ON THE FINITE ELEMENT APPROXIMATION OF p-STOKES SYSTEMS

L. BELENKI[†], L. C. BERSELLI[*], L. DIENING[†], AND M. RŮŽIČKA[†]

Abstract. In this paper we study the finite element approximation of systems of p-Stokes type for $p \in (1, \infty)$. We derive (in some cases optimal) error estimates for finite element approximation of the velocity and for the pressure in a suitable functional setting. The results are supported by numerical experiments.

Keywords. Error analysis, inf-sup condition, velocity, pressure, conforming elements.

1. Introduction. We study the numerical approximation of steady systems of p-Stokes type

$$\begin{aligned} -\operatorname{div} \mathbf{S}(\mathbf{D}\mathbf{v}) + \nabla q &= \mathbf{f} && \text{in } \Omega, \\ -\operatorname{div} \mathbf{v} &= g && \text{in } \Omega, \\ \mathbf{v} &= \mathbf{0} && \text{on } \partial\Omega, \end{aligned} \tag{1.1}$$

by means of conforming finite element spaces satisfying the classical discrete inf-sup condition. The physical problem which motivates this study is the steady motion of a homogeneous, incompressible fluid with shear-dependent viscosity, in the Stokes approximation of small velocities. Here $\Omega \subset \mathbb{R}^n$, $n \geq 2$, is a polyhedral, bounded domain. The unknowns are the velocity vector field $\mathbf{v} = (v_1, \ldots, v_n)$ and the scalar kinematic pressure q. The extra stress tensor $\mathbf{S}(\mathbf{D}\mathbf{v})$ depends on $\mathbf{D}\mathbf{v} := \frac{1}{2}(\nabla\mathbf{v} + \nabla\mathbf{v}^\top)$, the symmetric part of the velocity gradient $\nabla\mathbf{v}$. The vector $\mathbf{f} = (f_1, \ldots, f_n)$ is the external body force, and the prescribed divergence of the velocity g has to satisfy the compatibility condition $\int_\Omega g\, dx = 0$. Physical interpretation and discussion of some non-Newtonian fluid models can be found, e.g., in [8, 27, 26].

Throughout the paper we assume that the extra stress tensor $\mathbf{S}$ has (p, δ)-structure (cf. Assumption 2.3) and the relevant example which falls into this class is

$$\mathbf{S}(\mathbf{D}\mathbf{v}) = \mu(\delta + |\mathbf{D}\mathbf{v}|)^{p-2}\mathbf{D}\mathbf{v},$$

with $p \in (1, \infty)$, $\delta \geq 0$, and $\mu > 0$.

The mathematical investigation of fluids with shear-dependent viscosities started with the celebrated work of Ladyzhenskaya (cf. [22]). In recent years there has been an enormous progress in the understanding of this problem and we refer the reader to [25, 26, 4, 5, 15] and the references therein for a detailed discussion.

The first results regarding the numerical analysis date back to Sandri [32]. Later these results have been improved by Barrett and Liu [3], where the error estimates are presented in the setting of quasi-norms. The notion of quasi-norm is the natural one for this type of problem (cf. [2, 3, 24]). Note that Acerbi and Fusco already used in [1] another equivalent expression, which relies on the non-linear quantity $\mathbf{F}$ defined in (2.11) (cf. Remark 2.16 for a comparison of the different approaches). We refer to all these equivalent quantities as the *natural distance*.

[*]Dipartimento di Matematica Applicata "U.Dini", Università di Pisa, Via F. Buonarroti 1/c, I-56127 Pisa, ITALY. (berselli@dma.unipi.it)

[†]Institute of Applied Mathematics, Albert-Ludwigs-University Freiburg, Eckerstr. 1, D-79104 Freiburg, GERMANY. (belenki@mathematik.uni-freiburg.de, diening@mathematik.uni-freiburg.de, rose@mathematik.uni-freiburg.de)

Closely related to the p-Stokes problem is the p-Laplace equation. There has been an intense research activity regarding its numerical analysis in the last two decades, see for instance the recent results in [13, 14, 16, 17]. Going from p-Laplace equation to p-Stokes system involves additional difficulties arising from the pressure and the divergence constraint.

A fundamental tool in our analysis of systems with (p, δ)-structure is the use of Orlicz functions. Indeed, many of our non-linear estimates are deduced by means of linear estimates for a family of *shifted-N-functions* defined in (2.12) (cf. Theorems 3.5, 3.6, 4.2). The use of Orlicz functions enables a unified treatment of the cases $p \geq 2$ and $p \leq 2$ and makes the proofs simpler and clearer.

Outline of the paper: In the Section 2 we introduce the main notation, the basic assumptions and the precise formulation of the problem. Moreover, we present the main results of the paper, i.e. optimal error estimates for the velocity and (in some cases) for the pressure, under natural regularity assumptions. We provide a short outline of the ideas which leads to the results and compare our results with those of Barrett and Liu [3]. The proofs will be postponed to the following sections. We prove the best-approximation error for the velocity and for the pressure, in Section 3 and in Section 4, respectively. In Section 5 we use the best-error estimates in order to prove convergence rates in terms of the mesh size under natural regularity assumptions on the velocity and the pressure. In Section 6 we prove error estimates in terms of the mesh size under assumptions on the data $\mathbf{f}$ and the velocity. Finally, in Section 7 we present results of some numerical experiments and compare them with the theoretical results of the previous sections. An Appendix is also added, where we prove or recall some rather technical results which are used in the paper.

Results similar to the ones proved in the present paper have been obtained at the same time and independently by A. Hirn [20]. Instead of relying on inf-sup-stable elements he uses the so-called local *pressure stabilization* in the context of quadrilateral elements. We thank the author for having put at disposal a draft of his results and for interesting discussions on comparing the two different approaches.

2. The p-Stokes problem: notation and main results. In this section we introduce the notation we will use, we state the precise assumptions on the extra stress tensor $\mathbf{S}$, and we give the main existence and regularity results for the p-Stokes problem and its discrete counterpart.

2.1. Function spaces. We use c, C to denote generic constants, which may change from line to line, but not depending on the crucial quantities. Moreover we write $f \sim g$ if and only if there exists constants $c, C > 0$ such that $c\, f \leq g \leq C\, f$.

We will use the customary Lebesgue spaces $L^p(\Omega)$ and Sobolev spaces $W^{k,p}(\Omega)$, where $\Omega \subset \mathbb{R}^n$ is a bounded, polyhedral domain. We will denote by $\|\,.\,\|_p$ the norm in $L^p(\Omega)$ and by $\|\,.\,\|_{k,p}$ the norm in $W^{k,p}(\Omega)$. The space $W^{1,p}_0(\Omega)$ is the closure of the compactly supported, smooth functions $C^\infty_0(\Omega)$ in $W^{1,p}(\Omega)$. We equip $W^{1,p}_0(\Omega)$ (based on the Poincaré Lemma) with the gradient norm $\|\nabla\cdot\|_p$. For a normed space X we denote its topological dual space by X^*. We denote by $|M|$ the n-dimensional Lebesgue measure of a measurable set M. The mean value of a locally integrable function f over a measurable set $M \subset \Omega$ is denoted by $\langle f\rangle_M := \fint_M f\, dx = \frac{1}{|M|}\int_M f\, dx$. Moreover, we use the notation $\langle f, g\rangle := \int_\Omega fg\, dx$, whenever the right-hand side is well defined.

We will also use Orlicz and Sobolev–Orlicz spaces (cf. [29]). A real convex function

$\psi : \mathbb{R}^{\geq 0} \to \mathbb{R}^{\geq 0}$ is said to be an *N-function**, if $\psi(0) = 0$, $\psi(t) > 0$ for $t > 0$, $\lim_{t\to 0} \psi(t)/t = 0$, as well as $\lim_{t\to\infty} \psi(t)/t = \infty$. As a consequence there exists ψ', the right derivative of ψ, which is non-decreasing and satisfies $\psi'(0) = 0$, $\psi'(t) > 0$ for $t > 0$, and $\lim_{t\to\infty} \psi'(t) = \infty$. We define the *conjugate N-function* ψ^* by $\psi^*(t) := \sup_{s\geq 0}(st - \psi(s))$ for all $t \geq 0$. If ψ' is strictly increasing and therefore invertible, then $(\psi^*)' = (\psi')^{-1}$. A given N-function ψ satisfies the Δ_2–condition, if there exists $K > 0$ such that for all $t \geq 0$ holds $\psi(2\,t) \leq K\,\psi(t)$. We denote the smallest such constant by $\Delta_2(\psi)$. In the following we always assume that ψ and ψ^* satisfy the Δ_2–condition. Under this condition we have

$$\psi^*(\psi'(t)) \sim \psi(t). \tag{2.1}$$

We denote by $L^\psi(\Omega)$ and $W^{1,\psi}(\Omega)$ the classical Orlicz and Sobolev-Orlicz spaces, i.e, $f \in L^\psi(\Omega)$ if the *modular* $\int_\Omega \psi(|f|)\,dx$ is finite and $f \in W^{1,\psi}(\Omega)$ if $f, \nabla f \in L^\psi(\Omega)$. Equipped with the Luxembourg norm $\|f\|_\psi := \inf\{\lambda > 0 : \int_\Omega \psi(|f|/\lambda)\,dx \leq 1\}$ the space $L^\psi(\Omega)$ becomes a Banach space. The same holds for the space $W^{1,\psi}(\Omega)$ if it is equipped with the norm $\|\cdot\|_\psi + \|\nabla\cdot\|_\psi$. Note that the dual space $(L^\psi(\Omega))^*$ can be identified with the space $L^{\psi^*}(\Omega)$. By $W_0^{1,\psi}(\Omega)$ we denote the closure of $C_0^\infty(\Omega)$ in $W^{1,\psi}(\Omega)$ and equip it with the gradient norm $\|\nabla\cdot\|_\psi$. By $L_0^\psi(\Omega)$ and $C_{0,0}^\infty(\Omega)$ we denote the subspace of $L^\psi(\Omega)$ and $C_0^\infty(\Omega)$, respectively, consisting of functions f with vanishing mean value, i.e., $\langle f\rangle_\Omega = 0$.

We need the following refined version of the Young inequality: for all $\varepsilon > 0$ there exists $c_\varepsilon > 0$, depending only on $\Delta_2(\psi), \Delta_2(\psi^*) < \infty$, such that for all $s, t \geq 0$ it holds

$$\begin{aligned} ts &\leq \varepsilon\,\psi(t) + c_\varepsilon\,\psi^*(s)\,, \\ t\,\psi'(s) + \psi'(t)\,s &\leq \varepsilon\,\psi(t) + c_\varepsilon\,\psi(s)\,. \end{aligned} \tag{2.2}$$

2.2. Basic properties of the extra stress tensor. In the whole paper we assume that the extra stress tensor $\mathbf{S}$ has (p, δ)-structure, which will be defined now. A detailed discussion and full proofs can be found in [12, 31]. For a tensor $\mathbf{A} \in \mathbb{R}^{n\times n}$ we denote its symmetric part by $\mathbf{A}^{\text{sym}} := \frac{1}{2}(\mathbf{A} + \mathbf{A}^\top) \in \mathbb{R}^{n\times n}_{\text{sym}} := \{\mathbf{A} \in \mathbb{R}^{n\times n} \,|\, \mathbf{A} = \mathbf{A}^\top\}$. The scalar product between two tensors $\mathbf{A}, \mathbf{B}$ is denoted by $\mathbf{A} \cdot \mathbf{B}$, and we use the notation $|\mathbf{A}|^2 = \mathbf{A} \cdot \mathbf{A}^\top$.

Assumption 2.3 (extra stress tensor). *We assume that the extra stress tensor* $\mathbf{S}\colon \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}_{\text{sym}}$ *belongs to* $C^0(\mathbb{R}^{n\times n}, \mathbb{R}^{n\times n}_{\text{sym}}) \cap C^1(\mathbb{R}^{n\times n} \setminus \{\mathbf{0}\}, \mathbb{R}^{n\times n}_{\text{sym}})$*, satisfies* $\mathbf{S}(\mathbf{A}) = \mathbf{S}\big(\mathbf{A}^{\text{sym}}\big)$*, and* $\mathbf{S}(\mathbf{0}) = \mathbf{0}$*. Moreover, we assume that the tensor* $\mathbf{S}$ *has* (p, δ)*-*structure*, i.e., there exist* $p \in (1, \infty)$*,* $\delta \in [0, \infty)$*, and constants* $C_0, C_1 > 0$ *such that*

$$\sum\nolimits_{i,j,k,l=1}^n \partial_{kl} S_{ij}(\mathbf{A}) C_{ij} C_{kl} \geq C_0 \big(\delta + |\mathbf{A}^{\text{sym}}|\big)^{p-2} |\mathbf{C}^{\text{sym}}|^2, \tag{2.4a}$$

$$\big|\partial_{kl} S_{ij}(\mathbf{A})\big| \leq C_1 \big(\delta + |\mathbf{A}^{\text{sym}}|\big)^{p-2}, \tag{2.4b}$$

are satisfied for all $\mathbf{A}, \mathbf{C} \in \mathbb{R}^{n\times n}$ *with* $\mathbf{A}^{\text{sym}} \neq \mathbf{0}$ *and all* $i, j, k, l = 1, \ldots, n$*. The constants* C_0*,* C_1*, and* p *are called the* characteristics *of* $\mathbf{S}$*.*

Remark 2.5. We would like to emphasize that, if not otherwise stated, the constants in the paper depend only on the characteristics of $\mathbf{S}$ but are independent of $\delta \geq 0$.

*N stands for "nice".

Defining for $t \geq 0$ a special N-function $\varphi = \varphi_{p,\delta}$ by

$$\varphi(t) := \int_0^t \varphi'(s)\,ds \qquad \text{with} \quad \varphi'(t) := (\delta + t)^{p-2} t\,, \tag{2.6}$$

we can replace in the right-hand side of (2.4) $C_i\big(\delta + |\mathbf{A}^{\mathrm{sym}}|\big)^{p-2}$ by $\widetilde{C}_i\,\varphi''\big(|\mathbf{A}^{\mathrm{sym}}|\big)$, $i = 0, 1$. The function φ satisfies uniformly in t the important equivalence

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

since $\min\{1, p-1\}\,(\delta{+}t)^{p-2} \leq \varphi''(t) \leq \max\{1, p-1\}(\delta{+}t)^{p-2}$. Moreover, φ satisfies the Δ_2-condition with $\Delta_2(\varphi) \leq c\,2^{\max\{2,p\}}$ (hence independent of δ). This implies that, uniformly in t, we have

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

The conjugate function φ^* satisfies $\varphi^*(t) \sim (\delta^{p-1} + t)^{p'-2} t^2$ with $1 = \frac{1}{p} + \frac{1}{p'}$. Also φ^* satisfies the Δ_2-condition with $\Delta_2(\varphi^*) \leq c\,2^{\max\{2,p'\}}$.

