which acts on the test-functions, thus enabling the application of \mathbf{f} .

We make use of the smoothing operator introduced in [47, §4.1] for the Stokes equations (see also [34, §3]). The operator acts on $\mathbf{v}_h \in \mathbf{CR}$ as follows

$$\mathbf{E}\mathbf{v}_h = \mathbf{A}\mathbf{v}_h + \mathbf{B}\mathbf{v}_h + \mathbf{C}\mathbf{v}_h$$

where

- $\mathbf{A} : \mathbf{CR} \rightarrow (S_1^1)^d$ is an averaging operator, enabling the stability estimates (2.30) below;
- \mathbf{B} maps \mathbf{v}_h to a combination of face bubbles, so as to enforce the identity

$$(2.29a) \quad \int_F \mathbf{E}\mathbf{v}_h = \int_F \mathbf{v}_h \quad \forall F \in \mathcal{F}$$

which implies

$$(2.29b) \quad \int_\Omega \mathbf{Q}_h : \nabla_h \mathbf{v}_h = \int_\Omega \mathbf{Q}_h : \nabla \mathbf{E}\mathbf{v}_h \quad \forall \mathbf{Q}_h \in (S_0^0)^{d \times d},$$

in combination with element-wise integration by parts;

- \mathbf{C} maps \mathbf{v}_h to a combination of element bubbles, so as to enforce the identity

$$(2.29c) \quad \operatorname{div} \mathbf{E}\mathbf{v}_h = \operatorname{div}_h \mathbf{v}_h.$$

We refer to [47, 34] for the details of the construction and the proof of (2.29).

Remark 2.11 (Pressure robustness). *Owing to (2.29c), we readily infer*

$$\operatorname{div}_h \mathbf{v}_h = 0 \quad \implies \quad \operatorname{div} \mathbf{E} \mathbf{v}_h = 0$$

i.e. \mathbf{E} maps element-wise divergence-free functions to exactly divergence-free functions. This property, originally pointed out in [37], guarantees the pressure robustness of the discretizations proposed in the next sections, meaning that the discrete velocity is independent of the discrete pressure. This property is necessary to make sure that the velocity-error is independent of the pressure error, thus preventing from sub-optimal error decay rates as in (1.6).

Lemma 2.12 (Stability of \mathbf{E}). *The operator $\mathbf{E} : \mathbf{CR} \rightarrow \mathbf{W}_0^{1,\infty}(\Omega)$ from [47, 34] satisfies (2.29) and the local stability properties*

$$(2.30a) \quad \|\nabla \mathbf{E} \mathbf{v}_h\|_{1;K} \lesssim \|\nabla_h \mathbf{v}_h\|_{1;\omega_K} \quad \forall K \in \mathcal{T},$$

$$(2.30b) \quad \|\mathcal{D}_h(\mathbf{w}_h + \mathbf{E} \mathbf{v}_h)\|_{\infty;K} \lesssim \int_K |\mathcal{D}_h(\mathbf{w}_h + \mathbf{E} \mathbf{v}_h)| \quad \forall K \in \mathcal{T},$$

$$(2.30c) \quad \|\mathcal{D}_h(\mathbf{v}_h - \mathbf{E} \mathbf{v}_h)\|_{1;K} \lesssim \sum_{\substack{F \in \mathcal{F} \\ F \cap K \neq \emptyset}} \int_F |[\![\mathbf{v}_h]\!]_F| \quad \forall K \in \mathcal{T},$$

for $\mathbf{w}_h, \mathbf{v}_h \in \mathbf{CR}$ and $\mathcal{D} \in \{\nabla, \mathbf{D}\}$. The hidden constants depend only on the shape constant of \mathcal{T} .

Proof. Properties (2.29) are proved in [47, §4.1].

The bound (2.30c) is proved in [34, proposition 18] for L^2 -type norms and $\mathcal{D} = \nabla$. Recalling that $|\mathbf{D}(\mathbf{v}_h - \mathbf{E} \mathbf{v}_h)| \leq |\nabla(\mathbf{v}_h - \mathbf{E} \mathbf{v}_h)|$, the claimed bound follows from scaling.

Assertion (2.30a) follows from a triangle inequality

$$\|\nabla \mathbf{E} \mathbf{v}_h\|_{1;K} \leq \|\nabla(\mathbf{E} \mathbf{v}_h - \mathbf{v}_h)\|_{1;K} + \|\nabla \mathbf{v}_h\|_{1;K}$$

with (2.30c) and Lemma 2.9.

Since by the above indicated construction, \mathbf{E} maps to an affine equivalent finite element, therefore property (2.30b) follows from equivalence of norms on finite dimensional spaces and standard scaling of affine equivalent elements. \square

The local $\mathbf{W}^{1,1}$ -stability of \mathbf{E} (2.30a) implies local $\mathbf{W}^{1,\varphi}$ -stability in the following sense.

Lemma 2.13 (Local $W^{1,\psi}$ -stability of \mathbf{E}). *Let ψ be a uniformly convex N -function. Then $\mathbf{E} : \mathbf{CR} \rightarrow \mathbf{W}_0^{1,\infty}(\Omega)$ from [47, 34] satisfies*

$$(2.31) \quad \int_K \psi(|\nabla \mathbf{E} \mathbf{v}_h|) \lesssim \int_{\omega_K} \psi(|\nabla_h \mathbf{v}_h|) \quad \forall K \in \mathcal{T}.$$

The hidden constant only depends on the indices of uniform convexity of ψ and on the shape constant of \mathcal{T} .

Proof. Using (2.30b), the local $\mathbf{W}^{1,1}$ -stability (2.30a) of \mathbf{E} together with (2.21) and the Δ_2 -condition, we estimate for all $K \in \mathcal{T}$ and $\mathbf{v}_h \in \mathbf{CR}$

$$(2.32) \quad \int_K \psi(|\nabla \mathbf{E} \mathbf{v}_h|) \lesssim \int_K \psi\left(\int_K |\nabla \mathbf{E} \mathbf{v}_h|\right) \lesssim \int_K \psi\left(\int_{\omega_K} |\nabla_h \mathbf{v}_h|\right) \lesssim \int_{\omega_K} \psi(|\nabla_h \mathbf{v}_h|).$$

Here we have used Jensen's inequality for convex functions in the last step. \square

Lemma 2.13 implies in particular, that \mathbf{E} is $\mathbf{W}^{1,\varphi}$ -stable. This and the fact that only finite many of the ω_K , $K \in \mathcal{T}$, overlap imply for all $\mathbf{v}_h \in \mathbf{CR}$ that

$$(2.33) \quad \int_{\Omega} \varphi(|\nabla \mathbf{E} \mathbf{v}_h|) \lesssim \int_{\Omega} \varphi(|\nabla_h \mathbf{v}_h|) \quad \text{resp.} \quad \|\nabla \mathbf{E} \mathbf{v}_h\|_{\varphi} \lesssim \|\nabla_h \mathbf{v}_h\|_{\varphi}.$$

Finally, we observe that \mathbf{E} is a right inverse of the Crouzeix-Raviart interpolation operator as a consequence of (2.29a), cf. [45, Lemma 3.2].

Corollary 2.14. *For all $K \in \mathcal{T}$ and $\mathbf{v}_h \in \mathbf{CR}(\mathcal{T})$ we have*

$$\mathcal{I}^{\mathbf{C}\mathbf{R}} \mathbf{E} \mathbf{v}_h = \mathbf{v}_h \quad \text{and} \quad \int_K \nabla \mathbf{E} \mathbf{v}_h = \int_K \nabla \mathbf{v}_h = \nabla_h \mathbf{v}_h|_K \quad \forall K \in \mathcal{T}.$$

3. A PRESSURE ROBUST CROUZEIX-RAVIART METHOD FOR $\mathbf{D} = \nabla$.

In this section we propose a pressure robust Crouzeix-Raviart method for the nonlinear Stokes equation (1.1) for the case $\mathbf{D} = \nabla$. The method coincides with the ones in [47, 34] for $\varphi(t) = t^2$, i.e. for the linear Stokes equations.

Method 3.1. *For $\mathbf{f} \in \mathbf{W}^{-1,\varphi^*}(\Omega)$ compute $\mathbf{u}_h \in \mathbf{CR}$ and $p_h \in \widehat{S}_0^0$ such that*

$$\begin{aligned} \forall \mathbf{v}_h \in \mathbf{CR} \quad & \int_{\Omega} \mathbf{A}(\nabla_h \mathbf{u}_h) : \nabla_h \mathbf{v}_h - \int_{\Omega} p_h \operatorname{div}_h \mathbf{v}_h = \langle \mathbf{f}, \mathbf{E} \mathbf{v}_h \rangle, \\ \forall q_h \in \widehat{S}_0^0 \quad & \int_{\Omega} q_h \operatorname{div}_h \mathbf{u}_h = 0. \end{aligned}$$

Note that we do not require regularity beyond $\mathbf{f} \in \mathbf{W}^{-1,\varphi^*}(\Omega)$ from (2.15), because $\mathbf{E} \mathbf{v}_h \in \mathbf{W}_0^{1,\infty}(\Omega)$. Similarly to the derivation of (2.17), we have that \mathbf{u}_h is in the space of element-wise divergence-free functions

$$(3.1) \quad \mathbf{Z}_h := \{\mathbf{v}_h \in \mathbf{CR} \mid \operatorname{div}_h \mathbf{v}_h = 0\} \subset \mathbf{CR},$$

and uniquely solves the reduced problem

$$(3.2) \quad \forall \mathbf{v}_h \in \mathbf{Z}_h \quad \int_{\Omega} \mathbf{A}(\nabla_h \mathbf{u}_h) : \nabla_h \mathbf{v}_h = \langle \mathbf{f}, \mathbf{E} \mathbf{v}_h \rangle.$$

In view of the stability (2.33) and the inf-sup-condition (2.28), the well-posedness of Method 3.1 follows from standard monotone operator theory; cf. [9, 26].

Proposition 3.2 (Well-posedness). *Method 3.1 determines a unique pair $(\mathbf{u}_h, p_h) \in \mathbf{CR} \times \widehat{S}_0^0$ depending continuously on the datum $\mathbf{f} \in \mathbf{W}^{-1,\varphi^*}(\Omega)$.*

From the properties of \mathbf{E} and the φ -structure of \mathbf{A} we infer our first main result addressing the approximation properties of Method 3.1.

