OPTIMAL ERROR ESTIMATE FOR SEMI-IMPLICIT SPACE-TIME DISCRETIZATION FOR THE EQUATIONS DESCRIBING INCOMPRESSIBLE GENERALIZED NEWTONIAN FLUIDS

LUIGI C. BERSELLI*, LARS DIENING†, AND MICHAEL RŮŽIČKA‡

Abstract. In this paper we study the numerical error arising in the space-time approximation of unsteady generalized Newtonian fluids which possess a stress-tensor with $(p.\delta)$-structure. A semi-implicit time-discretization scheme coupled with conforming inf-sup stable finite element space discretization is analyzed. The main result, which improves previous suboptimal estimates as those in [A. Prohl, and M. Růžička, *SIAM J. Numer. Anal.,* 39 (2001), pp. 214–249] is the optimal $\mathcal{O}(k+h)$ error-estimate valid in the range $p \in (3/2, 2]$, where k and h are the time-step and the mesh-size, respectively. Our results hold in three-dimensional domains (with periodic boundary conditions) and are uniform with respect to the degeneracy parameter $\delta \in [0, \delta_0]$ of the extra stress tensor.

Key words. Non-Newtonian fluids, shear dependent viscosity, fully discrete problem, error analysis.

AMS subject classifications. 76A05, 35Q35, 65M15, 65M60.

1. Introduction. We study the (full) space-time discretization of a homogeneous (for simplicity the density ρ is set equal to 1), unsteady, and incompressible fluid with shear-dependent viscosity, governed by the following system of partial differential equations

$$\begin{aligned} \mathbf{u}_t - \operatorname{div} \mathbf{S}(\mathbf{Du}) + [\nabla \mathbf{u}]\mathbf{u} + \nabla \pi &= \mathbf{f} && \text{in } I \times \Omega, \\ \operatorname{div} \mathbf{u} &= 0 && \text{in } I \times \Omega, \\ \mathbf{u}(0) &= \mathbf{u}_0 && \text{in } \Omega, \end{aligned} \tag{NS$_p$}$$

where the vector field $\mathbf{u} = (u_1, u_2, u_3)$ is the velocity, $\mathbf{S}$ is the extra stress tensor, the scalar π is the kinematic pressure, the vector $\mathbf{f} = (f_1, f_2, f_3)$ is the external body force, and $\mathbf{u}_0$ is the initial velocity. We use the notation $([\nabla \mathbf{u}]\mathbf{u})_i = \sum_{j=1}^3 u_j \partial_j u_i$, $i = 1, 2, 3$, for the convective term, while $\mathbf{Du} := \frac{1}{2}(\nabla \mathbf{u} + \nabla \mathbf{u}^\top)$ denotes the symmetric part of the gradient $\nabla \mathbf{u}$. Throughout the paper we shall assume that $\Omega = (0, 2\pi)^3 \subset \mathbb{R}^3$ and we endow the problem with space periodic boundary conditions. As explained in [3, 4], this assumption simplifies the problem, allows us to prove suitable regularity results for both the continuous and the time-discrete problems, so we can concentrate on the difficulties that arise from the structure of the extra stress tensor. As usual $I = [0, T]$ denotes some non-vanishing time interval.

The most standard example of power-law like extra stress tensors in the class under consideration (cf. Assumption 2.6) is, for $p \in (1, \infty)$,

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

where $\mu > 0$ and $\delta \geq 0$ are given constants. The literature on this subject is very extensive (cf. [2, 3] and the discussion therein). Based on the results in [2, 3] we find

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

†Institute of Mathematics, LMU Munich, Theresienstr. 39, D-80333, Munich, GERMANY, (diening@mathematik.uni-muenchen.de)

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

here a suitable setting for the choice of the finite-element space-discretization and for the semi-implicit Euler scheme for time advancing, in order to show a convergence result, which is optimal apart from the h-k coupling. Previous results in this direction have been proved in [10] even if the lack of available precise regularity results lead to non-optimal results. In fact an error $\mathcal{O}(h+k)^{\frac{5p-6}{2p}}$ for $p \in \left]\frac{3+\sqrt{29}}{5}, 2\right[$ was obtained in the space-periodic three-dimensional case in [10] in the case of conforming and non-conforming finite elements.

Here, we combine the optimal estimates for the time-discretization from [3] with those for the stationary problem (without convective term) from [2], and also the results for parabolic systems in [7], to produce the optimal $\mathcal{O}(k+h)$ order of convergence for a natural distance, see Theorem 2.31 for the precise statement of the result.