Remark 2.9. An important example of an extra stress $\mathbf{S}$ satisfying assumption 2.3 is given by $\mathbf{S}(\mathbf{A}) = \varphi'(|\mathbf{A}^{\mathrm{sym}}|)|\mathbf{A}^{\mathrm{sym}}|^{-1}\mathbf{A}^{\mathrm{sym}}$. In this case the characteristics of $\mathbf{S}$, namely C_0, C_1, and p, depend only on p and are independent of $\delta \geq 0$.

Remark 2.10. It is possible to adapt the approach presented here to cover also the situation treated in [3]. In fact, the assumption (A) in that paper can be treated by choosing $\varphi'(t) = \big(t^\alpha (1+t)^{1-\alpha}\big)^{p-2} t$ in (2.6) and formulating Assumption 2.3 directly with this N-function.

Closely related to the extra stress tensor $\mathbf{S}$ with (p, δ)-structure is the function $\mathbf{F}\colon \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}_{\mathrm{sym}}$ defined through

$$\mathbf{F}(\mathbf{A}) := \big(\delta + |\mathbf{A}^{\mathrm{sym}}|\big)^{\frac{p-2}{2}} \mathbf{A}^{\mathrm{sym}}\,. \tag{2.11}$$

Another important tool are the shifted N-functions $\{\varphi_a\}_{a\geq 0}$, cf. [12, 13, 31], defined for $t \geq 0$ by

$$\varphi_a(t) := \int_0^t \varphi_a'(s)\,ds \qquad \text{with} \qquad \varphi_a'(t) := \varphi'(a+t)\frac{t}{a+t}. \tag{2.12}$$

For the (p, δ)-structure we have that $\varphi_a(t) \sim (\delta + a + t)^{p-2} t^2$ and also $(\varphi_a)^*(t) \sim ((\delta + a)^{p-1} + t)^{p'-2} t^2$. The families $\{\varphi_a\}_{a\geq 0}$ and $\{(\varphi_a)^*\}_{a\geq 0}$ satisfy the Δ_2-condition uniformly in $a \geq 0$, with $\Delta_2(\varphi_a) \leq c\,2^{\max\{2,p\}}$ and $\Delta_2((\varphi_a)^*) \leq c\,2^{\max\{2,p\}}$, respectively.

The connection between $\mathbf{S}$, $\mathbf{F}$, and $\{\varphi_a\}_{a\geq 0}$ is best explained by the following lemma (cf. [12, 31]).

Lemma 2.13. *Let* $\mathbf{S}$ *satisfy Assumption 2.3, let* φ *be defined in* (2.6)*, and let* $\mathbf{F}$ *be defined in* (2.11)*. Then*

$$\big(\mathbf{S}(\mathbf{P}) - \mathbf{S}(\mathbf{Q})\big) \cdot \big(\mathbf{P} - \mathbf{Q}\big) \sim \big|\mathbf{F}(\mathbf{P}) - \mathbf{F}(\mathbf{Q})\big|^2 \tag{2.14a}$$

$$\sim \varphi_{|\mathbf{P}^{\mathrm{sym}}|}\big(|\mathbf{P}^{\mathrm{sym}} - \mathbf{Q}^{\mathrm{sym}}|\big) \tag{2.14b}$$

$$\sim \varphi''\big(|\mathbf{P}^{\mathrm{sym}}| + |\mathbf{Q}^{\mathrm{sym}}|\big)|\mathbf{P}^{\mathrm{sym}} - \mathbf{Q}^{\mathrm{sym}}|^2 \tag{2.14c}$$

uniformly in $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^{n\times n}$. *Moreover, uniformly in* $\mathbf{Q} \in \mathbb{R}^{n\times n}$,

$$\mathbf{S}(\mathbf{Q}) \cdot \mathbf{Q} \sim |\mathbf{F}(\mathbf{Q})|^2 \sim \varphi(|\mathbf{Q}^{\text{sym}}|). \tag{2.14d}$$

The constants depend only on the characteristics of $\mathbf{S}$.

Note that if $\varphi''(0)$ does not exist, the expression in (2.14c) is continuously extended by zero for $|\mathbf{P}^{\text{sym}}| = |\mathbf{Q}^{\text{sym}}| = 0$. Moreover,

$$|\mathbf{S}(\mathbf{P}) - \mathbf{S}(\mathbf{Q})| \sim \varphi'_{|\mathbf{P}^{\text{sym}}|}\big(|\mathbf{P}^{\text{sym}} - \mathbf{Q}^{\text{sym}}|\big) \qquad \forall\, \mathbf{P}, \mathbf{Q} \in \mathbb{R}^{n\times n}. \tag{2.15}$$

Remark 2.16 (Natural distance). In view of the previous lemma we have, for all $\mathbf{u}, \mathbf{w} \in (W^{1,\varphi}(\Omega))^n$,

$$\langle \mathbf{S}(\mathbf{Du}) - \mathbf{S}(\mathbf{Dw}), \mathbf{Du} - \mathbf{Dw}\rangle \sim \|\mathbf{F}(\mathbf{Du}) - \mathbf{F}(\mathbf{Dw})\|_2^2 \sim \int_\Omega \varphi_{|\mathbf{Du}|}(|\mathbf{Du} - \mathbf{Dw}|)\, dx.$$

The constants depend only on the characteristics of $\mathbf{S}$. The last expression equals the quasi-norm introduced in [3] raised to the power $\rho = \max\{p, 2\}$. This ensures that our results can also be expressed in terms of the quasi-norm. We refer to all three equivalent quantities as the *natural distance.*

In view of Lemma 2.13 one can deduce many useful properties of the natural distance and of the quantities $\mathbf{F}$, $\mathbf{S}$ from the corresponding properties of the shifted N-functions $\{\varphi_a\}$. For example the following important estimates follow directly from (2.15), Young's inequality (2.2), and (2.14).

Lemma 2.17. *For all* $\varepsilon > 0$, *there exist a constant* $c_\varepsilon > 0$ *depending only on* $\varepsilon > 0$ *and the characteristics of* $\mathbf{S}$ *such that for all sufficiently smooth vector fields* $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *we have*

$$\langle \mathbf{S}(\mathbf{Du}) - \mathbf{S}(\mathbf{Dv}), \mathbf{Dw} - \mathbf{Dv}\rangle \le \varepsilon \|\mathbf{F}(\mathbf{Du}) - \mathbf{F}(\mathbf{Dv})\|_2^2 + c_\varepsilon \|\mathbf{F}(\mathbf{Dw}) - \mathbf{F}(\mathbf{Dv})\|_2^2 .$$

2.3. The p-Stokes problem.

Let us briefly recall some well-known facts about the p-Stokes system (1.1). We define the function spaces

$$\begin{aligned} X &:= \big(W^{1,p}(\Omega)\big)^n, & V &:= \big(W^{1,p}_0(\Omega)\big)^n, \\ Y &:= L^{p'}(\Omega), & Q &:= L^{p'}_0(\Omega) := \Big\{ f \in L^{p'}(\Omega) : \fint_\Omega f\, dx = 0 \Big\}. \end{aligned}$$

Remark 2.18. For the special N-function $\psi = \varphi_a$, with $a \in [0, a_0]$, $\delta \in [0, \delta_0]$, and $p \in [p_0, p_1]$, we get $L^{\psi^*}(\Omega) = L^{p'}(\Omega)$ and $W^{1,\psi}(\Omega) = W^{1,p}(\Omega)$ with uniform equivalence of the corresponding norms depending on a and p, since Ω is bounded.

With this notation the weak formulation of problem (1.1) is the following.

Problem (Q). For $(\mathbf{f}, g) \in V^* \times Y^*$ with $\langle g, 1\rangle = 0$ find $(\mathbf{v}, q) \in V \times Q$ such that

$$\begin{aligned} \langle \mathbf{S}(\mathbf{Dv}), \mathbf{D}\boldsymbol{\xi}\rangle - \langle \operatorname{div} \boldsymbol{\xi}, q\rangle &= \langle \mathbf{f}, \boldsymbol{\xi}\rangle & \forall\, \boldsymbol{\xi} &\in V, \\ -\langle \operatorname{div} \mathbf{v}, \eta\rangle &= \langle g, \eta\rangle & \forall\, \eta &\in Y. \end{aligned}$$

The condition $\langle g, 1\rangle = 0$ comes from the compatibility condition with the zero boundary values of the velocity.

Alternatively, we can reformulate the problem "hiding" the pressure:

Problem (P). For $(\mathbf{f}, g) \in V^* \times Y^*$ with $\langle g, 1\rangle = 0$ find $\mathbf{v} \in V(g)$ such that

$$\langle \mathbf{S}(\mathbf{D}\mathbf{v}), \mathbf{D}\boldsymbol{\xi}\rangle = \langle \mathbf{f}, \boldsymbol{\xi}\rangle \qquad \forall\, \boldsymbol{\xi} \in V(0),$$

where $V(g) := \{\mathbf{w} \in V \,:\, -\langle \operatorname{div} \mathbf{w}, \eta\rangle = \langle g, \eta\rangle \;\forall\, \eta \in Y\}$.

The names "Problem **(Q)**" and "Problem **(P)**" are traditional, see [9]. Note that $V(g) \neq \emptyset$ due to the solvability of the divergence equation (cf. Theorem 4.2). This and the theory of monotone operators (cf. [23]) easily yields the existence of an unique weak solution of the problem **(P)**. The Theorem of DeRham, the solvability of the divergence equation and the negative norm theorem then ensure the solvability of the problem **(Q)** (cf. [6] for more details).

The problems **(Q)** and **(P)** have a discrete counterpart, whose analysis is the ultimate goal of this paper. Let $\mathcal{T}_h$ be a family of shape regular triangulations of our domain Ω consisting of n-dimensional simplices K with diameter h_K less than h. For a simplex $K \in \mathcal{T}_h$ we denote by ρ_K the supremum of the diameters of inscribed balls. We assume that there exists a constant γ_0 independent on h and $K \in \mathcal{T}_h$ such that $h_K \rho_K^{-1} \leq \gamma_0$. Let S_K denote the neighborhood of K, i.e., the patch S_K is the union of all simplices of $\mathcal{T}_h$ touching K. One easily sees that under these assumptions we get that $|K| \sim |S_K|$ and that the number of simplices in S_K and the number of patches to which a simplex belongs to are uniformly bounded with respect to $h > 0$ and $K \in \mathcal{T}_h$.

We denote by $\mathcal{P}_m(\mathcal{T}_h)$, with $m \in \mathbb{N}_0$, the space of scalar or vector-valued continuous functions, which are polynomials of degree at most m on each simplex $K \in \mathcal{T}_h$. Given a triangulation of Ω with the above properties and given $k, m \in \mathbb{N}_0$ we denote by $X_h \subset \mathcal{P}_m(\mathcal{T}_h)$ and $Y_h \subset \mathcal{P}_k(\mathcal{T}_h)$ appropriate conforming finite element spaces defined on $\mathcal{T}_h$, i.e., X_h, Y_h satisfy $X_h \subset X$ and $Y_h \subset Y$. Moreover, we set $V_h := X_h \cap V$ and $Q_h := Y_h \cap Q$. Now the discrete counterpart of **(P)** and **(Q)** can be written as follows:

Problem (Q$_h$). For $(\mathbf{f}, g) \in V^* \times Y^*$ with $\langle g, 1\rangle = 0$ find $(\mathbf{v}_h, q_h) \in V_h \times Q_h$ such that

$$\begin{aligned} \langle \mathbf{S}(\mathbf{D}\mathbf{v}_h), \mathbf{D}\boldsymbol{\xi}_h\rangle - \langle \operatorname{div} \boldsymbol{\xi}_h, q_h\rangle &= \langle \mathbf{f}, \boldsymbol{\xi}_h\rangle \qquad \forall\, \boldsymbol{\xi}_h \in V_h, \\ -\langle \operatorname{div} \mathbf{v}_h, \eta_h\rangle &= \langle g, \eta_h\rangle \qquad \forall\, \eta_h \in Q_h. \end{aligned} \tag{2.19}$$

If $(\mathbf{v}_h, q_h) \in V_h \times Q_h$ is a solution of the "Problem (Q_h)" then $(2.19)_2$ is satisfied for all $\eta_h \in Y_h$, since $\operatorname{div} \mathbf{v}_h$ and g are orthogonal to constants.

Problem (P$_h$). For $(\mathbf{f}, g) \in V^* \times Y^*$ with $\langle g, 1\rangle = 0$ find $\mathbf{v}_h \in V_h(g)$ such that

$$\langle \mathbf{S}(\mathbf{D}\mathbf{v}_h), \mathbf{D}\boldsymbol{\xi}_h\rangle = \langle \mathbf{f}, \boldsymbol{\xi}_h\rangle \qquad \forall\, \boldsymbol{\xi}_h \in V_h(0),$$

where $V_h(g) := \{\mathbf{w}_h \in V_h \,:\, -\langle \operatorname{div} \mathbf{w}_h, \eta_h\rangle = \langle g, \eta_h\rangle \;\forall\, \eta_h \in Y_h\}$.

For the well-possedness of the problem **(P$_h$)** we certainly have to assume that $V_h(g) \neq \emptyset$. However, this fact follows immediately from our Assumption 2.20 on the interpolation operator (see below), since $-\operatorname{div} \Pi_h^{\operatorname{div}} \mathbf{v} = g$ in Y_h^* and $V(g) \neq \emptyset$. Also note that in general $V_h(g) \not\subset V(g)$, although $V_h \subset V$. The existence of an unique weak solution of the problem **(P$_h$)** follows as for the problem **(P)**. The solvability of the problem **(Q$_h$)** then follows from the discrete inf-sup condition of Lemma 4.1 (cf. [6] for more details).

2.4. Main results. In this section we state the main error estimates of the paper and compare them with the previous results in the literature. Throughout the paper we will make the following assumptions on the finite element spaces for velocity and pressure that we consider.