Theorem 3.3 (Velocity-error). *Let $(\mathbf{u}, p) \in \mathbf{W}_0^{1,\varphi}(\Omega) \times L_0^{\varphi^*}(\Omega)$ be the solution to (2.15) and let $(\mathbf{u}_h, p_h) \in \mathbf{CR} \times \widehat{S}_0^0$ be computed by Method 3.1. Then we have*

$$(3.3) \quad \|\mathbf{F}(\nabla \mathbf{u}) - \mathbf{F}(\nabla_h \mathbf{u}_h)\|_2^2 \lesssim \sum_{K \in \mathcal{T}} \inf_{\mathbf{Q}_K \in \mathbf{R}^{d \times d}} \|\mathbf{F}(\nabla \mathbf{u}) - \mathbf{Q}_K\|_{2;\omega_K}^2.$$

In particular, when $\mathbf{F}(\nabla \mathbf{u}) \in \mathbf{W}^{1,2}(\Omega)$, we have

$$(3.4) \quad \|\mathbf{F}(\nabla \mathbf{u}) - \mathbf{F}(\nabla_h \mathbf{u}_h)\|_2 \lesssim \|h \nabla \mathbf{F}(\nabla \mathbf{u})\|_2.$$

The hidden constants depend only on Ω , r^{\pm} , and the shape constant of \mathcal{T} .

Before we prove Theorem 3.3, some remarks are in order.

Remark 3.4 (Quasi-optimality). *Recall that, for conforming and divergence-free discretizations, the velocity-error estimate (1.4) is derived from the quasi-optimality stated in (1.3). Theorem 3.3 establishes a counterpart of (1.4) for Method 3.1, but quasi-optimality in the sense of (1.3) is not guaranteed. The critical aspect is the stability of the smoothing operator \mathbf{E} , as Lemma 2.13 for $\psi = \varphi_a$, requires a constant shift $a \geq 0$ on element neighbourhoods. Enforcing this through a change of the shift (2.9) introduces a suboptimal approximation term. For details consider the bound of I_2 in the proof of Theorem 3.3 below.*

Remark 3.5 (Rate of convergence and regularity). *We emphasize that two aspects of the velocity-error estimate (3.4) improve upon the existing results in [9] for the MINI element and in [26] for a stabilized $\mathbf{Q}_1/\mathbf{Q}_1$ discretization. First, the convergence rate for the velocity is optimal, whereas for problems with (r, ε) -structure, suboptimal rates are observed in [9, 26] both theoretically and in numerical simulations for $r > 2$, cf. (1.6). Second, no additional regularity of the pressure is necessary beyond the minimal requirement $p \in L_0^{\varphi^*}(\Omega)$ from (2.15). Both improvements are made possible by the pressure robustness observed in Remark 2.11.*

An $\mathcal{O}(h)$ velocity-error estimate can be found in [30, Corollary 4.2] for a local dG discretization, however, compared to (3.4) this result has several disadvantages. First, the discrete solution is obtained via a pseudo-monotone scheme, which does not guarantee uniqueness of the solution. Second, the constant hidden in $\mathcal{O}(h)$ is not robust with respect to ε and is unbounded as $\varepsilon \rightarrow 0$, and third, the additional regularity $\mathbf{f} \in \mathbf{L}^2(\Omega)$ is required. In contrast, Theorem 3.3 invokes only the minimal regularity $\mathbf{F}(\nabla \mathbf{u}) \in W^{1,2}(\Omega)^{d \times d}$ that is necessary according to the approximation theory.

The regularity for system (2.15) has not been fully developed. However, it is expected that this kind of regularity is the natural one. Results for the special case of (r, ε) -structure can be found in [8, 46, 7]. The case without pressure but for all uniformly convex N -functions is considered in [6, Theorem 2.4].

Proof of Theorem 3.3. We abbreviate $\mathbf{v}_h = \mathcal{I}^{\text{cr}} \mathbf{u} \in \mathbf{CR}$. We use Lemma 2.6 to estimate the error by

$$\begin{aligned} \|\mathbf{F}(\nabla \mathbf{u}) - \mathbf{F}(\nabla_h \mathbf{u}_h)\|_2^2 &\approx \int_{\Omega} (\mathbf{A}(\nabla_h \mathbf{u}_h) - \mathbf{A}(\nabla \mathbf{u})) : (\nabla_h \mathbf{u}_h - \nabla \mathbf{u}) \\ &= \underbrace{\int_{\Omega} (\mathbf{A}(\nabla_h \mathbf{u}_h) - \mathbf{A}(\nabla \mathbf{u})) : (\nabla_h \mathbf{v}_h - \nabla \mathbf{u})}_{=: I_1} \\ &\quad + \underbrace{\int_{\Omega} (\mathbf{A}(\nabla_h \mathbf{u}_h) - \mathbf{A}(\nabla \mathbf{u})) : (\nabla_h \mathbf{u}_h - \nabla_h \mathbf{v}_h)}_{=: I_2}. \end{aligned}$$

We bound the first term with the Young-type inequality (2.4) and Lemma 2.6 by

$$\begin{aligned} I_1 &\leq \int_{\Omega} \delta (\varphi_{|\nabla \mathbf{u}|})^* (|\mathbf{A}(\nabla_h \mathbf{u}_h) - \mathbf{A}(\nabla \mathbf{u})|) + \int_{\Omega} C_{\delta} \varphi_{|\nabla \mathbf{u}|} (|\nabla_h \mathbf{v}_h - \nabla \mathbf{u}|) \\ &\lesssim \delta \|\mathbf{F}(\nabla \mathbf{u}) - \mathbf{F}(\nabla_h \mathbf{u}_h)\|_2^2 + C_{\delta} \|\mathbf{F}(\nabla \mathbf{u}) - \mathbf{F}(\nabla_h \mathbf{v}_h)\|_2^2. \end{aligned}$$

For estimating the second term on the right-hand side, observe that the second equation of Method 3.1 and (2.25) imply $\text{div}_h(\mathbf{u}_h - \mathbf{v}_h)|_K = -\mathbf{f}_K \text{div} \mathbf{u} = 0$ for all $K \in \mathcal{T}$. This and (2.29c) reveal $\text{div} \mathbf{E}(\mathbf{u}_h - \mathbf{v}_h) = 0$. Hence, the first equation of

Method 3.1 and (2.15) imply

$$\begin{aligned}
I_2 &= \langle \mathbf{f}, \mathbf{E}(\mathbf{u}_h - \mathbf{v}_h) \rangle - \int_{\Omega} \mathbf{A}(\nabla \mathbf{u}) : (\nabla_h \mathbf{u}_h - \nabla_h \mathbf{v}_h) \\
&= \int_{\Omega} \mathbf{A}(\nabla \mathbf{u}) : (\nabla \mathbf{E}(\mathbf{u}_h - \mathbf{v}_h) - \nabla_h(\mathbf{u}_h - \mathbf{v}_h)) \\
&= \int_{\Omega} (\mathbf{A}(\nabla \mathbf{u}) - \mathbf{A}(\nabla_h \mathbf{v}_h)) : (\nabla \mathbf{E}(\mathbf{u}_h - \mathbf{v}_h) - \nabla_h(\mathbf{u}_h - \mathbf{v}_h)),
\end{aligned}$$

where we have used (2.29b) together with $\mathbf{A}(\nabla_h \mathbf{v}_h) \in (S_0^0)^{d \times d}$ in the last step.

By the Young-type inequality (2.4), the quasi-triangle inequality (2.3), and Lemma 2.6, we conclude for $\delta > 0$ that

$$\begin{aligned}
I_2 &\leq 2 \int_{\Omega} C_{\delta} (\varphi_{|\nabla \mathbf{u}|})^* (|\mathbf{A}(\nabla \mathbf{u}) - \mathbf{A}(\nabla_h \mathbf{v}_h)|) \\
&\quad + \int_{\Omega} \delta \varphi_{|\nabla \mathbf{u}|} (|\nabla_h(\mathbf{u}_h - \mathbf{E}\mathbf{u}_h)|) + \int_{\Omega} \delta \varphi_{|\nabla \mathbf{u}|} (|\nabla_h(\mathbf{v}_h - \mathbf{E}\mathbf{v}_h)|) \\
&\approx C_{\delta} \|\mathbf{F}(\nabla \mathbf{u}) - \mathbf{F}(\nabla_h \mathbf{v}_h)\|_2^2 \\
&\quad + \int_{\Omega} \delta \varphi_{|\nabla \mathbf{u}|} (|\nabla_h(\mathbf{u}_h - \mathbf{E}\mathbf{u}_h)|) + \int_{\Omega} \delta \varphi_{|\nabla \mathbf{u}|} (|\nabla_h(\mathbf{v}_h - \mathbf{E}\mathbf{v}_h)|).
\end{aligned}$$

The first term on the right-hand side is already fine and the latter terms can be treated in the same way. For $\mathbf{Q}_K \in \mathbb{R}^{d \times d}$ we first change the shift with (2.9) from $|\nabla \mathbf{u}|$ to $|\mathbf{Q}_K|$ to obtain

$$\int_K \varphi_{|\nabla \mathbf{u}|} (|\nabla_h(\mathbf{v}_h - \mathbf{E}\mathbf{v}_h)|) \lesssim \int_K \varphi_{|\mathbf{Q}_K|} (|\nabla_h(\mathbf{v}_h - \mathbf{E}\mathbf{v}_h)|) + \int_K |\mathbf{F}(\nabla \mathbf{u}) - \mathbf{F}(\mathbf{Q}_K)|^2.$$

Next, Lemma 2.12 and Lemma 2.9 yield that

$$\begin{aligned}
\|\nabla_h(\mathbf{v}_h - \mathbf{E}\mathbf{v}_h)\|_{\infty; K} &\lesssim \int_K |\nabla_h(\mathbf{v}_h - \mathbf{E}\mathbf{v}_h)| = \sum_{\substack{F \in \mathcal{F} \\ F \cap K \neq \emptyset}} \int_F h^{-1} \|[\mathbf{v}_h - \mathbf{u}]_F\| \\
&\lesssim \sum_{\substack{F \in \mathcal{F} \\ F \cap K \neq \emptyset}} \sum_{\substack{K' \in \mathcal{T} \\ F \subset K'}} \int_{K'} |\nabla_h(\mathbf{v}_h - \mathbf{u})| \approx \int_{\omega_K} |\nabla_h(\mathbf{v}_h - \mathbf{u})|.
\end{aligned}$$