Plan of the paper: In Section 2 we introduce the notation, the main hypotheses on the stress-tensor, and the properties of the numerical methods we consider. We also recall some technical results from previous papers which we will need later on. The proof of the main estimate on the numerical error is then postponed to Section 3.

2. Notation and preliminaries.

In this section we introduce the notation we will use and we also recall some technical results which will be needed in the proof of the main convergence result.

2.1. Function spaces.

We use c, C to denote generic constants, which may change from line to line, but are not depending on the crucial quantities. Moreover we write $f \sim g$ if and only if there exists constants $c, C > 0$ such that $c\, f \leq g \leq C\, f$. Given 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 use the customary Lebesgue spaces $L^p(\Omega)$ and Sobolev spaces $W^{k,p}(\Omega)$, where $\Omega = (0, 2\pi)^3$ and periodic conditions are enforced. As usual $p' := \frac{p}{p-1}$. In addition, $W^{k,p}_{\text{div}}(\Omega)$ denotes the subspace of (vector valued) functions with vanishing divergence. We will denote by $\|\,.\,\|_p$ the norm in $L^p(\Omega)$ and, in the case of zero mean value, we equip $W^{1,p}(\Omega)$ (based on the Poincaré Lemma) with the gradient norm $\|\nabla\,.\,\|_p$.

For the time-discretization, given $T > 0$ and $M \in \mathbb{N}$, we define the time-step size as $k := T/M > 0$, with the corresponding net $I^M := \{t_m\}_{m=0}^M$, $t_m := m\,k$, and we define the finite-differences backward approximation for the time derivative as:

$$d_t \mathbf{u}^m := \frac{\mathbf{u}^m - \mathbf{u}^{m-1}}{k}.$$

To deal with discrete problems we shall use the discrete spaces $l^p(I^M; X)$ consisting of X-valued sequences $\{a_m\}_{m=0}^M$, endowed with the norm

$$\|a_m\|_{l^p(I^M;X)} := \begin{cases} \left(k \displaystyle\sum_{m=0}^{M} \|a_m\|_X^p\right)^{1/p} & \text{if } 1 \leq p < \infty\,, \\ \displaystyle\max_{0\leq m\leq M} \|a_m\|_X & \text{if } p = \infty\,. \end{cases}$$

For the space discretization, $\mathcal{T}_h$ denotes a family of shape-regular, conformal triangulations, consisting of three-dimensional simplices K. We denote by h_K the diameter of K and by ρ_K the supremum of the diameters of inscribed balls. We assume that $\mathcal{T}_h$ is non-degenerate, i.e., $\max_{K\in\mathcal{T}_h}\frac{h_k}{\rho_K}\le\gamma_0$. The global mesh-size h is defined by $h:=\max_{K\in\mathcal{T}_h}h_K$. Let S_K denote the neighborhood of K, i.e., S_K is the union of all simplices of $\mathcal{T}$ touching K. By the assumptions we obtain that $|S_K|\sim|K|$ and that the number of patches S_K to which a simplex belongs is bounded uniformly with respect to h and K.

The function spaces which we will use are the following

$$X:=\big(W^{1,p}(\Omega)\big)^n,\qquad V:=\Big\{\mathbf{u}\in X:\ \fint_\Omega\mathbf{u}\,dx=0\Big\},$$

$$Y:=L^{p'}(\Omega)\,,\qquad Q:=L_0^{p'}(\Omega):=\Big\{f\in Y\ :\ \fint_\Omega f\,dx=0\Big\}.$$

In the finite element analysis, we denote by $\mathcal{P}_m(\mathcal{T}_h)$, $m\in\mathbb{N}_0$, 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 $k,m\in\mathbb{N}$ 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 $Q_h:=Y_h\cap Q$, while $\langle f,g\rangle_h:=\sum_{K\in\mathcal{T}_h}\int_K fg\,dx$ denotes the inner product in the appropriate spaces.

For the error estimates it is crucial to have projection operators well-behaving in terms of the natural norms. As in [2] we make the following assumptions on the projection operators associated with these spaces.