Assumption 2.20. *We assume that $\mathcal{P}_1(\mathcal{T}_h) \subset X_h$ and there exists a linear projection operator $\Pi_h^{\mathrm{div}} : X \to X_h$ which*

(a) *preserves divergence in the Y_h^*-sense, i.e.,*

$$\langle \operatorname{div} \mathbf{w}, \eta_h \rangle = \langle \operatorname{div} \Pi_h^{\mathrm{div}} \mathbf{w}, \eta_h \rangle \qquad \forall\, \mathbf{w} \in X,\ \forall\, \eta_h \in Y_h\,; \tag{2.21}$$

(b) *preserves zero boundary values, i.e. $\Pi_h^{\mathrm{div}}(V) \subset V_h$;*

(c) *is locally $W^{1,1}$-stable in the sense that*

$$\fint_K |\Pi_h^{\mathrm{div}} \mathbf{w}|\, dx \le c \fint_{S_K} |\mathbf{w}|\, dx + c \fint_{S_K} h_K |\nabla \mathbf{w}|\, dx \quad \forall\, \mathbf{w} \in X,\ \forall\, K \in \mathcal{T}_h. \tag{2.22}$$

Remark 2.23. Certainly, the existence of Π_h^{div} depends on the choice of X_h and Y_h. It is shown in [9], [18], [19] that Π_h^{div} exists for a variety of spaces X_h and Y_h. This includes the Taylor–Hood, the Crouzeix–Raviart, and the MINI element in dimension two and three (cf. Appendix A.1 where the proof for the MINI element is summarized). The abstract assumptions allow for an easy extension of our results to other choices of X_h and Y_h in future works.

Assumption 2.24. *We assume that Y_h contains the constant functions, i.e. $\mathbb{R} \subset Y_h$, and that there exists a linear projection operator $\Pi_h^Y : Y \to Y_h$ which is locally L^1-stable in the sense that*

$$\fint_K |\Pi_h^Y q|\, dx \le c \fint_{S_K} |q|\, dx \qquad \forall\, q \in Y,\ \forall\, K \in \mathcal{T}_h. \tag{2.25}$$

Remark 2.26. Note that the Clément interpolation operator [9] and the version of the Scott–Zhang interpolation operator (not preserving the boundary conditions) [33] satisfy Assumption 2.24.

Remark 2.27. It is possible to weaken the requirements on the projection operators Π_h^{div} and Π_h^Y. In fact, we can replace the requirement $\Pi_h^{\mathrm{div}} \mathbf{w}_h = \mathbf{w}_h$ for all $\mathbf{w}_h \in X_h$ by the requirement $\Pi_h^{\mathrm{div}} \mathfrak{q} = \mathfrak{q}$ for all linear polynomials (not in the piecewise sense), and the requirement $\Pi_h^Y q_h = q_h$ for all $q_h \in Y_h$ by the requirement $\Pi_h^Y c = c$ for all constants c.

Let us now state our main results and shortly explain the strategy of their proofs. First we prove that the error for the velocity is controlled by some best approximation error for the velocity (with prescribed divergence) and the pressure (cf. Lemma 3.3). This is the counterpart of the standard error estimate for the Stokes problem before applying the inf-sup condition. In our non-linear setting we deviate from the standard way and work directly with a divergence-preserving operator Π_h^{div} (cf. Assumption 2.20). From the local $W^{1,1}$-stability of Π_h^{div}, we derive its non-linear, local counterparts in terms of the natural distance (cf. Theorem 3.7). These new estimates are the main reason, why our results improve previous ones. Thus we can replace the best approximation error for the velocity (with prescribed divergence) by local averages of the solution $\mathbf{v}$ in terms of the natural distance (cf. Theorem 3.9)

Next, we prove that the error for the pressure is also controlled by a best approximation error for the pressure and the velocity (cf. Theorem 4.10 and also Remark 4.11). This result is sensitive to whether $p \ge 2$ or $p \le 2$. The proof is based on extensions of classical results (inf-sup condition, properties of the Bogovskiĭ operator, divergence-preserving projection) to Orlicz spaces.

Once we have at hand these best approximation estimates we obtain convergence rates in terms of the mesh size. More precisely, in Section 5 we will prove the following result:

THEOREM 2.28. *Let Π_h^{div} satisfy Assumption 2.20 and Π_h^Y satisfy Assumption 2.24. Let $(\mathbf{v},q)$ and $(\mathbf{v}_h,q_h)$ be solutions of the problems* **(Q)** *and* **($\mathbf{Q}_h$)**, *respectively. Furthermore, let $\mathbf{F}(\mathbf{D}\mathbf{v}) \in (W^{1,2}(\Omega))^{n\times n}$ and also let $q \in W^{1,\varphi^*}(\Omega) = W^{1,p'}(\Omega)$. Then*

$$\|\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\mathbf{D}\mathbf{v}_h)\|_2 \le c\, h^{\min\{1,\frac{p'}{2}\}}\,, \tag{2.29}$$

$$\|q - q_h\|_{p'} \le c\, h^{\min\left\{\frac{2}{p'},\frac{p'}{2}\right\}}\,, \tag{2.30}$$

$$\int_\Omega \varphi^*(|q-q_h|)\,dx \le c\, h^{\min\left\{2,\frac{(p')^2}{2}\right\}}. \tag{2.31}$$

Let us compare Theorem 2.28 with the previously-known best results, which can be found in [3]: It is an important feature of our paper that (i) the methods and results are independent of the degeneracy parameter $\delta \ge 0$ and (ii) include in particular the degenerate case $\delta = 0$. Therefore, in the following we compare our results only to those results of [3], which are valid for the full range[†] $\delta \ge 0$.

In [3, Theorem 4.1] concrete convergence rates in Sobolev norms for the error in the velocity and the pressure are given. These results are based on error estimates in the quasi-norm given in [3, Theorem 3.2] and certain natural approximation properties of the used finite elements. This step can be improved if one uses Theorem 3.5 of our paper. Therefore, we compare our results in Theorem 2.28 with these improved results of Barrett and Liu.

In [3] for the case $p \le 2$ one gets $\|\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\mathbf{D}\mathbf{v}_h)\|_2 = \mathcal{O}(h^{\frac{p}{2}})$, and $\|q-q_h\|_{p'} = \mathcal{O}(h^{p-1})$. Thus Theorem 2.28 improves both convergence rates by a factor of $\frac{2}{p}$.

In [3] for the case $p \ge 2$ one gets, under the additional assumption that $\mathbf{v} \in (W^{1,\infty}(\Omega))^n$, the estimate $\|\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\mathbf{D}\mathbf{v}_h)\|_2 + \|q-q_h\|_{p'} = \mathcal{O}(h^{\frac{p'}{2}})$. This is the same convergence rate as in Theorem 2.28.

Finally, we would like to mention that the numerical tests in Section 7 indicate that our error estimates are optimal in the case $p \le 2$, while they are only optimal for the velocity in the case $p \ge 2$.

In Section 6 we prove error estimates directly in terms of the regularity of the data $\mathbf{f}$ instead of the pressure q. Due to the lack of appropriate regularity results for the p-Stokes problem we still have to assume some regularity for the velocity (anyway cf. Lemma 6.1 for the space periodic setting, where such results are available). In particular, we prove:

THEOREM 2.32. *Let Π_h^{div} satisfy Assumption 2.20 and Π_h^Y satisfy Assumption 2.24. Let $(\mathbf{v},q)$ and $(\mathbf{v}_h,q_h)$ be solutions of the problems* **(Q)** *and* **($\mathbf{Q}_h$)**, *respectively. Furthermore, let $\mathbf{F}(\mathbf{D}\mathbf{v}) \in (W^{1,2}(\Omega))^{n\times n}$ and $\mathbf{f} \in (L^{\varphi^*}(\Omega))^n = (L^{p'}(\Omega))^n$. Then*

$$\|\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\mathbf{D}\mathbf{v}_h)\|_2 \le c\, h^{\min\{1,\frac{p'}{2}\}}\,. \tag{2.33}$$

[†]It is clear, that the estimates for $\delta > 0$ are better, since for example the case $p \ge 2$ with $\delta > 0$ and $\mathbf{v} \in W^{1,\infty}$ behaves just like an elliptic problem with convergence rate $\mathcal{O}(h)$. See also Remark 6.8 for $\mathbf{f} \in L^2$.

We would like to remark that for $p \le 2$ one can also show that the error estimates of the pressure (2.30) and (2.31) are still valid (cf. Remark 6.8). Moreover, for $p \ge 2$, $\delta > 0$, and $\mathbf{f} \in (L^2(\Omega))^n$ one can improve (2.33) to $\|\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\mathbf{D}\mathbf{v}_h)\|_2 \le c\,h$ (cf. Remark 6.8). These results seem to be new.

3. Best Approximation Error for the Velocity. In this section we prove error estimates for the velocity in terms of best approximation properties measured in the natural distance.

3.1. Equation for the error. Taking the difference between **(Q)** and **(Q$_h$)** we get the following equation for the numerical error

$$\begin{aligned} \langle \mathbf{S}(\mathbf{D}\mathbf{v}) - \mathbf{S}(\mathbf{D}\mathbf{v}_h), \mathbf{D}\boldsymbol{\xi}_h\rangle - \langle \operatorname{div} \boldsymbol{\xi}_h, q - q_h\rangle &= 0 \qquad \forall\, \boldsymbol{\xi}_h \in V_h, \\ -\langle \operatorname{div}(\mathbf{v} - \mathbf{v}_h), \eta_h\rangle &= 0 \qquad \forall\, \eta_h \in Y_h. \end{aligned} \tag{3.1}$$

By the definition of $V_h(0)$ it follows immediately that $\mathbf{v} - \mathbf{v}_h \in V_h(0)$ and

$$\langle \mathbf{S}(\mathbf{D}\mathbf{v}) - \mathbf{S}(\mathbf{D}\mathbf{v}_h), \mathbf{D}\boldsymbol{\xi}_h\rangle = \langle \operatorname{div} \boldsymbol{\xi}_h, q - \eta_h\rangle \quad \forall\, \boldsymbol{\xi}_h \in V_h(0),\ \forall\, \eta_h \in Y_h. \tag{3.2}$$

We start with a preliminary approximation result which will be improved later on in Theorem 3.9.

LEMMA 3.3. *Let* $(\mathbf{v}, q)$ *and* $(\mathbf{v}_h, q_h)$ *be the solutions of the problems* **(Q)** *and* **(Q$_h$)***, respectively. Then we have the following estimate*

$$\begin{aligned} \|\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\mathbf{D}\mathbf{v}_h)\|_2^2 \le c \inf_{\mathbf{w}_h \in V_h(g)} &\|\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\mathbf{D}\mathbf{w}_h)\|_2^2 \\ &+ c \inf_{\mu_h \in Y_h} \int_\Omega (\varphi_{|\mathbf{D}\mathbf{v}|})^*(|q - \mu_h|)\, dx. \end{aligned} \tag{3.4}$$

Proof. For $\mathbf{w}_h \in V_h(g)$ we have $\mathbf{v}_h - \mathbf{w}_h \in V_h(0)$. Consequently for all $\mu_h \in Y_h$, we obtain with Lemma 2.13 and (3.2) that

$$\begin{aligned} c\,\|\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\mathbf{D}\mathbf{v}_h)\|_2^2 &\le \langle \mathbf{S}(\mathbf{D}\mathbf{v}) - \mathbf{S}(\mathbf{D}\mathbf{v}_h), \mathbf{D}\mathbf{v} - \mathbf{D}\mathbf{v}_h\rangle \\ &= \langle \mathbf{S}(\mathbf{D}\mathbf{v}) - \mathbf{S}(\mathbf{D}\mathbf{v}_h), \mathbf{D}\mathbf{v} - \mathbf{D}\mathbf{w}_h\rangle + \langle \mathbf{S}(\mathbf{D}\mathbf{v}) - \mathbf{S}(\mathbf{D}\mathbf{v}_h), \mathbf{D}\mathbf{w}_h - \mathbf{D}\mathbf{v}_h\rangle \\ &= \langle \mathbf{S}(\mathbf{D}\mathbf{v}) - \mathbf{S}(\mathbf{D}\mathbf{v}_h), \mathbf{D}\mathbf{v} - \mathbf{D}\mathbf{w}_h\rangle - \langle \operatorname{div}(\mathbf{w}_h - \mathbf{v}_h), q - \mu_h\rangle. \end{aligned}$$

Now Lemma 2.17 shows that for any given $\varepsilon > 0$ there exists $c_\varepsilon > 0$ such that

$$\begin{aligned} \big|\langle \mathbf{S}(\mathbf{D}\mathbf{v}) - \mathbf{S}(\mathbf{D}\mathbf{v}_h), \mathbf{D}\mathbf{v} - \mathbf{D}\mathbf{w}_h\rangle\big| \le \varepsilon\, &\|\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\mathbf{D}\mathbf{v}_h)\|_2^2 \\ &+ c_\varepsilon\, \|\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\mathbf{D}\mathbf{w}_h)\|_2^2. \end{aligned}$$

Next, we estimate the term involving $q - \mu_h$. We add and subtract $\mathbf{D}\mathbf{v}$, use Young's inequality (2.2) for $\varphi_{|\mathbf{D}\mathbf{v}|}$, and apply Lemma 2.13 to obtain

$$\begin{aligned} &\big|\langle \operatorname{div}(\mathbf{v}_h - \mathbf{w}_h), q - \mu_h\rangle\big| \le \int_\Omega \big(|\mathbf{D}\mathbf{v}_h - \mathbf{D}\mathbf{v}| + |\mathbf{D}\mathbf{v} - \mathbf{D}\mathbf{w}_h|\big)\, |q - \mu_h|\, dx \\ &\le \varepsilon \int_\Omega \varphi_{|\mathbf{D}\mathbf{v}|}(|\mathbf{D}\mathbf{v}_h - \mathbf{D}\mathbf{v}|) + \varphi_{|\mathbf{D}\mathbf{v}|}(|\mathbf{D}\mathbf{w}_h - \mathbf{D}\mathbf{v}|)\, dx + c_\varepsilon \int_\Omega (\varphi_{|\mathbf{D}\mathbf{v}|})^*(|q - \mu_h|)\, dx \\ &\le \varepsilon\, c \Big(\|\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\mathbf{D}\mathbf{v}_h)\|_2^2 + \|\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\mathbf{D}\mathbf{w}_h)\|_2^2 \Big) + c_\varepsilon \int_\Omega (\varphi_{|\mathbf{D}\mathbf{v}|})^*(|q - \mu_h|)\, dx. \end{aligned}$$

Collecting the estimates and choosing $\varepsilon > 0$ small enough we obtain the assertion by noting that $\mathbf{w}_h \in V_h(g)$ and $\mu_h \in Y_h$ are arbitrary. ∎

This result is the counterpart of the standard error estimate for the linear Stokes problem, since the quantity $\mathbf{F}$ appears naturally in the non-linear p-Stokes problem (cf. Remark 2.16). The inequality (3.4) provides a (non-linear) error estimate in terms of best approximation quantities. The first term on the right-hand side of (3.4) denotes the best approximation of $\mathbf{v}$ in terms of the natural distance, among all discrete functions $\mathbf{w}_h \in V_h$ with prescribed divergence $-\operatorname{div} \mathbf{w}_h = g$ in Y_h^*. This drawback of estimate (3.4) is resolved by using the divergence-preserving operator $\Pi_h^{\text{div}} : X \to X_h$ from Assumption 2.20, since $\mathbf{v}$ has the correct divergence, i.e. $-\operatorname{div} \mathbf{v} = g$ in Y^*.