The last estimate follows from

$$\omega_K = \bigcup \{K' \in \mathcal{T} : F \subset K' \text{ for some } F \in \mathcal{F} \text{ with } F \cap K \neq \emptyset\}$$

and the fact that every element $K' \subset \omega_K$ appears in previous sum at least once and at most $(d+1)$ -times. Consequently, recalling $|K| \approx |\omega_K|$, we obtain with Jensen's inequality and a shift-change back to $|\nabla \mathbf{u}|$, that

$$\begin{aligned}
&\int_K \varphi_{|\nabla \mathbf{u}|} (|\nabla_h(\mathbf{v}_h - \mathbf{E}\mathbf{v}_h)|) \\
(3.5) \quad &\lesssim \int_{\omega_K} \varphi_{|\mathbf{Q}_K|} (|\nabla_h(\mathbf{v}_h - \mathbf{u})|) + \int_K |\mathbf{F}(\nabla \mathbf{u}) - \mathbf{F}(\mathbf{Q}_K)|^2 \\
&\lesssim \int_{\omega_K} \varphi_{|\nabla \mathbf{u}|} (|\nabla_h(\mathbf{v}_h - \mathbf{u})|) + \int_{\omega_K} |\mathbf{F}(\nabla \mathbf{u}) - \mathbf{F}(\mathbf{Q}_K)|^2 \\
&\approx \int_{\omega_K} |\mathbf{F}(\nabla_h \mathbf{v}_h) - \mathbf{F}(\nabla \mathbf{u})|^2 + \int_{\omega_K} |\mathbf{F}(\nabla \mathbf{u}) - \mathbf{F}(\mathbf{Q}_K)|^2.
\end{aligned}$$

Estimating the term $\int_K \varphi_{|\nabla \mathbf{u}|} (|\nabla_h(\mathbf{u}_h - \mathbf{E}\mathbf{u}_h)|)$ in exactly the same way, and summing over all $K \in \mathcal{T}$, we arrive at

$$\begin{aligned} I_2 &\lesssim C_\delta \|\mathbf{F}(\nabla \mathbf{u}) - \mathbf{F}(\nabla_h \mathbf{v}_h)\|_2^2 \\ &\quad + \delta \int_\Omega |\mathbf{F}(\nabla_h \mathbf{v}_h) - \mathbf{F}(\nabla \mathbf{u})|^2 + \delta \int_\Omega |\mathbf{F}(\nabla_h \mathbf{u}_h) - \mathbf{F}(\nabla \mathbf{u})|^2 \\ &\quad + \delta \sum_{K \in \mathcal{T}} \inf_{\mathbf{Q}_K \in \mathbb{R}^{d \times d}} \int_{\omega_K} \varphi_{|\nabla \mathbf{u}|} (|\nabla \mathbf{u} - \mathbf{Q}_K|) \end{aligned}$$

since $\mathbf{Q}_K \in \mathbb{R}^{d \times d}$ was arbitrary and the sets ω_K , $K \in \mathcal{T}$, overlap only finitely often.

Using again Lemma 2.6, combining the previous estimates, we arrive at

$$\begin{aligned} \|\mathbf{F}(\nabla \mathbf{u}) - \mathbf{F}(\nabla_h \mathbf{u}_h)\|_2^2 &\lesssim (C_\delta + \delta) \|\mathbf{F}(\nabla \mathbf{u}) - \mathbf{F}(\nabla_h \mathbf{v}_h)\|_2^2 \\ &\quad + \delta \|\mathbf{F}(\nabla \mathbf{u}) - \mathbf{F}(\nabla_h \mathbf{u}_h)\|_2^2 \\ &\quad + \delta \sum_{K \in \mathcal{T}} \inf_{\mathbf{Q}_K \in \mathbb{R}^{d \times d}} \|\mathbf{F}(\nabla \mathbf{u}) - \mathbf{F}(\mathbf{Q}_K)\|_{2;\omega_K}^2. \end{aligned}$$

Choosing δ sufficiently small, the second term can be compensated on the left hand side and the assertion follows from Corollary 2.10 recalling $\mathbf{v}_h = \mathcal{I}^{\mathbf{c}\mathbf{r}} \mathbf{u}$. \square

Remark 3.6. Note that we also have (3.3) for the post-processed approximation $\mathbf{E}\mathbf{u}_h$ instead of \mathbf{u}_h . Indeed, thanks to Lemma 2.6 and the quasi triangle inequality (2.3), we have

$$\|\mathbf{F}(\nabla \mathbf{u}) - \mathbf{F}(\nabla \mathbf{E}\mathbf{u}_h)\|_2^2 \lesssim \|\mathbf{F}(\nabla \mathbf{u}) - \mathbf{F}(\nabla_h \mathbf{u}_h)\|_2^2 + \int_\Omega \varphi_{|\nabla \mathbf{u}|} (|\nabla_h \mathbf{u}_h - \nabla \mathbf{E}\mathbf{u}_h|).$$

The claim follows estimating the former term on the right hand side by (3.3) and the latter term as in the proof of Theorem 3.3; compare with (3.5).

4. A PRESSURE ROBUST CROUZEIX-RAVIART METHOD FOR $\mathcal{D} = \mathbf{D}$

Recall the symmetric gradient $\mathbf{D}\mathbf{v} := \frac{1}{2}(\nabla \mathbf{v} + \nabla \mathbf{v}^T)$. According to [1], on some choices of Ω and \mathcal{T} , there exists $\mathbf{0} \neq \mathbf{v}_h \in \mathbf{CR}$ with $\mathbf{D}_h \mathbf{v}_h = \mathbf{0}$, where \mathbf{D}_h is the element-wise symmetric gradient; compare with (2.20).

This observation means that we cannot just apply Method 3.1 with ∇_h replaced by \mathbf{D}_h since the quasi-linear operator is not strictly monotone in general. We propose the following alternative, inspired by the so-called recovered finite elements [21].

Method 4.1. For $\mathbf{f} \in \mathbf{W}^{-1,\varphi^*}(\Omega)$ find $\mathbf{u}_h \in \mathbf{CR}$ and $p_h \in \widehat{S}_0^0$ such that

$$\begin{aligned} \forall \mathbf{v}_h \in \mathbf{CR} \quad &\int_\Omega \mathbf{A}(\mathbf{D}\mathbf{E}\mathbf{u}_h) : \mathbf{D}\mathbf{E}\mathbf{v}_h - \int_\Omega p_h \operatorname{div}_h \mathbf{v}_h = \langle \mathbf{f}, \mathbf{E}\mathbf{v}_h \rangle \\ \forall q_h \in \widehat{S}_0^0 \quad &\int_\Omega q_h \operatorname{div}_h \mathbf{u}_h = 0. \end{aligned}$$

Remark 4.2 (Equivalent conforming method). Method 4.1 is equivalent to a conforming and divergence-free discretization of (2.15): find $\tilde{\mathbf{u}}_h \in \mathbf{E}(\mathbf{CR})$ and $\tilde{p}_h \in \widehat{S}_0^0$ such that

$$\begin{aligned} \forall \tilde{\mathbf{v}}_h \in \mathbf{E}(\mathbf{CR}) \quad &\int_\Omega \mathbf{A}(\mathcal{D}\tilde{\mathbf{u}}_h) : \mathcal{D}\tilde{\mathbf{v}}_h - \int_\Omega \tilde{p}_h \operatorname{div} \tilde{\mathbf{v}}_h = \langle \mathbf{f}, \tilde{\mathbf{v}}_h \rangle, \\ \forall \tilde{q} \in \widehat{S}_0^0 \quad &\int_\Omega \tilde{q} \operatorname{div} \tilde{\mathbf{u}}_h = 0. \end{aligned}$$

Indeed, we have $\tilde{\mathbf{u}}_h = \mathbf{E}\mathbf{u}_h$ and $\tilde{p}_h = p_h$. The equivalence is obtained by (2.29c). Though the velocity space $\mathbf{E}(\mathbf{CR})$ appears to have a complicated structure, it can

be parametrised by \mathbf{CR} via \mathbf{E} , which is one-to-one according to the first part of Corollary 2.14.

In analogy with (2.17) and (3.2), we can characterize the velocity \mathbf{u}_h from Method 4.1 as the unique solution in \mathbf{Z}_h of the reduced problem

$$(4.2) \quad \forall \mathbf{v}_h \in \mathbf{Z}_h \quad \int_{\Omega} \mathbf{A}(\mathbf{DE}\mathbf{u}_h) : \mathbf{DE}\mathbf{v}_h = \langle \mathbf{f}, \mathbf{E}\mathbf{v}_h \rangle$$

which is equivalent to a conforming Galerkin discretization of (2.17), cf. Remark 4.2. Consequently, Method 4.1 is uniquely solvable by standard monotone operator theory and the inf-sup-condition (2.28). For an a priori bound of the velocity \mathbf{u}_h , we additionally recall the stability (2.33) of \mathbf{E} as well as the inverse bound

$$\int_{\Omega} \varphi(|\nabla_h \mathbf{v}_h|) = \int_{\Omega} \varphi(|\nabla_h \mathcal{I}^c \mathbf{E}\mathbf{v}_h|) \leq \int_{\Omega} \varphi(|\nabla \mathbf{E}\mathbf{v}_h|) \lesssim \int_{\Omega} \varphi(|\mathbf{DE}\mathbf{v}_h|) \quad \forall \mathbf{v}_h \in \mathbf{CR}.$$

resulting from Corollary 2.14, the stability of the Crouzeix-Raviart interpolation operator (2.26) and Korn's inequality from Proposition 2.1.

Proposition 4.3 (Well-posedness). *Method 4.1 determines a unique pair $(\mathbf{u}_h, p_h) \in \mathbf{CR} \times \widehat{S}_0^0$, which depends continuously on $\mathbf{f} \in \mathbf{W}^{-1, \varphi^*}(\Omega)$.*

The equivalence of Method 4.1 with a conforming and divergence-free discretization implies also that we have quasi-optimality in the sense of (1.3).