3.2. The divergence-preserving interpolation operator. In this section we derive the estimates for the divergence preserving operator Π_h^{div} from Assumption 2.20 in the natural distance. The technique is similar to the non-linear estimates for the Scott-Zhang interpolation operator obtained in [16]. Additional difficulties arise due to the symmetric gradient and the constraint on the divergence.

We start with the local continuity and approximability result of Π_h^{div} in terms of Orlicz spaces.

Theorem 3.5 (Orlicz-Continuity/Orlicz-Approximability). *Let ψ be an N-function with $\Delta_2(\psi) < \infty$ and let Π_h^{div} satisfy Assumption 2.20. Then Π_h^{div} has the local Orlicz-continuity property*

$$\int_K \psi\big(|\nabla \Pi_h^{\text{div}} \mathbf{w}|\big)\, dx \leq c \int_{S_K} \psi\big(|\nabla \mathbf{w}|\big)\, dx$$

and the local Orlicz-approximability property

$$\int_K \psi\big(|\mathbf{w} - \Pi_h^{\text{div}} \mathbf{w}|\big)\, dx + \int_K \psi\big(h_K |\nabla \mathbf{w} - \nabla \Pi_h^{\text{div}} \mathbf{w}|\big)\, dx \leq c \int_{S_K} \psi\big(h_K |\nabla \mathbf{w}|\big)\, dx,$$

for all $K \in \mathcal{T}_h$ and $\mathbf{w} \in (W^{1,\psi}(\Omega))^n$. The constant c depends only on $\Delta_2(\psi)$ and on the non-degeneracy constant γ_0 of the triangulation $\mathcal{T}_h$.

Proof. It follows from Assumption 2.20 and the usual inverse estimates that Π_h^{div} satisfies Assumption 4.1 of [16] with $l = l_0 = r_0 = 1$. Therefore, the local Orlicz-continuity follows from [16, Corollary 4.8] and the local Orlicz-approximability follows from [16, Theorem 4.6]. ∎

The above result is formulated in terms of $\nabla \mathbf{w}$ while in (3.4) appear symmetric gradients. To deal with this problem we need Korn's inequality in Orlicz spaces. The following result is a special case of [11, Theorem 6.13], proved for John domains. We can apply this result, since the neighborhood S_K are John domains with uniform John constant (depending on the mesh degeneracy).

Theorem 3.6 (Korn's inequality in Orlicz spaces). *Let ψ be an N-function with $\Delta_2(\psi)$, $\Delta_2(\psi^*) < \infty$. Then for all $K \in \mathcal{T}_h$ and all $\mathbf{w} \in (W^{1,\psi}(S_K))^n$ it holds that*

$$\int_{S_K} \psi\big(\big|\nabla \mathbf{w} - \langle \nabla \mathbf{w} \rangle_{S_K}\big|\big)\, dx \leq c \int_{S_K} \psi\big(|\mathbf{D}\mathbf{w} - \langle \mathbf{D}\mathbf{w} \rangle_{S_K}|\big)\, dx.$$

The constant c depends only on γ_0, $\Delta_2(\psi)$, and $\Delta_2(\psi^*)$.

Next, we present the estimates concerning Π_h^{div} in terms of the natural distance.

THEOREM 3.7. *Let Π_h^{div} satisfy Assumption 2.20. Then we have uniformly with respect to $K \in \mathcal{T}_h$ and to $\mathbf{w} \in (W^{1,\varphi}(\Omega))^n$*

$$\int_K \big|\mathbf{F}(\mathbf{D}\mathbf{w}) - \mathbf{F}(\mathbf{D}\Pi_h^{\mathrm{div}}\mathbf{w})\big|^2\,dx \le c \int_{S_K} \big|\mathbf{F}(\mathbf{D}\mathbf{w}) - \langle \mathbf{F}(\mathbf{D}\mathbf{w})\rangle_{S_K}\big|^2\,dx,$$

with c depending mainly on γ_0.

Proof. Let $\mathbf{w} \in (W^{1,\varphi}(\Omega))^n$, then Lemma 2.13 implies $\mathbf{F}(\mathbf{D}\mathbf{w}) \in (L^2(\Omega))^{n\times n}$. Fix one $K \in \mathcal{T}_h$ and choose a linear function $\mathfrak{p}$, defined on Ω, with $\nabla\mathfrak{p} = \langle \nabla \mathbf{w}\rangle_{S_K}$. Consequently we have also $\mathbf{D}\mathfrak{p} = \langle \mathbf{D}\mathbf{w}\rangle_{S_K}$. Using Lemma 2.13, the triangle inequality, and $\Pi_h^{\mathrm{div}}\mathfrak{p} = \mathfrak{p}$ we get

$$\begin{aligned}\int_K \big|\mathbf{F}(\mathbf{D}\mathbf{w}) - \mathbf{F}(\mathbf{D}\Pi_h^{\mathrm{div}}\mathbf{w})\big|^2\,dx &\le c\int_K \varphi_{|\mathbf{D}\mathbf{w}|}(|\mathbf{D}\mathbf{w} - \mathbf{D}\Pi_h^{\mathrm{div}}\mathbf{w}|)\,dx\\ &\le c\int_K \varphi_{|\mathbf{D}\mathbf{w}|}(|\mathbf{D}\mathbf{w} - \mathbf{D}\mathfrak{p}|)\,dx + c\int_K \varphi_{|\mathbf{D}\mathbf{w}|}(|\mathbf{D}\Pi_h^{\mathrm{div}}(\mathbf{w}-\mathfrak{p})|)\,dx =: (I) + (II).\end{aligned}$$

We cannot apply directly the Orlicz stability Theorem 3.5 to (II), since the shift $|\mathbf{D}\mathbf{w}|$ is not constant. The Orlicz function is space dependent and to avoid this problem we use the shift-change from Lemma A.3. Hence, we bound (II) in the following way

$$(II) \le c\int_K \varphi_{|\mathbf{D}\mathfrak{p}|}(|\mathbf{D}\Pi_h^{\mathrm{div}}(\mathbf{w}-\mathfrak{p})|)\,dx + c\int_K |\mathbf{F}(\mathbf{D}\mathbf{w}) - \mathbf{F}(\mathbf{D}\mathfrak{p})|^2\,dx =: (II)_1 + (II)_2\,.$$

We now estimate the symmetric gradient with the full gradient, use the Orlicz stability, and finally Korn's inequality from Theorem 3.6 to obtain

$$\begin{aligned}(II)_1 &\le c\int_K \varphi_{|\mathbf{D}\mathfrak{p}|}(|\nabla\Pi_h^{\mathrm{div}}(\mathbf{w}-\mathfrak{p})|)\,dx \le c\int_{S_K} \varphi_{|\mathbf{D}\mathfrak{p}|}(|\nabla(\mathbf{w}-\mathfrak{p})|)\,dx\\ &\le c\int_{S_K} \varphi_{|\mathbf{D}\mathfrak{p}|}(|\mathbf{D}(\mathbf{w}-\mathfrak{p})|)\,dx,\end{aligned}$$

where in the last step we used that $\nabla\mathfrak{p} = \langle \nabla\mathbf{w}\rangle_{S_K}$ and $\mathbf{D}\mathfrak{p} = \langle \mathbf{D}\mathbf{w}\rangle_{S_K}$. Collecting all results and using Lemma 2.13 we have overall shown

$$\int_K \big|\mathbf{F}(\mathbf{D}\mathbf{w}) - \mathbf{F}(\mathbf{D}\Pi_h^{\mathrm{div}}\mathbf{w})\big|^2\,dx \le c\int_{S_K} |\mathbf{F}(\mathbf{D}\mathbf{w}) - \mathbf{F}(\langle \mathbf{D}\mathbf{w}\rangle_{S_K})|^2\,dx.$$

The claim follows by using Lemma A.4. □

REMARK 3.8. (i) We recall that

$$\int_{S_K} |\mathbf{F}(\mathbf{D}\mathbf{w}) - \langle \mathbf{F}(\mathbf{D}\mathbf{w})\rangle_{S_K}|^2\,dx = \inf_{\mathbf{Q}\in\mathbb{R}^{n\times n}_{\mathrm{sym}}} \int_{S_K} \big|\mathbf{F}(\mathbf{D}\mathbf{w}) - \mathbf{F}(\mathbf{Q})\big|^2\,dx\,.$$

(ii) Theorem 3.7 also holds more generally. Assume that the N-function ψ belongs to $C^1(\mathbb{R}^{\ge 0}) \cap C^2(\mathbb{R}^{>0})$, satisfies $\Delta_2(\psi), \Delta_2(\psi^*) < \infty$, has a strictly increasing derivative ψ', and $\psi'(t) \sim t\,\psi''(t)$ uniformly with respect to $t \ge 0$. Then, Theorem 3.7 remains valid if we replace $\mathbf{F}$ by $\mathbf{F}_\psi(\mathbf{D}) := \sqrt{\frac{\psi'(|\mathbf{D}|)}{|\mathbf{D}|}}\,\mathbf{D}$.

3.3. Error of the velocity. Collecting the estimates and results of the previous sections we obtain the most useful error estimate.

THEOREM 3.9. *Let $\Pi_h^{\operatorname{div}}$ satisfy Assumption 2.20. Let* $(\mathbf{v}, q)$ *and* $(\mathbf{v}_h, q_h)$ *be solutions of the problems* **(Q)** *and* **($\mathbf{Q}_h$)**, *respectively. Then*

$$\|\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\mathbf{D}\mathbf{v}_h)\|_2^2 \le c \sum_{K \in \mathcal{T}_h} \int_{S_K} \big|\mathbf{F}(\mathbf{D}\mathbf{v}) - \langle \mathbf{F}(\mathbf{D}\mathbf{v})\rangle_{S_K}\big|^2 \, dx$$
$$+ c \inf_{\mu_h \in Y_h} \int_\Omega (\varphi_{|\mathbf{D}\mathbf{v}|})^*(|q - \mu_h|)\, dx.$$

Proof. Since $\Pi_h^{\operatorname{div}}$ is divergence-preserving (see (2.21)) $\mathbf{v} \in V(g)$ implies that $\Pi_h^{\operatorname{div}}\mathbf{v} \in V_h(g)$. The claim follows from Lemma 3.3 with $\mathbf{w}_h := \Pi_h^{\operatorname{div}}\mathbf{v}$ and Theorem 3.7. □

4. Best Approximation for the pressure.

We are now discussing best approximation results for the pressure. As in the classical Stokes problem we need the discrete inf-sup condition to recover information on the discrete pressure. We start by extending this condition to Orlicz spaces.

4.1. Inf-sup condition on Orlicz spaces.

The standard discrete inf-sup condition reads

$$\exists\, c > 0: \quad \|q_h\|_{Q_h} \le c \sup_{\|\boldsymbol{\xi}_h\|_{V_h} \le 1} \langle q_h, \operatorname{div} \boldsymbol{\xi}_h\rangle \qquad \forall\, q_h \in Q_h\,.$$

However, this is not enough in our setting and, different from the linear case, we need the inf-sup not only in terms of norms (as above) but also in a different form. In particular, we will use the following result.

LEMMA 4.1. *Let $\Pi_h^{\operatorname{div}}$ satisfy Assumption 2.20. Then for all $q_h \in Q_h$ holds*

$$\|q_h\|_{p'} \le c \sup_{\boldsymbol{\xi}_h \in V_h\,:\, \|\boldsymbol{\xi}_h\|_{1,p} \le 1} \langle q_h, \operatorname{div} \boldsymbol{\xi}_h\rangle$$

and also

$$\int_\Omega \varphi^*(|q_h|)\, dx \le \sup_{\boldsymbol{\xi}_h \in V_h} \bigg[\langle q_h, \operatorname{div} \boldsymbol{\xi}_h\rangle - \frac{1}{c} \int_\Omega \varphi(|\nabla \boldsymbol{\xi}_h|)\, dx\bigg],$$

where the constants depend ‡ *only on p and on Ω.*

We postpone the proof of this lemma and first recall the following result from [11, Theorem 6.6], which is proved for John domains (cf. [11] for the precise definition). Let us just mention that every Lipschitz domain is a John domain. We will apply the following two results to Ω and to S_K, which are both John domains.

THEOREM 4.2. *Let $G \subset \mathbb{R}^n$ be a bounded John domain. Then there exists a linear operator* $\mathtt{B} : C_{0,0}^\infty(G) \to (C_0^\infty(G))^n$ *which extends uniquely for all N-functions ψ with* $\Delta_2(\psi), \Delta_2(\psi^*) < \infty$ *to an operator* $\mathtt{B} : L_0^\psi(G) \to (W_0^{1,\psi}(G))^n$ *satisfying* $\operatorname{div} \mathtt{B} f = f$ *and*

$$\|\nabla \mathtt{B} f\|_{L^\psi(G)} \le c\, \|f\|_{L_0^\psi(G)},$$
$$\int_G \psi\big(|\nabla \mathtt{B} f|\big)\, dx \le c \int_G \psi\big(|f|\big)\, dx.$$

The constant c depends on $\Delta_2(\psi)$, $\Delta_2(\psi^*)$, and the John constant of G.

‡More precisely, on p and the John constant of Ω.

Next, by using this result we can prove the continuous inf-sup condition.

LEMMA 4.3. *Let $G \subset \mathbb{R}^n$ be a bounded John domain and let ψ be an N-function with $\Delta_2(\psi), \Delta_2(\psi^*) < \infty$. Then, for all $q \in L_0^{\psi^*}(G)$ we have*

$$\|q\|_{L_0^{\psi^*}(G)} \le c \sup_{\|\boldsymbol{\xi}\|_{(W_0^{1,\psi}(G))^n} \le 1} \langle q, \operatorname{div} \boldsymbol{\xi} \rangle$$

and also

$$\int_G \psi^*(|q|)\, dx \le \sup_{\boldsymbol{\xi} \in (W_0^{1,\psi}(G))^n} \left[\int_G q \operatorname{div} \boldsymbol{\xi}\, dx - \frac{1}{c} \int_G \psi(|\nabla \boldsymbol{\xi}|)\, dx \right],$$

where the constants depend only on $\Delta_2(\psi)$, $\Delta_2(\psi^*)$, and the John constant of G.

Proof. The first assertion follows from the isomorphism between $(L_0^{\varphi}(G))^*$ and $L_0^{\varphi^*}(G)$ (with constant bounded by 2). In fact, it follows from Theorem 4.2

$$\|q\|_{L_0^{\psi^*}(G)} \le 2 \sup_{\|\eta\|_{L_0^{\psi}(G)} \le 1} \int_G q\, \eta\, dx \le 2 \sup_{\|\boldsymbol{\xi}\|_{(W_0^{1,\psi}(G))^n} \le c} \int_G q\, \operatorname{div} \boldsymbol{\xi}\, dx\,.$$

The properties of conjugate functions (see [29]) and $q \in L_0^{\psi^*}(G)$ imply

$$\int_G \psi^*(|q|)\, dx = \sup_{\eta \in L_0^{\psi}(G)} \left[\int_G q\, \eta\, dx - \int_G \psi(|\eta|)\, dx \right].$$

Next, Theorem 4.2 and the Δ_2-condition for ψ imply

$$\int_G \psi^*(|q|)\, dx \le \sup_{\boldsymbol{\xi} \in (W_0^{1,\psi}(G))^n} \left[\int_G q \operatorname{div} \boldsymbol{\xi}\, dx - \frac{1}{c} \int_G \psi(|\nabla \boldsymbol{\xi}|)\, dx \right].$$

This proves the second assertion. ∎

We are now ready for the proof of Lemma 4.1.

Proof. [Proof of Lemma 4.1.] We use Lemma 4.3, Assumption 2.20, and Theorem 3.5 to get

$$\begin{aligned} \|q_h\|_{Q_h} &\le c \sup_{\|\boldsymbol{\xi}\|_V \le 1} \langle q_h, \operatorname{div} \boldsymbol{\xi} \rangle = c \sup_{\|\boldsymbol{\xi}\|_V \le 1} \langle q_h, \operatorname{div} \Pi_h^{\operatorname{div}} \boldsymbol{\xi} \rangle \\ &\le c \sup_{\|\Pi_h^{\operatorname{div}} \boldsymbol{\xi}\|_{V_h} \le 1} \langle q_h, \operatorname{div} \Pi_h^{\operatorname{div}} \boldsymbol{\xi} \rangle \le c \sup_{\|\boldsymbol{\xi}_h\|_{V_h} \le 1} \langle q_h, \operatorname{div} \boldsymbol{\xi}_h \rangle \end{aligned}$$

and, by the same tools,

$$\begin{aligned} \int_\Omega \varphi^*(|q_h|)\, dx &\le \sup_{\boldsymbol{\xi} \in V} \left[\langle q_h, \operatorname{div} \boldsymbol{\xi} \rangle - \frac{1}{c} \int_\Omega \varphi(|\nabla \boldsymbol{\xi}|)\, dx \right] \\ &\le \sup_{\boldsymbol{\xi} \in V} \left[\langle q_h, \operatorname{div} \Pi_h^{\operatorname{div}} \boldsymbol{\xi} \rangle - \frac{1}{c} \int_\Omega \varphi(|\nabla \Pi_h^{\operatorname{div}} \boldsymbol{\xi}|)\, dx \right] \\ &\le \sup_{\boldsymbol{\xi}_h \in V_h} \left[\langle q_h, \operatorname{div} \boldsymbol{\xi}_h \rangle - \frac{1}{c} \int_\Omega \varphi(|\nabla \boldsymbol{\xi}_h|)\, dx \right]. \end{aligned}$$

∎

REMARK 4.4. Lemma 4.1 remains correct if φ is an arbitrary N-function such that φ and φ^* satisfy the Δ_2-condition. In this case the constants depend on $\Delta_2(\varphi)$ and $\Delta_2(\varphi^*)$, instead of on p.

4.2. Error estimate for the pressure. We now derive a best approximation result for the numerical error of the pressure. This result is valid for all $p \in (1, \infty)$. Later on, however, we have to distinguish between the cases $p \in (1, 2]$ and $p \in [2, \infty)$.

LEMMA 4.5. *Let Π_h^{div} satisfy Assumption 2.20. Let $(\mathbf{v}, q)$ and $(\mathbf{v}_h, q_h)$ be solutions of the problems* **(Q)** *and* **($\mathbf{Q}_h$)***, respectively. Then, we have the following estimate*

$$\int_\Omega \varphi^*(|q - q_h|)\, dx \le c \int_\Omega \varphi^*(|\mathbf{S}(\mathbf{D}\mathbf{v}) - \mathbf{S}(\mathbf{D}\mathbf{v}_h)|)\, dx + c \inf_{\mu_h \in Q_h} \int_\Omega \varphi^*(|q - \mu_h|)\, dx\,.$$

Proof. We split the error $q - q_h$ into a best approximation error $q - \mu_h$ and the remaining part $\mu_h - q_h$, which we will control by means of the equation for q_h. In particular, for all $\mu_h \in Q_h$ it holds

$$\int_\Omega \varphi^*(|q - q_h|)\, dx \le c \int_\Omega \varphi^*(|q - \mu_h|)\, dx + c \int_\Omega \varphi^*(|\mu_h - q_h|)\, dx,$$

by a triangle inequality, where we used $\Delta_2(\varphi^*) < \infty$. The second term is estimated with the help of Lemma 4.1 as follows

$$\int_\Omega \varphi^*(|\mu_h - q_h|)\, dx \le \sup_{\boldsymbol{\xi}_h \in V_h} \bigg[\langle \mu_h - q_h, \operatorname{div} \boldsymbol{\xi}_h \rangle - \frac{1}{c} \int_\Omega \varphi(|\nabla \boldsymbol{\xi}_h|)\, dx \bigg].$$

Let us take a closer look at the term $\langle \mu_h - q_h, \operatorname{div} \boldsymbol{\xi}_h \rangle$. By using the equation for the error (3.1), we get

$$\begin{aligned} \langle \mu_h - q_h, \operatorname{div} \boldsymbol{\xi}_h \rangle &= \langle \mu_h - q, \operatorname{div} \boldsymbol{\xi}_h \rangle + \langle q - q_h, \operatorname{div} \boldsymbol{\xi}_h \rangle \\ &= \langle \mu_h - q, \operatorname{div} \boldsymbol{\xi}_h \rangle + \langle \mathbf{S}(\mathbf{D}\mathbf{v}) - \mathbf{S}(\mathbf{D}\mathbf{v}_h), \mathbf{D}\boldsymbol{\xi}_h \rangle. \end{aligned}$$

Thus, with Young's inequality (2.2) and the previous estimate we obtain

$$\int_\Omega \varphi^*(|\mu_h - q_h|)\, dx \le c \int_\Omega \varphi^*(|q - \mu_h|)\, dx + c \int_\Omega \varphi^*(|\mathbf{S}(\mathbf{D}\mathbf{v}) - \mathbf{S}(\mathbf{D}\mathbf{v}_h)|)\, dx.$$

The claim follows, since $\mu_h \in Q_h$ was arbitrary. □

Unfortunately, the estimate for the error of the pressure $q - q_h$ involves the error of the stresses $\mathbf{S}(\mathbf{D}\mathbf{v}) - \mathbf{S}(\mathbf{D}\mathbf{v}_h)$. Our error estimates for the velocity in Theorem 3.9 are however expressed in terms of $\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\mathbf{D}\mathbf{v}_h)$. In the linear case $p = 2$, this is no source of problems, since then $\mathbf{S}(\mathbf{Q}) = \mathbf{F}(\mathbf{Q}) = \mathbf{Q}^{\text{sym}}$. In our non-linear setting $p \neq 2$ this causes a serious problem. To handle the difficulties, we have to distinguish the sub-quadratic case $p \in (1, 2]$ and the super-quadratic case $p \in [2, \infty)$.

To this end we need the following simple estimates for φ and φ^*:

$$\begin{aligned} \varphi(\lambda\, t) &\le c \max\{\lambda^p, \lambda^2\}\, \varphi(t), \\ \varphi^*(\lambda\, t) &\le c \max\{\lambda^{p'}, \lambda^2\}\, \varphi^*(t)\,, \end{aligned} \tag{4.6}$$

which are valid for all $\lambda, t \ge 0$. Observe that we can omit the terms with λ^2 if $\delta = 0$. However, this does not improve our estimates. The following lemma represents the missing link between the error in terms of $\mathbf{S}$ and the error in terms of $\mathbf{F}$.

LEMMA 4.7. *For all $p \in (1, 2]$ it holds*

$$\int_\Omega \varphi^*(|\mathbf{S}(\mathbf{D}\mathbf{v}) - \mathbf{S}(\mathbf{D}\mathbf{v}_h)|)\, dx \le c \int_\Omega |\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\mathbf{D}\mathbf{v}_h)|^2\, dx, \tag{4.8}$$

while for all $p \in (2,\infty)$ *it holds*

$$\begin{aligned}\int_\Omega \varphi^*(|\mathbf{S}(\mathbf{Dv}) - \mathbf{S}(\mathbf{Dv}_h)|)\,dx \le c \bigg(\int_\Omega |\mathbf{F}(\mathbf{Dv}) - \mathbf{F}(\mathbf{Dv}_h)|^2\,dx\bigg)^{\frac{p'}{2}} \cdot \\ \cdot \bigg(\int_\Omega \varphi(|\mathbf{Dv}| + |\mathbf{Dv}_h|)\,dx\bigg)^{\frac{2-p'}{2}}.\end{aligned} \tag{4.9}$$

Proof. Let us define the functions $G := |\mathbf{Dv}| + |\mathbf{Dv}_h|$, $\lambda := \frac{|\mathbf{Dv}-\mathbf{Dv}_h|}{|\mathbf{Dv}|+|\mathbf{Dv}_h|}$. Then by (2.15), the definition (2.12) of $\varphi_a'(t)$, the equivalence $|\mathbf{A}| + |\mathbf{A}-\mathbf{B}| \sim |\mathbf{A}| + |\mathbf{B}|$, (4.6), (2.1), and $\lambda \le 1$ it follows

$$\varphi^*(|\mathbf{S}(\mathbf{Dv}) - \mathbf{S}(\mathbf{Dv}_h)|) \sim \varphi^*(\varphi'(G)\,\lambda) \le c\,\varphi^*(\varphi'(G)) \max\{\lambda^2, \lambda^{p'}\} \sim \varphi(G)\,\lambda^{\min\{2,p'\}}.$$

From Lemma 2.13, (2.7), and (2.8) we also see that

$$|\mathbf{F}(\mathbf{Dv}) - \mathbf{F}(\mathbf{Dv}_h)|^2 \sim \varphi(|G|)\,\lambda^2.$$

So if $p \in (1,2]$, then $\varphi^*(|\mathbf{S}(\mathbf{Dv}) - \mathbf{S}(\mathbf{Dv}_h)|) \le c\,|\mathbf{F}(\mathbf{Dv}) - \mathbf{F}(\mathbf{Dv}_h)|^2$ and the claim follows.

If $p \in (2,\infty)$ Young's inequality implies $\lambda^{p'} \le \big(\frac{\lambda}{\gamma}\big)^2 + c\,\gamma^{\frac{2p'}{2-p'}}$ for all $\lambda > 0$ and $\gamma > 0$. Thus, we get

$$\varphi^*(|\mathbf{S}(\mathbf{Dv}) - \mathbf{S}(\mathbf{Dv}_h)|) \le c\,\gamma^{-2}|\mathbf{F}(\mathbf{Dv}) - \mathbf{F}(\mathbf{Dv}_h)|^2 + c\,\varphi(|G|)\gamma^{\frac{2p'}{2-p'}}.$$

Now the claim follows by integration over $x \in \Omega$ and minimizing with respect to γ. ∎

Combining Lemma 4.5 and Lemma 4.7 we get our desired error estimate for the pressure.

THEOREM 4.10. *Let* Π_h^{div} *satisfy Assumption 2.20. Let* $(\mathbf{v}, q)$ *and* $(\mathbf{v}_h, q_h)$ *be solutions of the problems* **(Q)** *and* **(Q$_h$)**, *respectively. Then we have*

$$\int_\Omega \varphi^*(|q - q_h|)\,dx \le c\,\|\mathbf{F}(\mathbf{Dv}) - \mathbf{F}(\mathbf{Dv}_h)\|_2^{\min\{p',2\}} + c \inf_{\mu_h \in Q_h} \int_\Omega \varphi^*(|q - \mu_h|)\,dx\,.$$

REMARK 4.11. By a slightly different argument it is possible to replace $\varphi^*(t)$ by $t^{p'} =: \omega^*(t)$ in the estimates of Lemma 4.5 and of Theorem 4.10. In such a way we avoid the implicit appearance of δ and get $\|q - q_h\|_{p'}^{p'}$ instead.

To achieve this we first note that the proof and the statement of Lemma 4.7 remain both valid with $\varphi^*(t)$ replaced by $\omega^*(t)$. We have to modify the proof of Lemma 4.7. For $p \le 2$ we use the estimate $\omega^*(t) \le \varphi^*(t)$ applied to $t = |\mathbf{S}(\mathbf{Dv}) - \mathbf{S}(\mathbf{Dv}_h)|$ in (4.8). In the case $p > 2$ one computes directly the estimate (4.9). Note that on the right-hand side of (4.9) φ is not changed to t^p.

5. Estimates using the regularity of ∇q and $\nabla \mathbf{F}$. In this section we prove estimates for the error of the velocity and of the pressure using additional assumptions on the regularity of the pressure and of the velocity. This allows to obtain convergence rates in terms of the mesh size h. In particular, we will assume $\mathbf{F}(\mathbf{Dv}) \in (W^{1,2}(\Omega))^{n\times n}$ and $q \in W^{1,\varphi^*}(\Omega)$. The assumption on $\mathbf{F}(\mathbf{Dv})$ corresponds naturally to the monotonicity of $-\operatorname{div}(\mathbf{S}(\mathbf{Dv}))$ expressed by Lemma 2.13 and it represents the non-linear extension of $\mathbf{v} \in W^{2,2}(\Omega)$ for the linear case $p = 2$. The assumption on q however is the natural extension of $q \in L^{\varphi*}(\Omega)$ for weak solutions, by assuming control over one more derivative. In the linear case this corresponds to $q \in W^{1,2}(\Omega)$.