Theorem 4.4 (Approximation by $\mathbf{E}\mathbf{u}_h$). *Let $(\mathbf{u}, p) \in \mathbf{W}_0^{1, \varphi}(\Omega) \times L_0^{\varphi^*}(\Omega)$ be the solution of (2.15) and let $(\mathbf{u}_h, p_h) \in \mathbf{CR} \times \widehat{S}_0^0$ be computed by Method 4.1. Then*

$$\|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{DE}\mathbf{u}_h)\|_2 \lesssim \inf_{\mathbf{v}_h \in \mathbf{Z}_h} \|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{DE}\mathbf{v}_h)\|_2.$$

The hidden constant depends only on r^{\pm} .

Proof. Let $\mathbf{v}_h \in \mathbf{Z}_h$ be arbitrary. We employ Lemma 2.6, the scaled Young inequality (2.4) and the Galerkin orthogonality of Method 4.1 to obtain that

$$\begin{aligned} \|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{DE}\mathbf{u}_h)\|_2^2 &\lesssim \int_{\Omega} (\mathbf{A}(\mathbf{D}\mathbf{u}) - \mathbf{A}(\mathbf{DE}\mathbf{u}_h)) : (\mathbf{D}\mathbf{u} - \mathbf{DE}\mathbf{v}_h) \\ &\lesssim \int_{\Omega} \delta (\varphi_{|\mathbf{D}\mathbf{u}|})^* (|\mathbf{A}(\mathbf{D}\mathbf{u}) - \mathbf{A}(\mathbf{DE}\mathbf{u}_h)|) \\ &\quad + \int_{\Omega} C_{\delta} \varphi_{|\mathbf{D}\mathbf{u}|} (|\mathbf{D}\mathbf{u} - \mathbf{DE}\mathbf{v}_h|) \\ &\approx \delta \|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{DE}\mathbf{u}_h)\|_2^2 \\ &\quad + C_{\delta} \|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{DE}\mathbf{v}_h)\|_2^2. \end{aligned}$$

The claimed bound follows for $\delta > 0$ small enough. \square

If we consider the approximation of \mathbf{u} by \mathbf{u}_h instead of $\mathbf{E}\mathbf{u}_h$, we have the following a priori error bounds. Note that, in this case, quasi-optimality is not guaranteed.

Theorem 4.5 (Approximation by \mathbf{u}_h). *Let $(\mathbf{u}, p) \in \mathbf{W}_0^{1, \varphi}(\Omega) \times L_0^{\varphi^*}(\Omega)$ be the solution of (2.15) and let $(\mathbf{u}_h, p_h) \in \mathbf{CR} \times \widehat{S}_0^0$ be computed by Method 4.1. Then*

$$\|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{D}_h \mathbf{u}_h)\|_2 \lesssim \inf_{\mathbf{v}_h \in \mathbf{Z}_h} \{ \|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{DE}\mathbf{v}_h)\|_2 + \|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{D}_h \mathbf{v}_h)\|_2 \}.$$

The hidden constant depends only on r^{\pm} , d and the shape constant of \mathcal{T} .

Proof. Let $\mathbf{v}_h \in \mathbf{Z}_h$ be arbitrary. Lemma 2.6, Corollary 2.14 and (2.26) reveal

$$\begin{aligned} \|\mathbf{F}(\mathbf{D}_h \mathbf{u}_h) - \mathbf{F}(\mathbf{D}_h \mathbf{v}_h)\|_2^2 &\approx \int_{\Omega} \varphi_{|\mathbf{D}_h \mathbf{u}_h|} (|\mathbf{D}_h \mathcal{I}^{\text{cr}} \mathbf{E}(\mathbf{u}_h - \mathbf{v}_h)|) \\ &\leq \sum_{K \in \mathcal{T}} \int_K \varphi_{|\mathbf{D}_h \mathbf{u}_h|} \left(\int_K |\mathbf{D}\mathbf{E}(\mathbf{u}_h - \mathbf{v}_h)| \right) \\ &\leq \int_{\Omega} \varphi_{|\mathbf{D}_h \mathbf{u}_h|} (|\mathbf{D}\mathbf{E}(\mathbf{u}_h - \mathbf{v}_h)|), \end{aligned}$$

with Jensen's inequality in the last step. Changing the shift to $|\mathbf{D}\mathbf{u}|$ with (2.9) for some $\delta > 0$ and applying the quasi-triangle inequality (2.3) yields

$$\begin{aligned} &\int_{\Omega} \varphi_{|\mathbf{D}_h \mathbf{u}_h|} (|\mathbf{D}\mathbf{E}(\mathbf{u}_h - \mathbf{v}_h)|) \\ &\leq C_{\delta} \int_{\Omega} \varphi_{|\mathbf{D}\mathbf{u}|} (|\mathbf{D}\mathbf{E}(\mathbf{u}_h - \mathbf{v}_h)|) + \delta \int_{\Omega} |\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{D}_h \mathbf{u}_h)|_2^2 \\ &\lesssim C_{\delta} \int_{\Omega} \varphi_{|\mathbf{D}\mathbf{u}|} (|\mathbf{D}\mathbf{u} - \mathbf{D}\mathbf{E}\mathbf{u}_h|) + C_{\delta} \int_{\Omega} \varphi_{|\mathbf{D}\mathbf{u}|} (|\mathbf{D}\mathbf{u} - \mathbf{D}\mathbf{E}\mathbf{v}_h|) \\ &\quad + \delta \|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{D}_h \mathbf{u}_h)\|_2^2 \\ &\approx C_{\delta} (\|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{D}\mathbf{E}\mathbf{u}_h)\|_2^2 + \|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{D}\mathbf{E}\mathbf{v}_h)\|_2^2) \\ &\quad + \delta \|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{D}_h \mathbf{u}_h)\|_2^2 \\ &\lesssim C_{\delta} \|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{D}\mathbf{E}\mathbf{v}_h)\|_2^2 + \delta \|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{D}_h \mathbf{u}_h)\|_2^2, \end{aligned}$$

where we have used Theorem 4.4 in the last step. With a quasi-triangle inequality, we thus obtain

$$\begin{aligned} \|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{D}_h \mathbf{u}_h)\|_2^2 &\lesssim \|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{D}_h \mathbf{v}_h)\|_2^2 + \|\mathbf{F}(\mathbf{D}_h \mathbf{u}_h) - \mathbf{F}(\mathbf{D}_h \mathbf{v}_h)\|_2^2 \\ &\lesssim \|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{D}_h \mathbf{v}_h)\|_2^2 + C_{\delta} \|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{D}\mathbf{E}\mathbf{v}_h)\|_2^2 \\ &\quad + \delta \|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{D}_h \mathbf{u}_h)\|_2^2. \end{aligned}$$

From this the claimed bound follows for $\delta > 0$ small enough. \square

In order to establish velocity-error estimates like (1.4) and (1.8), we must assess the approximation power of the space $\mathbf{E}(\mathbf{CR})$ in the natural distance.

Lemma 4.6 (Smoothing error). *For all $\mathbf{v}_h \in \mathbf{CR}$ we have*

$$\|\mathbf{F}(\mathbf{D}_h \mathbf{v}_h) - \mathbf{F}(\mathbf{D}\mathbf{E}\mathbf{v}_h)\|_2^2 \lesssim \sum_{K \in \mathcal{T}} \sum_{\substack{F \in \mathcal{F} \\ F \cap K \neq \emptyset}} \int_F h \varphi_{|\mathbf{D}_h \mathbf{v}_h|_K} (h^{-1} |[\mathbf{v}_h]_F|) =: \mathcal{J}(\mathbf{v}_h).$$

The hidden constant depends only on r^{\pm} , d and the shape constant of \mathcal{T} .

Proof. Property (2.30c) implies for all $\mathbf{v}_h \in \mathbf{CR}$ and $K \in \mathcal{T}$ that

$$(4.3) \quad \int_K |\mathbf{D}_h \mathbf{v}_h - \mathbf{D}\mathbf{E}\mathbf{v}_h| \lesssim \sum_{\substack{F \in \mathcal{F} \\ K \cap F \neq \emptyset}} \int_F h^{-1} |[\mathbf{v}_h]_F|$$

Using Lemmas 2.6 and 2.12 and (4.3), we estimate

$$\begin{aligned} \|\mathbf{F}(\mathbf{D}_h \mathbf{v}_h) - \mathbf{F}(\mathbf{D}\mathbf{E}\mathbf{v}_h)\|_{2;K}^2 &\approx \int_K \varphi_{|\mathbf{D}_h \mathbf{v}_h|} (|\mathbf{D}_h \mathbf{v}_h - \mathbf{D}\mathbf{E}\mathbf{v}_h|) \\ &\lesssim \int_K \varphi_{|\mathbf{D}_h \mathbf{v}_h|} \left(\int_K |\mathbf{D}_h \mathbf{v}_h - \mathbf{D}\mathbf{E}\mathbf{v}_h| \right) \\ &\lesssim \int_K \varphi_{|\mathbf{D}_h \mathbf{v}_h|} \left(\sum_{\substack{F \in \mathcal{F} \\ F \cap K \neq \emptyset}} \int_F h^{-1} |[\mathbf{v}_h]_F| \right). \end{aligned}$$

It follows from the quasi-triangle inequality (2.3) and Jensen's-inequality that

$$\begin{aligned} \|\mathbf{F}(\mathbf{D}_h \mathbf{v}_h) - \mathbf{F}(\mathbf{D}\mathbf{E}\mathbf{v}_h)\|_{2;K}^2 &\lesssim \int_K \sum_{\substack{F \in \mathcal{F} \\ F \cap K \neq \emptyset}} \varphi_{|\mathbf{D}_h \mathbf{v}_h|} \left(\int_F h^{-1} |[\mathbf{v}_h]_F| \right) \\ &\lesssim \int_K \sum_{\substack{F \in \mathcal{F} \\ F \cap K \neq \emptyset}} \int_F \varphi_{|\mathbf{D}_h \mathbf{v}_h|_K} (h^{-1} |[\mathbf{v}_h]_F|) \\ &\approx \sum_{\substack{F \in \mathcal{F} \\ F \cap K \neq \emptyset}} \int_F h \varphi_{|\mathbf{D}_h \mathbf{v}_h|_K} (h^{-1} |[\mathbf{v}_h]_F|). \end{aligned}$$

By summing over all $K \in \mathcal{T}$, we arrive at the desired estimate. \square

Motivated by the above result, we next investigate the approximation in **CR** in the natural distance augmented with jumps.