Unfortunately, there is a certain mismatch between the assumptions $\mathbf{F}(\mathbf{D}\mathbf{v}) \in (W^{1,2}(\Omega))^{n\times n}$ and $q \in W^{1,\varphi^*}(\Omega)$ since the latter condition corresponds roughly to the condition $\mathbf{S}(\mathbf{D}\mathbf{v}) \in (W^{1,\varphi^*}(\Omega))^{n\times n}$. Observe that the condition $\mathbf{F}(\mathbf{D}\mathbf{v}) \in (W^{1,2}(\Omega))^{n\times n}$ is stronger than $\mathbf{S}(\mathbf{D}\mathbf{v}) \in (W^{1,\varphi^*}(\Omega))^{n\times n}$ for $p \geq 2$, and a weaker condition for $p \leq 2$. This mismatch is already indicated in the proof of Lemma 4.7. This problem is not easy to solve even in the case of the p-Laplacian and is related to many open questions and conjectures concerning the regularity of solutions of the p-Laplacian.

It remains to control the terms at the right-hand side in Theorem 3.9 and Theorem 4.10 in terms of the mesh size and the assumed regularity. We begin with the terms involving the velocity.

THEOREM 5.1. *Let* $\mathbf{F}(\mathbf{D}\mathbf{v}) \in (W^{1,2}(\Omega))^{n\times n}$*, then*

$$\sum_{K\in\mathcal{T}_h} \int_{S_K} \big|\mathbf{F}(\mathbf{D}\mathbf{v}) - \langle \mathbf{F}(\mathbf{D}\mathbf{v})\rangle_{S_K}\big|^2 \, dx \leq c\, h^2 \,\big\|\nabla \mathbf{F}(\mathbf{D}\mathbf{v})\big\|_2^2. \tag{5.2}$$

Proof. The claim follows immediately by using Poincaré's inequality applied to $\mathbf{F}(\mathbf{D}\mathbf{v})$ in $(L^2(S_K))^{n\times n}$. □

We now turn to the more complicated terms involving the pressure. We have the following local continuity and approximability property of Π_h^Y.

LEMMA 5.3. *Let* Π_h^Y *satisfy Assumption 2.24. Let* ψ *be an N-function with* $\Delta_2(\psi) < \infty$*. Then for all* $K \in \mathcal{T}_h$ *and* $q \in L^{\psi}(\Omega)$ *we have*

$$\int_K \psi\big(|\Pi_h^Y q|\big)\, dx \leq c \int_{S_K} \psi\big(|q|\big)\, dx.$$

Moreover, for all $K \in \mathcal{T}_h$ *and* $q \in W^{1,\psi}(\Omega)$ *we have*

$$\int_K \psi\big(|q - \Pi_h^Y q|\big)\, dx \leq c \int_{S_K} \psi\big(h_K |\nabla q|\big)\, dx.$$

The constants depend only on $\Delta_2(\psi)$ *and on* γ_0*.*

Proof. Due to Assumption 2.24 the operator Π_h^Y satisfies assumption 4.1 of [16] both for $r_0 = l_0 = l = 0$ and $r_0 = l_0 = 0$, $l = 1$. The first choice and [16, Corollary 4.8] imply the first assertion, while the second one and [16, Theorem 4.6] yield the second assertion. □

As a direct consequence of this lemma, we can estimate the term involving $q - \mu_h$ in Theorem 4.10 by choosing $\mu_h = \Pi_h^Y q$.

LEMMA 5.4. *Let* Π_h^Y *satisfy Assumption 2.24. Let* $(\mathbf{v}, q)$ *be a solution of the problem* **(Q)***. Then*

$$\inf_{\mu_h \in Q_h} \int_\Omega \varphi^*(|q - \mu_h|)\, dx \leq c \int_\Omega \varphi^*(h\,|\nabla q|)\, dx\,.$$

However, the expression with $q - \mu_h$ in Theorem 3.9 requires a little bit more of work. We have the following result.

Lemma 5.5. *Let Let $\Pi_h^{\rm div}$ satisfy Assumption 2.20 and let Π_h^Y satisfy Assumption 2.24. Let $(\mathbf{v}, q)$ and $(\mathbf{v}_h, q_h)$ be solutions of the problems* **(Q)** *and* **($\mathbf{Q}_h$)**, *respectively. Then*

$$\int_K (\varphi_{|\mathbf{Dv}|})^*(|q - \Pi_h^Y q|)\,dx \le c\int_{S_K} (\varphi_{|\mathbf{Dv}|})^*(h\,|\nabla q|)\,dx + c\int_{S_K} \big|\mathbf{F}(\mathbf{Dv}) - \langle \mathbf{F}(\mathbf{Dv})\rangle_{S_K}\big|^2\,dx.$$

Proof. The claim follows by first using a shift-change from $|\mathbf{Dv}|$ to $|\langle \mathbf{Dv}\rangle_{S_K}|$ (see Lemma A.3), second applying Lemma 5.3 with $\psi = (\varphi_{|\langle \mathbf{Dv}\rangle_{S_K}|})^*$, third reversing the shift back to $|\mathbf{Dv}|$, and then fourth using Lemma A.4 to pass from $\mathbf{F}(\langle \mathbf{Dv}\rangle_{S_K})$ to $\langle \mathbf{F}(\mathbf{Dv})\rangle_{S_K}$. □

We are now ready to prove our error estimates for the velocity and the pressure in terms of the mesh size h.

Proof. [Proof of Theorem 2.28] It follows from Theorem 3.9, Lemma 5.5, and Theorem 5.1 that

$$\|\mathbf{F}(\mathbf{Dv}) - \mathbf{F}(\mathbf{Dv}_h)\|_2^2 \le c\,h^2 + c\int_\Omega (\varphi_{|\mathbf{Dv}|})^*(h\,|\nabla q|)\,dx.$$

Now estimate (2.29) follows from the elementary inequalities

$$(\varphi_{|\mathbf{Q}|})^*(h\,t) \le c\,h^{\min\{2,p'\}}(\varphi_{|\mathbf{Q}|})^*(t),$$
$$(\varphi_{|\mathbf{Q}|})^*(t) \le c\,\varphi^*(t) + c\,\varphi(|\mathbf{Q}|),$$

where the first estimate follows from the definition of the shifted N-functions, while the second estimate is a consequence of Lemma A.3 with $\mathbf{P} = \mathbf{0}$ and $|\mathbf{F}(\mathbf{Q})|^2 \sim \varphi(|\mathbf{Q}|)$.

From Theorem 4.10, Lemma 5.4, (4.6) and estimate (2.29) it follows

$$\int_\Omega \varphi^*(|q - q_h|)\,dx \le c\,\big(h^{\min\{2,p'\}}\big)^{\frac{\min\{2,p'\}}{2}} + \int_\Omega \varphi^*(h|\nabla q|)\,dx \le c\,h^{\min\left\{2,\frac{(p')^2}{2}\right\}} + c\,h^{\min\{2,p'\}}.$$

This proves estimate (2.31). As in Remark 4.11 it is possible to replace $\varphi^*(t)$ in (2.31) by $t^{p'}$, which proves estimate (2.30). □

6. Error Estimates in Terms of f. Sometimes it is of interest to derive estimates for the error purely in terms of the data $\mathbf{f}$. In the linear case this is done by combining the error estimates in terms of the regularity of the solution with the regularity results of the solutions in terms of the data. So for example if $p = 2$ and the domain is convex, then $\mathbf{f} \in (L^2(\Omega))^n$ and $g \in W^{1,2}(\Omega) \cap L^2_0(\Omega)$ imply $\mathbf{v} \in (W^{2,2}(\Omega))^n$, $q \in W^{1,2}(\Omega)$ and therefore $\|\nabla \mathbf{v} - \nabla \mathbf{v}_h\|_2 + \|q - q_h\|_2 \le c\,h$. However, such optimal regularity results are not available in the non-linear context. The best results in this direction can be shown in the space periodic setting. Using ideas from [30, 7] it is possible to show the following:

Lemma 6.1. *Consider $(1.1)_{1,2}$ with $g = 0$ in the space periodic setting with $\Omega = [-L, L]^n$, $n \ge 3$. Assume that $\mathbf{S}$ satisfies Assumption 2.3 with $\delta > 0$. Then we have:*

(i) *For $p \ge 2$ and $\mathbf{f} \in (L^2_0(\Omega))^n$ there exists a solution $(\mathbf{v}, q)$ with $\mathbf{F}(\mathbf{Dv}) \in (W^{1,2}(\Omega))^{n\times n}$ and $q \in W^{1,p'}(\Omega) \cap L^{p'}_0(\Omega)$.*

(ii) *For* $p \le 2$ *and* $\mathbf{f} \in (L_0^{p'}(\Omega))^n$ *there exists a solution* $(\mathbf{v}, q)$ *with* $\mathbf{F}(\mathbf{D}\mathbf{v}) \in (W^{1,2}(\Omega))^{n\times n}$ *and* $q \in W^{1,\frac{np}{p+n-2}}(\Omega) \cap L_0^{p'}(\Omega)$.

REMARK 6.2. If we consider the system (1.1) with Dirichlet boundary conditions, $g = 0$ and $\mathbf{S}$ satisfying Assumption 2.3 with $\delta > 0$, the best results at the moment are contained in [5]. It is shown there that for a sufficiently smooth $\partial\Omega$ and $p \ge 2$ one gets $\mathbf{F}(\mathbf{D}\mathbf{v}) \in (W^{1,\frac{2r}{p+r-2}}(\Omega))^{n\times n}$ with $r = (np + 2 - p)/(n-2)$, if $n \ge 3$, and arbitrary $r < \infty$, if $n = 2$.

The previous lemma indicates that the regularity $\mathbf{F}(\mathbf{D}\mathbf{v}) \in (W^{1,2}(\Omega))^{n\times n}$ and $q \in W^{1,\varphi^*}(\Omega)$ seem at least natural in the non-linear setting. However, the term $(\varphi_{|\mathbf{D}\mathbf{v}|})^*(|q - \mu_h|)$ appearing in Theorem 3.9 can not be treated directly using $q \in W^{1,p'}(\Omega)$. Thus we have to proceed differently.

We will also need a Orlicz-Sobolev version of the Poincaré's inequality for functions vanishing at the boundary

LEMMA 6.3 (Poincaré inequality). *Let* $\Omega \subset \mathbb{R}^n$ *be open and bounded. Let* ψ *be an N-function with* $\Delta_2(\psi), \Delta_2(\psi^*) < \infty$. *Then, there exists* $c > 0$ *only depending on* $\Delta_2(\psi)$ *and* $\Delta_2(\psi^*)$ *such that*

$$\int_\Omega \psi\bigg(\frac{|u|}{\operatorname{diam}(\Omega)}\bigg)\,dx \le c\int_\Omega \psi(|\nabla u|)\,dx \qquad \forall\, u \in W_0^{1,\psi}(\Omega). \tag{6.4}$$

Proof. If follows from [28] that for $u \in W_0^{1,\psi}(\Omega) \subset W_0^{1,1}(\Omega)$ holds

$$|u(x)| \le c\,\operatorname{diam}(\Omega)\,\mathcal{M}(|\nabla u|)(x),$$

where $\mathcal{M}(\,\cdot\,)(x)$ is the Hardy-Littlewood maximal function. Since $\Delta_2(\psi) < \infty$ and $\Delta_2(\psi^*) < \infty$, if follows from [21] that $\mathcal{M}$ satisfies a modular estimate on $L^\psi(\Omega)$, so

$$\int_\Omega \psi(|\mathcal{M}(|\nabla u|)|)\,dx \le c\int_\Omega \psi(|\nabla u|)\,dx,$$

which finishes the proof. ◻

LEMMA 6.5. *Let* Π_h^Y *satisfy Assumption 2.24. Let* $(\mathbf{v}, q)$ *be solutions of the problem* **(Q)**. *Then, for all* $K \in \mathcal{T}_h$, *it holds*

$$\int_K (\varphi_{|\mathbf{D}\mathbf{v}|})^*(|q - \Pi_h^Y q|)\,dx \le c\int_K \big(\varphi_{|\mathbf{D}\mathbf{v}|}\big)^*(h|\mathbf{f}|)\,dx + c\int_{S_K} |\mathbf{F}(\mathbf{D}\mathbf{v}) - \langle\mathbf{F}(\mathbf{D}\mathbf{v})\rangle_{S_K}|^2\,dx.$$

Proof. We use the identity $q - \Pi_h^Y q = (q - \langle q\rangle_{S_K}) - \Pi_h^Y(q - \langle q\rangle_{S_K})$, the triangle inequality together with $\Delta_2(\varphi^*) < \infty$, and the local stability of Π_h^Y from Lemma 5.3 to conclude that

$$\begin{aligned}\int_K \big(\varphi_{|\mathbf{D}\mathbf{v}|}\big)^*(|q - \Pi_h^Y q|)\,dx &\le c\int_K \big(\varphi_{|\mathbf{D}\mathbf{v}|}\big)^*(|q - \langle q\rangle_{S_K}|)\,dx \\ &\quad + c\int_K \big(\varphi_{|\mathbf{D}\mathbf{v}|}\big)^*(|\Pi_h^Y(q - \langle q\rangle_{S_K})|)\,dx \\ &\le c\int_{S_K} \big(\varphi_{|\mathbf{D}\mathbf{v}|}\big)^*(|q - \langle q\rangle_{S_K}|)\,dx.\end{aligned}$$

With the shift-change Lemma A.3 we further get

$$\begin{aligned}\int_{S_K} \big(\varphi_{|\mathbf{D}\mathbf{v}|}\big)^*(|q - \langle q\rangle_{S_K}|)\,dx &\le \int_{S_K} \big(\varphi_{|\langle\mathbf{D}\mathbf{v}\rangle_{S_K}|}\big)^*(|q - \langle q\rangle_{S_K}|)\,dx \\ &\quad + \int_{S_K} |\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\langle\mathbf{D}\mathbf{v}\rangle_{S_K})|^2\,dx.\end{aligned}$$

From the inf-sup condition in modular form of Lemma 4.3 for the domain S_K, it follows that

$$\begin{aligned}\int_{S_K} \big(\varphi_{|\langle \mathbf{D}\mathbf{v}\rangle_{S_K}|}\big)^*(|q-\langle q\rangle_{S_K}|)\,dx \\ \le \sup_{\boldsymbol{\xi}\in C_0^\infty(S_K)} \bigg[\int_{S_K} q \operatorname{div}\boldsymbol{\xi}\,dx - \frac{1}{c}\int_{S_K} \varphi_{|\langle \mathbf{D}\mathbf{v}\rangle_{S_K}|}(|\nabla\boldsymbol{\xi}|)\,dx\bigg].\end{aligned} \tag{6.6}$$