Lemma 4.7. *For $\mathbf{v} \in \mathbf{W}_0^{1,\varphi}(\Omega)$, we have*

$$\|\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\mathbf{D}_h \mathcal{I}^{cr} \mathbf{v})\|_2^2 + \mathcal{J}(\mathcal{I}^{cr} \mathbf{v}) \lesssim \sum_{K \in \mathcal{T}} \inf_{\mathbf{Q}_K \in \mathbb{R}^{d \times d}} \|\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\mathbf{Q}_K)\|_{2;\omega_K}^2.$$

The hidden constant depends only on r^\pm , d and the shape constant of \mathcal{T} .

Proof. Owing to Corollary 2.10, we have for all $K \in \mathcal{T}$ that

$$(4.4) \quad \|\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\mathbf{D}_h \mathcal{I}^{cr} \mathbf{v})\|_{2;K}^2 \lesssim \inf_{\mathbf{Q}_K \in \mathbb{R}^{d \times d}} \|\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\mathbf{Q}_K)\|_{2;K}^2.$$

In order to bound the term $\mathcal{J}(\mathcal{I}^{cr} \mathbf{v})$, we first employ (2.9) to change the shift to some arbitrary $\mathbf{Q} \in \mathbb{R}^{d \times d}$ and get

$$(4.5) \quad \begin{aligned} &\sum_{\substack{K \in \mathcal{T} \\ F \cap K \neq \emptyset}} \int_F h \varphi_{|\mathbf{D} \mathcal{I}^{cr} \mathbf{v}|_K} (h^{-1} |[\mathcal{I}^{cr} \mathbf{v}]_F|) \\ &\lesssim \sum_{\substack{K \in \mathcal{T} \\ F \cap K \neq \emptyset}} \int_F h \varphi_{|\mathbf{Q}|} (h^{-1} |[\mathcal{I}^{cr} \mathbf{v}]_F|) + \sum_{\substack{K \in \mathcal{T} \\ F \cap K \neq \emptyset}} \int_K |\mathbf{F}(\mathbf{D}_h \mathcal{I}^{cr} \mathbf{v}) - \mathbf{F}(\mathbf{Q})|^2, \end{aligned}$$

where we have used that $\mathbf{D}_h \mathcal{I}^{cr} \mathbf{v}$ is piecewise constant. Using again Corollary 2.10, The latter term can be bounded by

$$(4.6) \quad \sum_{\substack{K \in \mathcal{T} \\ F \cap K \neq \emptyset}} \|\mathbf{F}(\mathbf{D}_h \mathcal{I}^{cr} \mathbf{v}) - \mathbf{F}(\mathbf{Q})\|_{2;K}^2 \lesssim \sum_{\substack{K \in \mathcal{T} \\ F \cap K \neq \emptyset}} \|\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\mathbf{Q})\|_{2;K}^2.$$

For the first term at the right hand side of (4.5), we have

$$\int_F h \varphi_{|\mathbf{Q}|} (h^{-1} |[\mathcal{I}^{cr} \mathbf{v}]_F|) \lesssim |F| h_F \varphi_{|\mathbf{Q}|} \left(\int_F h^{-1} |[\mathcal{I}^{cr} \mathbf{v}]_F| \right),$$

where we used that $|\llbracket \mathcal{I}^{\text{cr}} \mathbf{v} \rrbracket_F| \lesssim f_F |\llbracket \mathcal{I}^{\text{cr}} \mathbf{v} \rrbracket_F|$ since $\llbracket \mathcal{I}^{\text{cr}} \mathbf{v} \rrbracket_F$ is linear on F . Note that thanks to $\mathbf{v} \in \mathbf{W}_0^{1,\varphi}(\Omega)$, we have $\llbracket \mathbf{v} \rrbracket_F = 0$ in $F \in \mathcal{F}$. Consequently, $\llbracket \mathcal{I}^{\text{cr}} \mathbf{v} \rrbracket_F = \llbracket \mathcal{I}^{\text{cr}} \mathbf{v} - \mathbf{v} \rrbracket_F$ in $F \in \mathcal{F}$ and it follows from Lemma 2.9 that

$$\int_F h^{-1} |\llbracket \mathcal{I}^{\text{cr}} \mathbf{v} \rrbracket_F| = \int_F h^{-1} |\llbracket \mathcal{I}^{\text{cr}} \mathbf{v} - \mathbf{v} \rrbracket_F| \lesssim \sum_{\substack{K \in \mathcal{T} \\ F \subset K}} \int_K |\nabla(\mathcal{I}^{\text{cr}} \mathbf{v} - \mathbf{v})|.$$

This and Jensen's inequality imply

$$\begin{aligned} \int_F h \varphi_{|\mathbf{Q}|} (h^{-1} |\llbracket \mathcal{I}^{\text{cr}} \mathbf{v} \rrbracket_F|) &\lesssim \int_F h \varphi_{|\mathbf{Q}|} \left(\sum_{\substack{K \in \mathcal{T} \\ F \subset K}} \int_K |\nabla(\mathcal{I}^{\text{cr}} \mathbf{v} - \mathbf{v})| \right) \\ &\lesssim \sum_{\substack{K \in \mathcal{T} \\ F \subset K}} \int_K \varphi_{|\mathbf{Q}|} (|\nabla(\mathcal{I}^{\text{cr}} \mathbf{v} - \mathbf{v})|). \end{aligned}$$

Thanks to (2.25) we can apply the Korn inequality (Proposition 2.1) and obtain

$$\int_F h \varphi_{|\mathbf{Q}|} (h^{-1} |\llbracket \mathcal{I}^{\text{cr}} \mathbf{v} \rrbracket_F|) \lesssim \sum_{\substack{K \in \mathcal{T} \\ F \subset K}} \int_K \varphi_{|\mathbf{Q}|} (|\mathbf{D}(\mathcal{I}^{\text{cr}} \mathbf{v} - \mathbf{v})|).$$

At this point, we again invoke (2.9) to change the shift to $\mathbf{D}\mathbf{v}$ and obtain with Lemma 2.6

$$\begin{aligned} \int_F h \varphi_{|\mathbf{Q}|} (h^{-1} |\llbracket \mathcal{I}^{\text{cr}} \mathbf{v} \rrbracket_F|) &\lesssim \sum_{\substack{K \in \mathcal{T} \\ F \subset K}} \left(\int_K \varphi_{|\mathbf{D}\mathbf{v}|} (|\mathbf{D}(\mathcal{I}^{\text{cr}} \mathbf{v} - \mathbf{v})|) + \int_K |\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\mathbf{Q})|^2 \right) \\ &\lesssim \sum_{\substack{K \in \mathcal{T} \\ F \subset K}} (\|\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\mathbf{D}\mathcal{I}^{\text{cr}} \mathbf{v})\|_{2;K}^2 + \|\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\mathbf{Q})\|_{2;K}^2) \\ &\lesssim \sum_{\substack{K \in \mathcal{T} \\ F \subset K}} \|\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\mathbf{Q})\|_{2;K}^2, \end{aligned}$$

where we used (4.4) in the last step. Combining this with (4.5) and (4.6) and recalling that $\mathbf{Q} \in \mathbb{R}^{d \times d}$ was arbitrary, we have proved for $F \in \mathcal{F}$ that

$$\sum_{\substack{K \in \mathcal{T} \\ F \cap K \neq \emptyset}} \int_F h \varphi_{|\mathbf{D}\mathcal{I}^{\text{cr}} \mathbf{v}|_K} (h^{-1} |\llbracket \mathcal{I}^{\text{cr}} \mathbf{v} \rrbracket_F|) \lesssim \sum_{\substack{K \in \mathcal{T} \\ F \cap K \neq \emptyset}} \|\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\mathbf{Q})\|_{2;K}^2.$$

Note that $\mathbf{F} : \mathbb{R}^{d \times d} \rightarrow \mathbb{R}^{d \times d}$ is surjective and since $\mathbf{Q} \in \mathbb{R}^{d \times d}$ is arbitrary, we have

$$\sum_{\substack{K \in \mathcal{T} \\ F \cap K \neq \emptyset}} \int_F h \varphi_{|\mathbf{D}\mathcal{I}^{\text{cr}} \mathbf{v}|_K} (h^{-1} |\llbracket \mathcal{I}^{\text{cr}} \mathbf{v} \rrbracket_F|) \lesssim \inf_{\mathbf{Q} \in \mathbb{R}^{d \times d}} \sum_{\substack{K \in \mathcal{T} \\ F \cap K \neq \emptyset}} \|\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{Q}\|_{2;K}^2.$$

Together with (4.4), summing over all $F \in \mathcal{F}$ finally proves the assertion. \square

With this preparation, we can prove the following main result.

Theorem 4.8 (Velocity-error). *Let $(\mathbf{u}, p) \in \mathbf{W}_0^{1,\varphi}(\Omega) \times L_0^{\varphi^*}(\Omega)$ be the solution to (2.15) and let $(\mathbf{u}_h, p_h) \in \mathbf{CR} \times \widehat{S}_0^0$ be computed by Method 4.1. Then we have*

$$\|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{D}_h \mathbf{u}_h)\|_2^2 + \|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{D}\mathbf{E}\mathbf{u}_h)\|_2^2 \lesssim \sum_{K \in \mathcal{T}} \inf_{\mathbf{Q}_K \in \mathbb{R}^{d \times d}} \|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{Q}_K\|_{2;\omega_K}^2.$$

In particular, we have for $\mathbf{F}(\mathbf{D}\mathbf{u}) \in \mathbf{W}^{1,2}(\Omega)$ the estimate

$$\|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{D}_h\mathbf{u}_h)\|_2 + \|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{D}\mathbf{E}\mathbf{u}_h)\|_2 \lesssim \|h\nabla\mathbf{F}(\mathbf{D}\mathbf{u})\|_2.$$

The hidden constant depends only on r^\pm , d and the shape constant of \mathcal{T} .