Next, by using the weak formulation of problem **(Q)** we obtain

$$\int_{S_K} q \operatorname{div}\boldsymbol{\xi}_K\,dx = \int_{S_K} \mathbf{S}(\mathbf{D}\mathbf{v})\cdot\mathbf{D}\boldsymbol{\xi}_K - \mathbf{f}\cdot\boldsymbol{\xi}_K\,dx. \tag{6.7}$$

By using (2.15), Young's inequality (2.2), and Lemma 2.13 we get

$$\begin{aligned}\bigg|\int_{S_K} \mathbf{S}(\mathbf{D}\mathbf{v})\cdot\mathbf{D}\boldsymbol{\xi}_K\,dx\bigg| &= \bigg|\int_{S_K} \big(\mathbf{S}(\mathbf{D}\mathbf{v})-\mathbf{S}(\langle\mathbf{D}\mathbf{v}\rangle_{S_K})\big)\cdot\mathbf{D}\boldsymbol{\xi}_K\,dx\bigg| \\ &\le c_\varepsilon \int_{S_K} \varphi_{|\langle \mathbf{D}\mathbf{v}\rangle_{S_K}|}(|\mathbf{D}\mathbf{v}-\langle\mathbf{D}\mathbf{v}\rangle_{S_K}|)\,dx + \varepsilon\int_{S_K}\varphi_{|\langle \mathbf{D}\mathbf{v}\rangle_{S_K}|}(|\nabla\boldsymbol{\xi}_K|)\,dx \\ &\le c_\varepsilon \int_{S_K} \big|\mathbf{F}(\mathbf{D}\mathbf{v})-\mathbf{F}(\langle\mathbf{D}\mathbf{v}\rangle_{S_K})\big|^2\,dx + \varepsilon\int_{S_K}\varphi_{|\langle \mathbf{D}\mathbf{v}\rangle_{S_K}|}(|\nabla\boldsymbol{\xi}_K|)\,dx\,.\end{aligned}$$

Moreover, with Young's inequality (2.2), and Poincaré's inequality (6.4) we obtain

$$\begin{aligned}\bigg|\int_{S_K}\mathbf{f}\cdot\boldsymbol{\xi}_K\,dx\bigg| &\le c_\varepsilon\int_{S_K}\big(\varphi_{|\langle \mathbf{D}\mathbf{v}\rangle_{S_K}|}\big)^*(h\,|\mathbf{f}|)\,dx + \varepsilon\int_{S_K}\varphi_{|\langle \mathbf{D}\mathbf{v}\rangle_{S_K}|}\bigg(\bigg|\frac{\boldsymbol{\xi}_K}{h}\bigg|\bigg)\,dx \\ &\le c_\varepsilon\int_{S_K}\big(\varphi_{|\langle \mathbf{D}\mathbf{v}\rangle_{S_K}|}\big)^*(h\,|\mathbf{f}|)\,dx + \varepsilon\,c\int_{S_K}\varphi_{|\langle \mathbf{D}\mathbf{v}\rangle_{S_K}|}(|\nabla\boldsymbol{\xi}_K|)\,dx.\end{aligned}$$

By another shift-change from $|\langle\mathbf{D}\mathbf{v}\rangle_{S_K}|$ to $|\mathbf{D}\mathbf{v}|$ and by collecting all estimates after (6.6) we get

$$\begin{aligned}\int_{S_K} \big(\varphi_{|\langle \mathbf{D}\mathbf{v}\rangle_{S_K}|}\big)^*(|q-\langle q\rangle_{S_K}|)\,dx \\ \le c\int_{S_K}\big(\varphi_{|\mathbf{D}\mathbf{v}|}\big)^*(h|\mathbf{f}|)\,dx + c\int_{S_K}|\mathbf{F}(\mathbf{D}\mathbf{v})-\mathbf{F}(\langle\mathbf{D}\mathbf{v}\rangle_{S_K})|^2\,dx,\end{aligned}$$

where we chose $\varepsilon>0$ so small that all terms involving ε could be absorbed into the negative term on the right-hand side of (6.6). This together with Lemma A.4 proves the claim. □

Proof. [Proof of Theorem 2.32] The proof of the statement is exactly the same as the proof of (2.29) if we use Lemma 6.5 instead of Lemma 5.5. □

REMARK 6.8. (i) For $p\le 2$ we can also obtain an error estimate for the pressure. Indeed, Theorem 4.10 and the elementary estimate $\varphi^*(t)\le c\,(\varphi_a)^*(t)$ imply

$$\int_\Omega \varphi^*(|q-q_h|)\,dx \le c\,h^2$$

in the same way as in the proof of (2.31) if we use again Lemma 6.5 instead of Lemma 5.5. Again one can replace $\varphi^*(t)$ in (2.31) by $t^{p'}$, which proves $\|q-q_h\|_{p'}\le c\,h^{\frac{2}{p'}}$.

(ii) For $p \geq 2$ we can improve (2.33) to

$$\|\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\mathbf{D}\mathbf{v}_h)\|_2 \leq c\, \delta^{\frac{2-p}{2}}\, h$$

if $\mathbf{f} \in (L^2(\Omega))^n$ and $\delta > 0$. This can be achieved by using $(\varphi_a)^*(t) \sim \big(t + (a + \delta)^{p-1}\big)^{p'-2} t^2$ when estimating $\int_\Omega (\varphi_{|\mathbf{D}\mathbf{v}|})^*(h|\mathbf{f}|)\, dx$.

7. Numerical Experiments. In this section we report on some numerical experiments with a fully practical approximation of (1.1). The computations have been performed with an extension of the finite element toolbox ALBERTA. This extension was developed in framework of the DFG research unit "Non-Linear partial differential equations: Theoretical and numerical analysis". For simplicity we took Ω to be the square $[-1, 1] \times [-1, 1]$. The computational domain was uniformly refined by first inserting a SW–NE diagonal and then using the newest vertex bisection method. For the solution of the saddle point problem (1.1) we used a preconditioned CG method for the non-linear Schur complement operator $S := B^* A^{-1} B$, where $B = \nabla$, $B^* = \operatorname{div}$ and $A = -\operatorname{div} \mathbf{S}$. The preconditioner was especially designed to handle the non-linear elliptic operator in dependence on p and δ (cf. Assumption 2.3). In the case $\mathbf{S} = \mu(\delta + |\mathbf{D}\mathbf{v}|)^{p-2}\mathbf{D}\mathbf{v}$ the approximation to the inverse of the Schur complement operator reads

$$S^{-1}\pi \approx \mu(\delta + n^{-1/2}|\pi|)^{p-2}\pi\,,$$

where n is the dimension of the computational domain. Note that for $p = 2$ the preconditioner reduces to the one used in the linear case. The resulting non-linear elliptic problem was solved using Newton's method with step-size control. It uses standard strategies aiming at the reduction of the residuum. In fact we combine the residual monotonicity test and the natural monotonicity test in our stopping criterion (cf. [10]). The resulting linearized equations are solved by a SSOR preconditioned CG method. This code showed reliable results for a wide range of parameters p and δ. In our test problem we have chosen $\delta = 10^{-4}$, while p was varying between 1.25 and 3.0. Since we are mostly interested in the sub-linear case we included more experiments for $p \leq 2$.

We considered solutions with a point singularity at the origin both in the velocity and in the pressure. More precisely, the exact solution was given by[§]

$$\mathbf{v}(\mathbf{x}) = |\mathbf{x}|^{\alpha-1}(x_2, -x_1)^\top\,, \qquad q(\mathbf{x}) = |\mathbf{x}|^\gamma\,, \tag{7.1}$$

where α and γ have been chosen such that just $\mathbf{F}(\mathbf{D}\mathbf{v}) \in (W^{1,2}(\Omega))^{2\times 2}$ and $q \in W^{1,p'}(\Omega)$. This requirement is ensured for $\alpha > 1$ and $\gamma > -1 + \frac{2}{p}$. In our experiments we have chosen $\alpha = 1.01$ and $\gamma = \frac{2}{p} - 1 + 0.01$ in order to be very close to the critical regularity. This regularity corresponds to $\mathbf{v} \in (W^{2,2}(\Omega))^2$ and $q \in W^{1,2}(\Omega)$ in the linear case $p = 2$.

Due to the regularity $\mathbf{F}(\mathbf{D}\mathbf{v}) \in (W^{1,2}(\Omega))^{2\times 2}$ and $q \in W^{1,p'}(\Omega)$ we expect the errors $\|\mathbf{F}(\mathbf{D}\mathbf{v}) - \mathbf{F}(\mathbf{D}\mathbf{v}_h)\|_2$ and $\|q - q_h\|_{p'}$ to converge at most with order 1.0. In our experiments we used the MINI element (P_1 plus bubble - P_1). In particular, the

[§]The exact solution is not zero on the boundary of the computational domain. However, the error is clearly concentrated around the singularity and thus this small inconsistency with the setup of the theory does not have any influence on the results. Of course the problem can be easily modified by a suitable cut-off function to ensure zero boundary conditions.

velocity element can only resolve $\mathbf{F}(\mathbf{Dv})$ locally by constants, which fits naturally with the regularity $\mathbf{F}(\mathbf{Dv}) \in (W^{1,2}(\Omega))^{2\times 2}$. However, due the nonlinearity $-\operatorname{div}(\mathbf{S}(\mathbf{Dv}))$ it is not clear if the finite element solution will achieve this linear convergence. Indeed, in our main result (Theorem 2.28) we were only able to show linear convergence of $\|\mathbf{F}(\mathbf{Dv}) - \mathbf{F}(\mathbf{Dv}_h)\|_2$ for $p \le 2$ and we have sublinear convergence for $p \ge 2$. For the pressure we even have sublinear convergence for all p. It is the aim of the experiments to see if the rates in Theorem 2.28 are optimal. In the Tables 7.1 and 7.2 we present the experimental order of convergence (EOC) for $\|\mathbf{F}(\mathbf{Dv}) - \mathbf{F}(\mathbf{Dv}_h)\|_2$ and $\|q - q_h\|_{p'}$, respectively. The last line of the tables shows the order of convergence as expected from Theorem 2.28.

	p							
h	1.25	1.33	1.5	1.67	1.8	2.0	2.5	3.0
1.77e-01	0.86	0.90	0.85	0.91	0.86	0.85	0.26	0.66
8.84e-02	0.88	0.90	0.87	0.91	0.88	0.88	0.84	0.74
4.42e-02	0.89	0.91	0.89	0.92	0.90	0.90	0.85	0.75
2.21e-02	0.91	0.92	0.91	0.93	0.91	0.91	0.86	0.76
1.10e-02	0.92	0.93	0.92	0.94	0.92	0.92	0.85	0.76
5.52e-03	0.93	0.94	0.93	0.94	0.93	0.93	0.85	0.76
$\min\{1, \frac{p'}{2}\}$	1.00	1.00	1.00	1.00	1.00	1.00	0.83	0.75

TABLE 7.1
EOC of $\|\mathbf{F}(\mathbf{Dv}) - \mathbf{F}(\mathbf{Dv}_h)\|_2$

	p							
h	1.25	1.33	1.5	1.67	1.8	2.0	2.5	3.0
1.77e-01	1.02	1.05	1.06	1.01	0.98	1.24	0.90	0.91
8.84e-02	0.58	0.68	0.86	0.96	0.99	1.07	0.99	1.00
4.42e-02	0.49	0.61	0.82	0.95	0.99	1.05	1.01	1.02
2.21e-02	0.45	0.57	0.78	0.94	0.99	1.03	1.01	1.02
1.10e-02	0.42	0.54	0.75	0.93	0.99	1.02	1.01	1.02
5.52e-03	0.41	0.52	0.73	0.92	0.99	1.02	1.01	1.01
$\min\{\frac{2}{p'}, \frac{p'}{2}\}$	0.40	0.50	0.67	0.80	0.89	1.00	0.83	0.75

TABLE 7.2
EOC of $\|q - q_h\|_{p'}$

As can be seen from the Tables 7.1 and 7.2, the EOC for $\|\mathbf{F}(\mathbf{Dv}) - \mathbf{F}(\mathbf{Dv}_h)\|_2$ agrees for all p with the one predicted in Theorem 2.28. Therefore, our convergence result for the velocity is optimal. However, the EOC for $\|q - q_h\|_{p'}$ agrees only for $p \le 2$ with the one predicted in Theorem 2.28. So our convergence result for the pressure is optimal only for $p \le 2$. In the case $p \ge 2$, we observe linear convergence of $\|q - q_h\|_{p'}$ in our experiments.

Let us discuss the experimentally observed convergence rate for the pressure for $p \ge 2$. It is not clear if the predicted convergence rate is suboptimal or if the example (7.1) is not the best choice. The reason for the theoretical convergence rate $\frac{p'}{2}$ for the pressure is Lemma 4.5, where the error of the pressure is directly estimated by the error of the stress $\mathbf{S}$. However, we have no estimate for the numerical error in $\mathbf{S}$. When we estimate the numerical error of $\mathbf{S}$ by the one for $\mathbf{F}$, as done in

Lemma 4.7, we get the reduced rate $\frac{p'}{2}$. Though, in our special example (7.1), we have $\mathbf{S}(\mathbf{Dv}) \in (W^{1,p'}(\Omega))^{2\times 2}$ for all $p \geq 2$. In fact, this is the reason why the error of $\mathbf{S}$ might converges linearly, which is experimentally confirmed in Table 7.3 for $p \geq 2$. So by Lemma 4.5 this convergence rate transfers to the convergence rate of the pressure q. This explains the linear convergence of the pressure for $p \geq 2$. The last line in Table 7.3 shows the maximal possible convergence rate based on the regularity of $\mathbf{S}$ for our example (7.1), which is limited by 1 due to the ansatz functions of the velocity.

Based on this observation we suggest to investigate error estimates for the velocity in terms of $\mathbf{S}$ or to derive estimates which directly bound the error of the pressure in terms of $\mathbf{F}$.

	p							
h	1.25	1.33	1.5	1.67	1.8	2.0	2.5	3.0
1.77e-01	0.40	0.50	0.63	0.77	0.79	0.85	0.36	0.88
8.84e-02	0.40	0.50	0.65	0.78	0.81	0.88	1.02	0.98
4.42e-02	0.40	0.50	0.66	0.78	0.83	0.90	1.03	1.01
2.21e-02	0.40	0.50	0.66	0.79	0.84	0.91	1.04	1.01
1.10e-02	0.40	0.50	0.67	0.80	0.85	0.92	1.03	1.02
5.52e-03	0.40	0.50	0.67	0.80	0.86	0.93	1.03	1.02
$\min\{1, \frac{2}{p'}\}$	0.40	0.50	0.67	0.80	0.89	1.00	1.00	1.00

TABLE 7.3
EOC of $\|\mathbf{S}(\mathbf{Dv}) - \mathbf{S}(\mathbf{Dv}_h)\|_{p'}$

Appendix A.