Proof. We conclude from Theorems 4.4 and 4.5, and Lemma 4.6 that

$$\begin{aligned} \|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{D}_h\mathbf{u}_h)\|_2^2 + \|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{D}\mathbf{E}\mathbf{u}_h)\|_2^2 \\ \lesssim \inf_{\mathbf{v}_h \in \mathbf{Z}_h} \{ \|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{D}_h\mathbf{v}_h)\|_2^2 + \mathcal{J}(\mathbf{v}_h) \}. \end{aligned}$$

Choosing $\mathbf{v}_h = \mathcal{I}^{\mathbf{c}}\mathbf{u} \in \mathbf{Z}_h$, the assertion follows from Lemma 4.7. \square

Remark 4.9 (Stabilization for the symmetric gradient problem). *In order to handle the non-definiteness of the symmetric gradient, one can also consider a jump penalization of our method as it is common for discontinuous Galerkin (DG) methods; compare with [12]. This has been done by Kaltenbach and Růžička in [29, 30, 31] for a DG approximation of (r, ε) -structure systems. However, their approach suffers from non-monotonicity of the stabilized nonlinear operator so that uniqueness of the discrete solution cannot be guaranteed. In contrast, our Method 4.1 features monotonicity and hence unique solvability.*

5. ESTIMATES OF THE PRESSURE-ERROR

Although the above velocity-error estimates for Methods 3.1 and 4.1 are pressure robust, i.e. independent of the pressure error, the pressure-error itself depends on the velocity-error. Consequently, improving the velocity approximation can improve also the pressure approximation in some cases. To make this more precise, we combine ideas from [9, Theorem 4.7] with the results in the previous sections. For the ease of presentation, we restrict ourselves to problems with (r, ε) -structure.

Theorem 5.1 (Pressure-error). *Let \mathbf{A} have (r, ε) -structure (2.10) for some $1 < r < \infty$ and $\varepsilon \geq 0$. Let $(\mathbf{u}, p) \in \mathbf{W}_0^{1,r}(\Omega) \times L_0^{r'}(\Omega)$ be the unique solution to (2.15) and let $(\mathbf{u}_h, p_h) \in \mathbf{CR} \times \widehat{S}_0^0$ be obtained with either Method 3.1 or 4.1. Then we have*

$$\|p - p_h\|_{r'} \lesssim \left(\sum_{K \in \mathcal{T}} \inf_{\mathbf{Q}_K \in \mathbf{R}^{d \times d}} \|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{Q}_K\|_{2;\omega_K}^2 \right)^{\frac{1}{2} \min\{1, \frac{2}{r'}\}} + \inf_{q_h \in \widehat{S}_0^0} \|p - q_h\|_{r'}.$$

In particular, when $\mathbf{F}(\nabla\mathbf{u}) \in \mathbf{W}^{1,2}(\Omega)$ and $p \in W^{1,r'}(\Omega)$, we have

$$(5.1) \quad \|p - p_h\|_{r'} \lesssim \|h\nabla\mathbf{F}(\mathbf{D}\mathbf{u})\|_2^{\min\{1, \frac{2}{r'}\}} + \|h\nabla p\|_{r'}.$$

Here $\mathbf{D} = \nabla$ or $\mathbf{D} = \mathbf{D}$ depending on the context and for $r > 2$, the hidden constant depends also on \mathbf{f} but is robust for $\varepsilon \rightarrow 0$.

Proof. We focus on the case of Method 3.1. For Method 4.1 the assertion follows with minor changes. Let $\pi p \in \widehat{S}_0^0$ be the L^2 -orthogonal projection of p and apply the triangle inequality

$$\|p - p_h\|_{r'} \leq \|p - \pi p\|_{r'} + \|\pi p - p_h\|_{r'}.$$

Owing to Jensen's inequality, the projection π is bounded, cf. (2.26) where the same argument is used. Therefore, the first term on the right-hand side above is bounded by

$$\|p - \pi p\|_{r'} \leq 2 \inf_{q_h \in \widehat{S}_0^0} \|p - q_h\|_{r'}.$$

For the second term we invoke the inf-sup stability (2.28), $\operatorname{div}_h \mathbf{CR} = \widehat{S}_0^0$ and (2.29c), to obtain

$$\beta \|\pi p - p_h\|_{r'} \leq \sup_{\mathbf{v}_h \in \mathbf{CR}} \frac{\int_{\Omega} (\pi p - p_h) \operatorname{div}_h \mathbf{v}_h}{\|\nabla_h \mathbf{v}_h\|_r} = \sup_{\mathbf{v}_h \in \mathbf{CR}} \frac{\int_{\Omega} (p - p_h) \operatorname{div} \mathbf{E} \mathbf{v}_h}{\|\nabla_h \mathbf{v}_h\|_r}.$$

For the enumerator, the continuous equations (2.15) and Method 3.1 imply

$$\begin{aligned} \int_{\Omega} (p - p_h) \operatorname{div} \mathbf{E} \mathbf{v}_h &= \int_{\Omega} (\mathbf{A}(\nabla \mathbf{u}) - \mathbf{A}(\nabla_h \mathbf{u}_h)) : \nabla \mathbf{E} \mathbf{v}_h \\ &\leq \|\mathbf{A}(\nabla \mathbf{u}) - \mathbf{A}(\nabla_h \mathbf{u}_h)\|_{r'} \|\nabla \mathbf{E} \mathbf{v}_h\|_r. \end{aligned}$$

Combining the above estimates with (2.33) and [9, Lemma 4.6] or [26, Lemma 2.4] thus yields

$$\begin{aligned} \|\pi p - p_h\|_{r'} &\lesssim \|\mathbf{A}(\nabla \mathbf{u}) - \mathbf{A}(\nabla_h \mathbf{u}_h)\|_{r'} \\ &\lesssim \begin{cases} \|\mathbf{F}(\nabla \mathbf{u}) - \mathbf{F}(\nabla_h \mathbf{u}_h)\|_2^{\frac{2}{r'}} & \text{if } r \in (1, 2] \\ \|\varepsilon + |\nabla \mathbf{u}| + |\nabla_h \mathbf{u}_h|\|_r^{\frac{r-2}{2}} \|\mathbf{F}(\nabla \mathbf{u}) - \mathbf{F}(\nabla_h \mathbf{u}_h)\|_2 & \text{if } r \in (2, \infty). \end{cases} \end{aligned}$$

The first assertion then follows from Theorem 3.3, and, in the case $r > 2$, from the boundedness of Galerkin solutions $\|\nabla \mathbf{u}\|_r + \|\nabla_h \mathbf{u}_h\|_r \lesssim \|\mathbf{f}\|_{-1, r'}^{\frac{1}{r-1}}$, with a hidden constant depending only on r ; compare with [27, (2.17)]. The claimed a priori rate follows again from Theorem 3.3 and a Poincaré inequality. \square

Remark 5.2 (Rate of convergence). *Theorem 5.1 predicts the same rate of convergence of the pressure-error as in [31], under less restrictive regularity assumptions and including $\varepsilon = 0$; compare also with Remark 3.5. The same rate is observed also in the numerical experiments of [9] where, however, the theoretical expectation is suboptimal for $r > 2$, cf. (1.7).*

6. NUMERICAL EXPERIMENTS

In this section, we illustrate the performance and the pressure robustness of the Methods 3.1 and 4.1 for the r -Stokes equations, i.e. for

$$(6.1) \quad \mathbf{A}(\mathbf{Q}) = |\mathbf{Q}|^{r-2} \mathbf{Q}$$

in Assumption 2.3. Note that the results from [30, 31] do not apply to this case, as they require the (r, ε) -structure and are not robust in the limit $\varepsilon \rightarrow 0$.

To highlight pressure robustness, we compare also with the following counterpart of Method 3.1 without smoothing operator.

Method 6.1. For $\mathbf{f} \in \mathbf{W}^{-1, r'}(\Omega)$ compute $\mathbf{u}_h \in \mathbf{CR}$ and $p_h \in \widehat{S}_0^0$ such that

$$\begin{aligned} \forall \mathbf{v}_h \in \mathbf{CR} \quad & \int_{\Omega} \mathbf{A}(\nabla_h \mathbf{u}_h) : \nabla_h \mathbf{v}_h - \int_{\Omega} p_h \operatorname{div}_h \mathbf{v}_h = \langle \mathbf{f}, \mathbf{v}_h \rangle, \\ \forall q_h \in \widehat{S}_0^0 \quad & \int_{\Omega} q_h \operatorname{div}_h \mathbf{u}_h = 0. \end{aligned}$$

Recall that we could not apply this method with the gradient replaced by the symmetric gradient, cf. [1].

Our implementation is realized with the help of the C library ALBERTA 3.1 [24, 41]. We solve the nonlinear system of equations stemming from each method via a relaxed Kačanov iteration, based on [13, 2]. For the solution of the linearized system, we apply SymmLQ with block diagonal preconditioning.

We consider two test cases with prescribed manufactured solution. In both cases, the domain is a square and the initial mesh \mathcal{T}_0 is obtained by drawing the two main diagonals. We obtain subsequent meshes $\mathcal{T}_1, \mathcal{T}_2, \dots$ by performing each time two uniform refinements with newest vertex bisection. We estimate the decay rate of the

velocity- and of the pressure-error with respect to the number of mesh refinements via the experimental order of convergence,

$$\text{EOC}_k := \log(e_{k-1}/e_k) / \log(2)$$

where e_k is the error of interest on \mathcal{T}_k for $k \geq 1$. For Method 4.1, we consider e.g. the velocity error with smoother $\|\mathbf{F}(\mathbf{D}\mathbf{u}) - \mathbf{F}(\mathbf{D}\mathbf{E}\mathbf{u}_h)\|_2$.

Test case 1: power functions. Following [9, section 7], we compute the data so that the solution of (2.15) with (6.1) is given by

$$(6.2) \quad \mathbf{u}(\mathbf{x}) = |\mathbf{x}|^\alpha (x_2, -x_1) \quad \text{and} \quad p(\mathbf{x}) = \eta |\mathbf{x}|^\gamma$$

for $\mathbf{x} = (x_1, x_2) \in \Omega = (-1, 1)^2$. Different from (2.15), we have nonzero Dirichlet datum on $\partial\Omega$, but a similar (undisplayed) test case on the unit circle with zero Dirichlet datum confirmed the results discussed below. In analogy with [9], we set

$$\alpha = 0.01, \quad \gamma = \frac{2}{r} - 1 + 0.01, \quad \eta = \begin{cases} 0.01 & \text{if } r \in (1, 2), \\ 1 & \text{if } r \in [2, \infty). \end{cases}$$

The factor η scales the pressure best-error relative to the velocity-error on the right hand side of (5.1) for $r \in (1, 2)$ such that the suboptimal convergence of the pressure-error can be better observed in Figure 2.