A.1. Construction of divergence-preserving interpolation operator. In the case of the MINI element con can easily construct a divergence preserving interpolation operator Π_h^{div} satisfying Assumption 2.20. We follow [18]¶. Let b_K, $K \in \mathcal{T}_h$, be the bubble function in the simplex K and let Π_h^1 be the Scott–Zhang interpolation operator preserving boundary values (cf. [33]). Then Π_h^{div} can be constructed explicitly for $\mathbf{w} \in X$ by

$$\Pi_h^{\text{div}}\mathbf{w} := \Pi_h^1\mathbf{w} - \sum_{K\in\mathcal{T}_h} \mathbf{c}_K b_K \,,$$

where $\mathbf{c}_K$ is chosen such that $\mathbf{c}_K \fint_K b_K\,dx = \fint_K \Pi_h^1\mathbf{w} - \mathbf{w}\,dx$. Note that the choice for $\mathbf{c}_K$ and $\Pi_h^1\mathbf{w} - \mathbf{w} \in V$ ensure condition (a) in Assumption 2.20. Condition (b) in Assumption 2.20 is also satisfied since $b_K \in W_0^{1,p}(K)$, $K \in \mathcal{T}_h$, and $\Pi_h^1\mathbf{w} \in V$ for $\mathbf{w} \in V$. The stability and approximation properties of the Scott–Zhang interpolation operator (cf. [33]) and the choice for $\mathbf{c}_K$ yield the local $W^{1,1}$-stability of Π_h^{div}, since

$$\begin{aligned}
&\fint_K |\Pi_h^{\text{div}}\mathbf{w}|\,dx \leq \fint_K |\Pi_h^1\mathbf{w}|\,dx + |\mathbf{c}_K| \fint_K |b_K|\,dx \\
&\leq c\fint_{S_K} |\mathbf{w}|\,dx + c\fint_{S_K} h_K|\nabla\mathbf{w}|\,dx + \fint_K |\Pi_h^1\mathbf{w} - \mathbf{w}|\,dx \leq c\fint_{S_K} |\mathbf{w}|\,dx + c\fint_{S_K} h_K|\nabla\mathbf{w}|\,dx\,.
\end{aligned}$$

¶Note, that one can also adapt the approach in [9, Sec. VI.4] to obtain the same results in two dimensions (cf. [6]).

Thus we obtain the Orlicz-continuity and Orlicz-approximability by Theorem 3.5. Moreover, we also get by standard arguments

$$\|\Pi_h^{\mathrm{div}}\mathbf{w}\|_{W^{k,q}(K)} \le c\, h^{s-k+n(\frac{1}{q}-\frac{1}{r})}\|\nabla^s\mathbf{w}\|_{L^r(S_K)}\,,$$

where $k = 0, 1,\ s = 1, 2$, and $q, r \in [1, \infty]$.

A.2. Shifted N-functions. In this section we present a few results on N-functions which we used in the previous sections. The results presented here are valid for the choice of φ and $\mathbf{F}$ as in Section 2.2. However, the results hold also in a more general setting. In particular, we make the following assumption, which includes the explicit one that we use in this paper.

Assumption A.1. *Let ψ be an N-function such that ψ is C^1 on $[0,\infty)$ and C^2 on $(0,\infty)$. Further assume that*

$$\psi'(t) \sim t\,\psi''(t) \tag{A.2}$$

uniformly in $t > 0$. The constants in (A.2) *are called the* characteristics of ψ.

We remark that under these assumptions ψ and ψ^* automatically satisfy the Δ_2-condition, where the Δ_2-constants depend only on the characteristics of ψ (cf. [14]). We further assume in the following that $\mathbf{S}$ and $\mathbf{F}$ are given by

$$\begin{aligned}\mathbf{S}(\mathbf{Q}) &= \mathbf{S}_\psi(\mathbf{Q}) = \psi'(|\mathbf{Q}^{\mathrm{sym}}|)\frac{\mathbf{Q}^{\mathrm{sym}}}{|\mathbf{Q}^{\mathrm{sym}}|},\\ \mathbf{F}(\mathbf{Q}) &= \mathbf{F}_\psi(\mathbf{Q}) = \sqrt{\frac{\psi'(|\mathbf{Q}^{\mathrm{sym}}|)}{|\mathbf{Q}^{\mathrm{sym}}|}}\,\mathbf{Q}^{\mathrm{sym}}\end{aligned}$$

for $\mathbf{Q} \in \mathbb{R}^{n\times n}$. For $\psi = \varphi$ this corresponds to $\mathbf{S}$ and $\mathbf{F}$ in Remark 2.9 and (2.11), respectively. It is shown [12, 31] that under this conditions ψ, $\mathbf{S}$, and $\mathbf{F}$ satisfy Lemma 2.13. The following auxiliary result can be found in [13, 31].

Lemma A.3 (Change of Shift). *For each $\delta > 0$ there exists $C_\delta \ge 1$ (depending only on δ and the characteristics of ψ) such that*

$$\begin{aligned}\psi_{|\mathbf{Q}|}(t) &\le C_\delta\,\psi_{|\mathbf{P}|}(t) + \delta\,|\mathbf{F}(\mathbf{Q}) - \mathbf{F}(\mathbf{P})|^2,\\ (\psi_{|\mathbf{Q}|})^*(t) &\le C_\delta\,(\psi_{|\mathbf{P}|})^*(t) + \delta\,|\mathbf{F}(\mathbf{Q}) - \mathbf{F}(\mathbf{P})|^2\end{aligned}$$

for all $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^{n\times n}_{\mathrm{sym}}$ and $t \ge 0$.

The following shows that we can take different mean values when considering the mean oscillation of $\mathbf{F}$.

Lemma A.4. *Let Ω be a bounded, open set. Then it follows for all functions $\mathbf{H} \in (L^\psi(\Omega))^{n\times n}$ that*

$$\int_\Omega |\mathbf{F}(\mathbf{H}) - \langle\mathbf{F}(\mathbf{H})\rangle_\Omega|^2\,dx \sim \int_\Omega |\mathbf{F}(\mathbf{H}) - \mathbf{F}(\langle\mathbf{H}\rangle_\Omega)|^2\,dx\,,$$

where the constants depend only on the characteristics of ψ.

Proof. From $\mathbf{H} \in (L^\psi(\Omega)))^{n\times n}$ and Lemma 2.13 follows $\mathbf{F}(\mathbf{H}) \in (L^2(\Omega)))^{n\times n}$. We denote the two terms above by (I) and (II). Since

$$(I) = \inf_{\mathbf{H}_0\in\mathbb{R}^{n\times n}} \int_\Omega |\mathbf{F}(\mathbf{H}) - \mathbf{F}(\mathbf{H}_0)|^2\,dx\,,$$

we have $(I) \leq (II)$. By Lemma 2.13 we have

$$(II) \sim \int_\Omega (\mathbf{S}(\mathbf{H}) - \mathbf{S}(\langle \mathbf{H} \rangle_\Omega)) \cdot (\mathbf{H} - \langle \mathbf{H} \rangle_\Omega)\, dx.$$

Since $\mathbf{H} - \langle \mathbf{H} \rangle_\Omega$ has mean value zero, we can change the constant $\mathbf{S}(\langle \mathbf{H} \rangle_\Omega)$ to any other constant without changing the integral. In particular,

$$(II) \sim \int_\Omega (\mathbf{S}(\mathbf{H}) - \mathbf{S}(\mathbf{H}_1)) \cdot (\mathbf{H} - \langle \mathbf{H} \rangle_\Omega)\, dx,$$

where we define the constant $\mathbf{H}_1$ by $\mathbf{F}(\mathbf{H}_1) = \langle \mathbf{F}(\mathbf{H}) \rangle_\Omega$. Note that $\mathbf{H}_1$ is well defined, since $\mathbf{F}$ is injective and coercive and therefore invertible. Now we use (2.15), Young's inequality (2.2) with $\psi_{|\mathbf{H}|}$ and Lemma 2.13 to get

$$(II) \leq \int_\Omega c\, |\mathbf{F}(\mathbf{H}) - \mathbf{F}(\mathbf{H}_1)|^2\, dx + \frac{1}{2} \int_\Omega |\mathbf{F}(\mathbf{H}) - \mathbf{F}(\langle \mathbf{H} \rangle_\Omega)|^2\, dx = c\,(I) + \frac{1}{2}(II).$$

This proves $(II) \leq c\,(I)$. $\square$

Acknowledgement. L. Diening and M. Růžička have been supported by the SFB/TR "Geometric Partial Differential Equations". L. Belenki has been supported by the DFG Research Unit "Nonlinear Partial Differential Equations: Theoretical and Numerical Analysis". Luigi Berselli thanks the University of Freiburg for their hospitality. The authors would like to thank the referees for their helpful remarks.

REFERENCES

[1] E. Acerbi and N. Fusco, *Regularity for minimizers of nonquadratic functionals: the case* $1 < p < 2$, J. Math. Anal. Appl. **140** (1989), no. 1, 115–135.
[2] J.W. Barrett and W.B. Liu, *Finite element approximation of the p-Laplacian*, Math. Comp. **61** (1993), no. 204, 523–537.
[3] J.W. Barrett and W.B. Liu, *Quasi-norm error bounds for the finite element approximation of a non-Newtonian flow*, Numer. Math. **68** (1994), no. 4, 437–456.
[4] H. Beirão da Veiga, *Navier-Stokes equations with shear thinning viscosity. Regularity up to the boundary*, J. Math. Fluid Mech. **11** (2009), no. 2, 258–273.
[5] H. Beirão da Veiga, P. Kaplický, and M. Růžička, *Boundary regularity of shear–thickening flows*, J. Math. Fluid Mech. (2010), Online First, DOI 10.1007/s00021-010-0025-y.
[6] L. Belenki, L.C. Berselli, L. Diening, and M. Růžička, *On the Finite Element approximation of p-Stokes systems*, (2011), preprint SFB/TR 71.
[7] L.C. Berselli, L. Diening, and M. Růžička, *Existence of strong solutions for incompressible fluids with shear dependent viscosities*, J. Math. Fluid Mech. (2010), **12**, 101–132.
[8] R.B. Bird, R.C. Armstrong, and O. Hassager, *Dynamic of polymer liquids*, John Wiley, 1987.
[9] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, Springer Series in Computational Mathematics, vol. 15, Springer-Verlag, New York, 1991.
[10] P. Deuflhard, *Newton methods for nonlinear problems*, Springer Series in Computational Mathematics, vol. 35, Springer-Verlag, Berlin, 2004, Affine invariance and adaptive algorithms.
[11] L. Diening, M. Růžička, and K. Schumacher, *A decomposition technique for John domains*, Ann. Acad. Sci. Fenn. Ser. A. I. Math. **35** (2010), 87–114.
[12] L. Diening and F. Ettwein, *Fractional estimates for non-differentiable elliptic systems with general growth*, Forum Mathematicum **20** (2008), no. 3, 523–556.
[13] L. Diening and Ch. Kreuzer, *Linear convergence of an adaptive finite element method for the p-Laplacian equation*, SIAM J. Numer. Anal. **46** (2008), 614–638.
[14] L. Diening and Ch. Kreuzer, *Optimality of an adaptive finite element method for the p-Laplacian equation*, IMA J. Num. Anal., accepted (2011), DOI 10.1093/imanum/DRR016.
[15] L. Diening, J. Málek, and M. Steinhauer, *On Lipschitz truncations of Sobolev functions (with variable exponent) and their selected applications*, ESAIM: COVC **14** (2008), 211–232.

[16] L. Diening and M. Růžička, *Interpolation operators in Orlicz Sobolev spaces*, Num. Math. **107** (2007), no. 1, 107–129.

[17] C. Ebmeyer and W.B. Liu, *Quasi-norm interpolation error estimates for finite element approximations of problems with p–structure*, Numer. Math. **100** (2005), 233–258.

[18] V. Girault and J.L. Lions, *Two-grid finite-element schemes for the steady Navier-Stokes problem in polyhedra*, Port. Math. (N.S.) **58** (2001), no. 1, 25–57.

[19] V. Girault and L.R. Scott, *A quasi-local interpolation operator preserving the discrete divergence*, Calcolo **40** (2003), no. 1, 1–19.

[20] A. Hirn, *Finite element approximation of the p-Stokes equations: equal-order interpolation*, Tech. report, Institut für Angewandte Mathematik, Heidelberg, 2009, To appear.

[21] V. Kokilashvili and M. Krbec, *Weighted inequalities in Lorentz and Orlicz spaces*, Singapore etc.: World Scientific Publishing Co. Pte. Ltd. xii, 233 p., 1991 (English).

[22] O.A. Ladyzhenskaya, *The mathematical theory of viscous incompressible flow*, Gordon and Breach, New York, 1969, 2nd edition.

[23] J.L. Lions, *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Dunod, Gauthier-Villars, Paris, 1969.

[24] W.B. Liu and J.W. Barrett, *Finite element approximation of some degenerate monotone quasilinear elliptic systems*, SIAM J. Numer. Anal. **33** (1996), no. 1, 88–106.

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

[26] J. Málek and K.R. Rajagopal, *Mathematical issues concerning the Navier-Stokes equations and some of its generalizations*, Evolutionary equations. Vol. II, Handb. Differ. Equ., Elsevier/North-Holland, Amsterdam, 2005, pp. 371–459.

[27] J. Málek, K.R. Rajagopal, and M. Růžička, *Existence and regularity of solutions and the stability of the rest state for fluids with shear dependent viscosity*, M3AS **5** (1995), 789–812.

[28] J. Malý and W.P. Ziemer, *Fine regularity of solutions of elliptic partial differential equations*, American Mathematical Society, Providence, RI, 1997.

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

[30] M. Růžička, *Flow of shear dependent electrorheological fluids: unsteady space periodic case*, Applied nonlinear analysis (A. Sequeira, ed.), Kluwer/Plenum, New York, 1999, pp. 485–504.

[31] M. Růžička and L. Diening, *Non–Newtonian fluids and function spaces*, Nonlinear Analysis, Function Spaces and Applications, Proceedings of NAFSA 2006 Prague, vol. 8, 2007, pp. 95–144.

[32] D. Sandri, *Sur l'approximation numérique des écoulements quasi-newtoniens dont la viscosité suit la loi puissance ou la loi de Carreau*, RAIRO Modél. Math. Anal. Numér. **27** (1993), no. 2, 131–155.

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