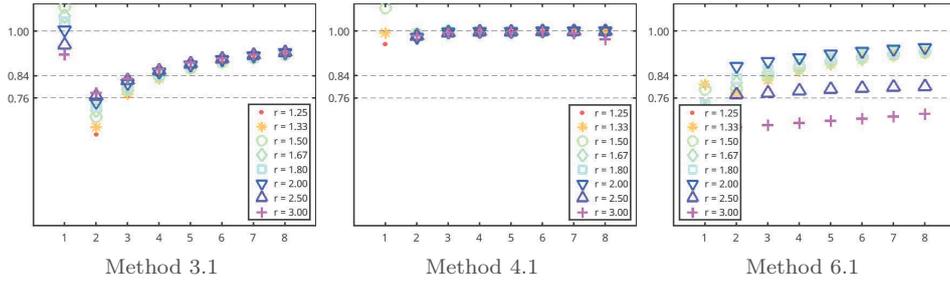


FIGURE 1. Test case 1. Experimental order of convergence EOC_k of the velocity-error versus the number k of mesh refinements.

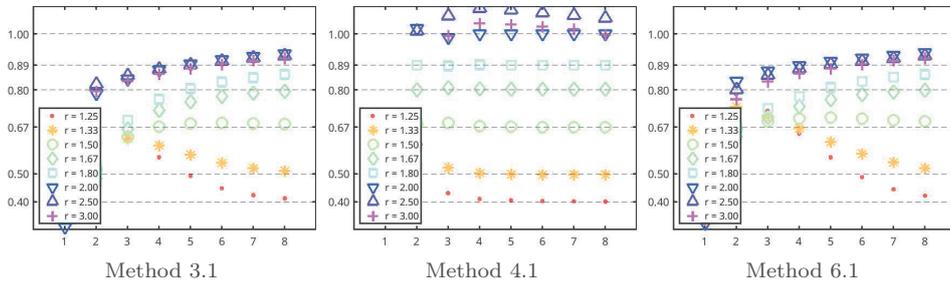


FIGURE 2. Test case 1. Experimental order of convergence EOC_k of the pressure-error versus the number k of mesh refinements.

Since $\mathbf{F}(\mathbf{D}\mathbf{u}) \in W^{1,2}(\Omega)^{2 \times 2}$, we expect that EOC_k for the velocity-error of both Method 3.1 and 4.1 converges to 1 as k increases, owing to Theorems 3.3 and 4.8, respectively. Similarly, since $p \in W^{1,r'}(\Omega)$ for all $r > 1$, we expect that EOC_k for the pressure-error of both methods converges to $\min\{1, \frac{2}{r'}\}$, as a consequence

of Theorem 5.1. The data in Figure 1 and 2 (left and center panel) confirm our expectation. The convergence to the expected value is faster for Method 4.1 than for Method 3.1. We verified that the gap is somehow related to this specific example and it is less evident if, e.g., we increase the exponent α in (6.2), see also the next test case.

Regarding Method 6.1, we observe that EOC_k for the velocity-error converges to the suboptimal value $\min\{1, \frac{r'}{2}\}$ (Figure 1, right panel), which is in line with the estimate (1.6) for another method lacking pressure robustness. For the pressure-error (Figure 2, right panel), the results suggest convergence of EOC_k to $\min\{1, \frac{2}{r'}\}$, as for the other methods. Thus, for $r > 2$, the convergence is faster than expected from (1.7), as reported also in [9].

Test case 2: jumping pressure. In analogy with [47, section 6.3] and [34, section 4.2], we consider a test case with smooth velocity and discontinuous pressure

$$\mathbf{u}(\mathbf{x}) = |\mathbf{x}|^{0.5}(x_2, -x_1) \quad \text{and} \quad p(\mathbf{x}) = \begin{cases} -\frac{3}{2}, & x_1 < \frac{2}{3} \\ 3, & x_1 > \frac{2}{3} \end{cases}$$

for $\mathbf{x} = (x_1, x_2) \in \Omega = (0, 1)^2$. As before, we expect that EOC_k for the velocity-error of both Method 3.1 and 4.1 converges to 1 as k increases. In contrast, since $p \in W^{s, r'}(\Omega)$ only for $s < \frac{1}{r'}$, we expect that EOC_k for the pressure-error of both methods converges to $\frac{1}{r'}$ for all $r > 1$. Our expectation is confirmed by the data in Figure 3 and 4 (left and center panel).

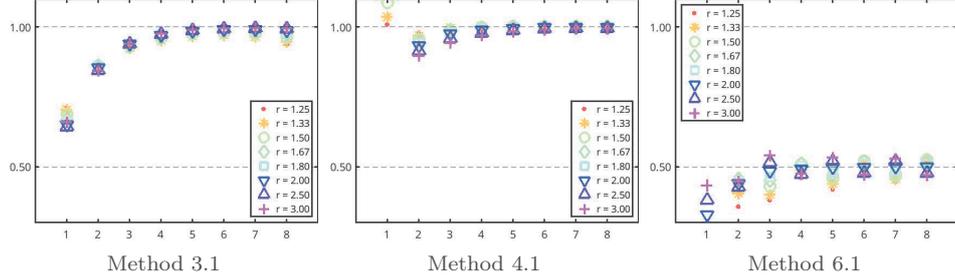


FIGURE 3. Test case 2. Experimental order of convergence EOC_k of the velocity-error versus the number k of mesh refinements.

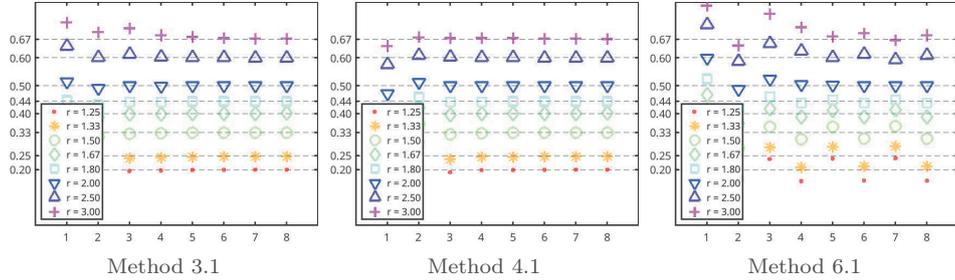


FIGURE 4. Test case 2. Experimental order of convergence EOC_k of the pressure-error versus the number k of mesh refinements.

Compared with the previous test case, the pressure regularity is much lower and the importance of pressure robustness is even more apparent. Indeed, for the

Method 6.1 (Figure 3, right panel), we observe that EOC_k for the velocity-error converges to the sub-optimal rate 0.5, irrespective of r , whereas for the pressure-error we have similar results as for the other methods (Figure 4, right panel).

7. CONCLUSIONS

We have proposed two numerical methods for the nonlinear Stokes equations (1.1) using nonconforming Crouzeix-Raviart velocity elements and piecewise constant pressure elements (Method 3.1 and 4.1) with $\mathcal{D} = \nabla$ and $\mathcal{D} = \mathbf{D}$ respectively. Both schemes feature a monotone nonlinear operator acting on the velocity and thus the discrete problems are uniquely solvable; see Propositions 3.2 and 4.3. Both methods are pressure robust. The error estimates in Theorems 3.3, 4.8 and 5.1 improve upon existing results for similar methods and are in line with the best known results for conforming and divergence-free methods.

Funding. Christian Kreuzer acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 321270008. Lars Diening research is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - SFB 1283/2 2021. Pietro Zanotti was supported by the GNCS-INdAM project CUP E53C23001670001.

REFERENCES

- [1] D. N. Arnold. “On nonconforming linear-constant elements for some variants of the Stokes equations”. In: *Istit. Lombardo Accad. Sci. Lett. Rend. A* 127.1 (1993), pp. 83–93.
- [2] A. K. Balcı, L. Diening, and J. Storn. “Relaxed Kačanov scheme for the p -Laplacian with large exponent”. In: *SIAM J. Numer. Anal.* 61.6 (2023), pp. 2775–2794. DOI: 10.1137/22M1528550.
- [3] J. Barrett and W. Liu. “Finite element approximation of the p -Laplacian”. In: *Math. Comp.* 61.204 (1993), pp. 523–537. DOI: 10.2307/2153239.
- [4] J. Barrett and W. Liu. “Finite element error analysis of a quasi-Newtonian flow obeying the Carreau or power law”. In: *Numer. Math.* 64.4 (1993), pp. 433–453. DOI: 10.1007/BF01388698.
- [5] J. Barrett and W. Liu. “Quasi-norm error bounds for the finite element approximation of a non-Newtonian flow”. In: *Numer. Math.* 68.4 (1994), pp. 437–456. DOI: 10.1007/s002110050071.
- [6] L. Behn and L. Diening. “Global regularity for nonlinear systems with symmetric gradients”. In: *Calc. Var. Partial Differential Equations* 63.3 (2024), Paper No. 60, 25. DOI: 10.1007/s00526-024-02666-z.
- [7] H. Beirão da Veiga. “Navier-Stokes equations with shear thinning viscosity. Regularity up to the boundary”. In: *J. Math. Fluid Mech.* 11.2 (2009), pp. 258–273. DOI: 10.1007/s00021-008-0258-1.
- [8] H. Beirão da Veiga, P. Kaplický, and M. Růžička. “Boundary regularity of shear thickening flows”. In: *J. Math. Fluid Mech.* 13.3 (2011), pp. 387–404. DOI: 10.1007/s00021-010-0025-y.
- [9] L. Belenki, L. C. Berselli, L. Diening, and M. Růžička. “On the finite element approximation of p -Stokes systems”. In: *SIAM J. Numer. Anal.* 50.2 (2012), pp. 373–397. DOI: 10.1137/10080436X.
- [10] L. Belenki, L. Diening, and C. Kreuzer. “Optimality of an adaptive finite element method for the p -Laplacian equation”. In: *IMA J. Numer. Anal.* 32.2 (2012), pp. 484–510. DOI: 10.1093/imanum/drr016.
- [11] M. Crouzeix and P.-A. Raviart. “Conforming and nonconforming finite element methods for solving the stationary Stokes equations. I”. In: *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge* 7 (1973), pp. 33–75.

- [12] D. A. Di Pietro and A. Ern. *Mathematical aspects of discontinuous Galerkin methods*. Vol. 69. Mathématiques & Applications (Berlin) [Mathematics & Applications]. Springer, Heidelberg, 2012, pp. xviii+384.
- [13] L. Diening, M. Fornasier, R. Tomasi, and M. Wank. “A relaxed Kačanov iteration for the p -Poisson problem”. In: *Numer. Math.* 145.1 (2020), pp. 1–34. DOI: 10.1007/s00211-020-01107-1.
- [14] L. Diening, P. Kaplický, and S. Schwarzacher. “BMO estimates for the p -Laplacian”. In: *Nonlinear Anal.* 75.2 (2012), pp. 637–650. DOI: 10.1016/j.na.2011.08.065.
- [15] L. Diening and M. Růžička. “Interpolation operators in Orlicz-Sobolev spaces”. In: *Numer. Math.* 107.1 (2007), pp. 107–129. DOI: 10.1007/s00211-007-0079-9.
- [16] L. Diening, M. Růžička, and K. Schumacher. “A decomposition technique for John domains”. In: *Ann. Acad. Sci. Fenn. Math.* 35.1 (2010), pp. 87–114. DOI: 10.5186/aasfm.2010.3506.
- [17] L. Diening, C. Ebmeyer, and M. Růžička. “Optimal convergence for the implicit space-time discretization of parabolic systems with p -structure”. In: *SIAM J. Numer. Anal.* 45.2 (2007), pp. 457–472. DOI: 10.1137/05064120X.
- [18] L. Diening and F. Ettwein. “Fractional estimates for non-differentiable elliptic systems with general growth”. In: *Forum Math.* 20.3 (2008), pp. 523–556. DOI: 10.1515/FORUM.2008.027.
- [19] C. Ebmeyer and W. Liu. “Quasi-norm interpolation error estimates for piecewise linear finite element approximation of p -Laplacian problems”. In: *Numer. Math.* 100.204 (2005), pp. 233–258.
- [20] A. Ern and J.-L. Guermond. *Finite elements II—Galerkin approximation, elliptic and mixed PDEs*. Vol. 73. Texts in Applied Mathematics. Springer, Cham, 2021, pp. ix+492. DOI: 10.1007/978-3-030-56923-5.
- [21] E. H. Georgoulis and T. Pryer. “Recovered finite element methods”. In: *Comput. Methods Appl. Mech. Engrg.* 332 (2018), pp. 303–324. DOI: 10.1016/j.cma.2017.12.026.
- [22] J. Guzmán and M. Neilan. “Conforming and divergence-free Stokes elements on general triangular meshes”. In: *Math. Comp.* 83.285 (2014), pp. 15–36. DOI: 10.1090/S0025-5718-2013-02753-6.
- [23] J. Guzmán and M. Neilan. “inf-sup stable finite elements on barycentric refinements producing divergence-free approximations in arbitrary dimensions”. In: *SIAM J. Numer. Anal.* 56.5 (2018), pp. 2826–2844. DOI: 10.1137/17M1153467.
- [24] K.-J. Heine, D. Köster, O. Kriessl, A. Schmidt, and K. Siebert. *ALBERTA - an adaptive hierarchical finite element toolbox*. <http://www.alberta-fem.de>.
- [25] C. Helanow and J. Ahlkrone. “Stabilized equal low-order finite elements in ice sheet modeling—accuracy and robustness”. In: *Comput. Geosci.* 22.4 (2018), pp. 951–974. DOI: 10.1007/s10596-017-9713-5.
- [26] A. Hirn. “Approximation of the p -Stokes equations with equal-order finite elements”. In: *J. Math. Fluid Mech.* 15.1 (2013), pp. 65–88. DOI: 10.1007/s00021-012-0095-0.
- [27] A. Hirn. “Finite element approximation of singular power-law systems”. In: *Math. Comp.* 82.283 (2013), pp. 1247–1268. DOI: 10.1090/S0025-5718-2013-02668-3.
- [28] V. John, A. Linke, C. Merdon, M. Neilan, and L. G. Rebholz. “On the divergence constraint in mixed finite element methods for incompressible flows”. In: *SIAM Rev.* 59.3 (2017), pp. 492–544. DOI: 10.1137/15M1047696.
- [29] A. Kaltenbach and M. Růžička. “A local discontinuous Galerkin approximation for the p -Navier-Stokes system, Part I: Convergence analysis”. In: *SIAM J. Numer. Anal.* 61.4 (2023), pp. 1613–1640. DOI: 10.1137/22M151474X.
- [30] A. Kaltenbach and M. Růžička. “A local discontinuous Galerkin approximation for the p -Navier-Stokes system, Part II: Convergence rates for the velocity”. In: *SIAM J. Numer. Anal.* 61.4 (2023), pp. 1641–1663. DOI: 10.1137/22M1514751.
- [31] A. Kaltenbach and M. Růžička. “A local discontinuous Galerkin approximation for the p -Navier-stokes system, Part III: Convergence rates for the pressure”. In: *SIAM J. Numer. Anal.* 61.4 (2023), pp. 1763–1782. DOI: 10.1137/22M1541472.

- [32] V. Kokilashvili and M. Krbeč. *Weighted inequalities in Lorentz and Orlicz spaces*. World Scientific Publishing Co., Inc., River Edge, NJ, 1991, pp. xii+233. DOI: 10.1142/9789814360302.
- [33] C. Kreuzer. “Reliable and efficient a posteriori error estimates for finite element approximations of the parabolic p -Laplacian”. In: *Calcolo* 50.2 (2013), pp. 79–110. DOI: 10.1007/s10092-012-0059-z.
- [34] C. Kreuzer, R. Verfürth, and P. Zanotti. “Quasi-optimal and pressure robust discretizations of the Stokes equations by moment- and divergence-preserving operators”. In: *Comput. Methods Appl. Math.* 21.2 (2021), pp. 423–443. DOI: 10.1515/cmam-2020-0023.
- [35] C. Kreuzer and P. Zanotti. “Quasi-optimal and pressure-robust discretizations of the Stokes equations by new augmented Lagrangian formulations”. In: *IMA J. Numer. Anal.* 40.4 (2020), pp. 2553–2583. DOI: 10.1093/imanum/drz044.
- [36] A. Linke, C. Merdon, M. Neilan, and F. Neumann. “Quasi-optimality of a pressure-robust nonconforming finite element method for the Stokes-problem”. In: *Math. Comp.* 87.312 (2018), pp. 1543–1566. DOI: 10.1090/mcom/3344.
- [37] A. Linke. “On the role of the Helmholtz decomposition in mixed methods for incompressible flows and a new variational crime”. In: *Comput. Methods Appl. Mech. Engrg.* 268 (2014), pp. 782–800. DOI: 10.1016/j.cma.2013.10.011.
- [38] W. Liu. “Degenerate quasilinear elliptic equations arising from bimaterial problems in elastic-plastic mechanics”. In: *Nonlinear Anal.* 35.4 (1999), pp. 517–529. DOI: 10.1016/S0362-546X(98)00014-5.
- [39] J. Málek, J. Nečas, M. Rokyta, and M. Růžička. *Weak and measure-valued solutions to evolutionary PDEs*. Vol. 13. Applied Mathematics and Mathematical Computation. Chapman & Hall, London, 1996, pp. xii+317. DOI: 10.1007/978-1-4899-6824-1.
- [40] M. M. Rao and Z. D. Ren. *Theory of Orlicz spaces*. Vol. 146. Monographs and Textbooks in Pure and Applied Mathematics. Marcel Dekker, Inc., New York, 1991, pp. xii+449.
- [41] A. Schmidt and K. G. Siebert. *Design of Adaptive Finite Element Software*. Vol. 42. Lecture Notes in Computational Science and Engineering. The Finite Element Toolbox ALBERTA. Springer-Verlag, Berlin, 2005, pp. xii+315.
- [42] L. R. Scott and M. Vogelius. “Norm estimates for a maximal right inverse of the divergence operator in spaces of piecewise polynomials”. In: *RAIRO Modél. Math. Anal. Numér.* 19.1 (1985), pp. 111–143. DOI: 10.1051/m2an/1985190101111.
- [43] A. Veese. “Approximating gradients with continuous piecewise polynomial functions”. In: *Found. Comput. Math.* 16.3 (2016), pp. 723–750. DOI: 10.1007/s10208-015-9262-z.
- [44] A. Veese and P. Zanotti. “Quasi-optimal nonconforming methods for symmetric elliptic problems. I—Abstract theory”. In: *SIAM J. Numer. Anal.* 56.3 (2018), pp. 1621–1642. DOI: 10.1137/17M1116362.
- [45] A. Veese and P. Zanotti. “Quasi-optimal nonconforming methods for symmetric elliptic problems. II—Overconsistency and classical nonconforming elements”. In: *SIAM J. Numer. Anal.* 57.1 (2019), pp. 266–292. DOI: 10.1137/17M1151651.
- [46] H. B. de Veiga. “On the global regularity for singular p -systems under non-homogeneous Dirichlet boundary conditions”. In: *J. Math. Anal. Appl.* 398.2 (2013), pp. 527–533. DOI: 10.1016/j.jmaa.2012.08.058.
- [47] R. Verfürth and P. Zanotti. “A quasi-optimal Crouzeix-Raviart discretization of the Stokes equations”. In: *SIAM J. Numer. Anal.* 57.3 (2019), pp. 1082–1099. DOI: 10.1137/18M1177688.
- [48] S. Zhang. “On the family of divergence-free finite elements on tetrahedral grids for the Stokes equations”. In: *Numerical Methods for Partial Differential Equations* 40 (2023). DOI: 10.1002/num.23055.

(Lars Diening) UNIVERSITÄT BIELEFELD, FAKULTÄT FÜR MATHEMATIK, D-33615 BIELEFELD,
GERMANY

Email address: `lars.diening@uni-bielefeld.de`

(Adrian Hirn) HOCHSCHULE ESSLINGEN, ROBERT-BOSCH-STRASSE 1, 73037 GÖPPINGEN,
GERMANY

Email address: `adrian.hirn@hs-esslingen.de`

(Christian Kreuzer) TU DORTMUND, FAKULTÄT FÜR MATHEMATIK, D-44221 DORTMUND,
GERMANY

Email address: `christian.kreuzer@tu-dortmund.de`

(Pietro Zanotti) UNIVERSITÀ DEGLI STUDI DI MILANO, DIPARTIMENTO DI MATEMATICA, 20133
MILANO, ITALY

Email address: `pietro.zanotti@unimi.it`