Assumption 2.1. *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*

1. *preserves divergence in the Y_h^*-sense, i.e.,*

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

2. *preserves periodic conditions, i.e. $\Pi_h^{\mathrm{div}}(X)\subset X_h$;*
3. *is locally $W^{1,1}$-continuous in the sense that*

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

Assumption 2.2. *We assume that Y_h contains the constant functions, i.e. that $\mathbb{R}\subset Y_h$, and that there exists a linear projection operator $\Pi_h^Y:Y\to Y_h$ which is locally L^1-continuous in the sense that*

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

For a discussion and consequences of these assumptions we refer to [2]. In particular we will need the following results:

Proposition 2.3. *Let $r\in(1,\infty)$ and let Π_h^{div} satisfy Assumption 2.1. Then Π_h^{div} has the following local continuity property*

$$\int_K\left|\nabla\Pi_h^{\mathrm{div}}\mathbf{w}\right|^r dx\le c\int_{S_K}\left|\nabla\mathbf{w}\right|^r dx$$

and the following local approximation property

$$\int_K |\mathbf{w} - \Pi_h^{\mathrm{div}} \mathbf{w}|^r \, dx \le c \int_{S_K} h_K^r |\nabla \mathbf{w}|^r \, dx,$$

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

Proof. This is special case of Thm. 3.5 in [2]. ☐

Proposition 2.4. *Let $r \in (1, \infty)$ and let Π_h^Y satisfy Assumption 2.2. Then for all $K \in \mathcal{T}_h$ and $q \in L^r(\Omega)$ we have*

$$\int_K |\Pi_h^Y q|^r \, dx \le c \int_{S_K} |q|^r \, dx.$$

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

$$\int_K |q - \Pi_h^Y q|^r \, dx \le c \int_{S_K} h_K^r |\nabla q|^r \, dx.$$

The constants depend only on r and on γ_0.

Proof. This is special case of Lemma 5.3 in [2]. ☐

Remark 2.5. By summing over $K \in \mathcal{T}_h$ one can easily get global analogues of the statements in the above Propositions.

As usual, to have a stable space-discretization, we use the following tri-linear form in the weak formulation of (space) discrete problems

$$b(\mathbf{u}, \mathbf{v}, \mathbf{w}) := \frac{1}{2}\big[\langle [\nabla \mathbf{v}]\mathbf{u}, \mathbf{w}\rangle_h - \langle [\nabla \mathbf{w}]\mathbf{u}, \mathbf{v}\rangle_h\big],$$

observing that for periodic divergence-free functions (in the continuous sense) it holds $b(\mathbf{u}, \mathbf{v}, \mathbf{w}) := \langle [\nabla \mathbf{v}]\mathbf{u}, \mathbf{w}\rangle$.

2.2. Basic properties of the extra stress tensor. For a second-order tensor $\mathbf{A} \in \mathbb{R}^{n\times n}$ we denote its symmetric part by $\mathbf{A}^{\mathrm{sym}} := \frac{1}{2}(\mathbf{A} + \mathbf{A}^\top) \in \mathbb{R}^{n\times n}_{\mathrm{sym}} := \{\mathbf{A} \in \mathbb{R}^{n\times n} \,|\, \mathbf{A} = \mathbf{A}^\top\}$. The scalar product between two tensors $\mathbf{A}, \mathbf{B}$ is denoted by $\mathbf{A} \cdot \mathbf{B}$, and we use the notation $|\mathbf{A}|^2 = \mathbf{A} \cdot \mathbf{A}^\top$. We assume that the extra stress tensor $\mathbf{S}$ has (p, δ)-structure, which will be defined now. A detailed discussion and full proofs of the following results can be found in [8, 11].

Assumption 2.6. *We assume that the extra stress tensor $\mathbf{S}\colon \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}_{\mathrm{sym}}$ belongs to $C^0(\mathbb{R}^{n\times n}, \mathbb{R}^{n\times n}_{\mathrm{sym}}) \cap C^1(\mathbb{R}^{n\times n} \setminus \{\mathbf{0}\}, \mathbb{R}^{n\times n}_{\mathrm{sym}})$, satisfies $\mathbf{S}(\mathbf{A}) = \mathbf{S}(\mathbf{A}^{\mathrm{sym}})$, 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*

$$\sum\nolimits_{i,j,k,l=1}^{n} \partial_{kl} S_{ij}(\mathbf{A}) C_{ij} C_{kl} \ge C_0 \big(\delta + |\mathbf{A}^{\mathrm{sym}}|\big)^{p-2} |\mathbf{C}^{\mathrm{sym}}|^2, \tag{2.7a}$$

$$\big|\partial_{kl} S_{ij}(\mathbf{A})\big| \le C_1 \big(\delta + |\mathbf{A}^{\mathrm{sym}}|\big)^{p-2}, \tag{2.7b}$$

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

Remark 2.8. 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 $\delta \ge 0$.

Another important set of tools are the shifted N-functions $\{\varphi_a\}_{a\geq 0}$, cf. [8, 9, 11]. To this end we define for $t \geq 0$ a special N-function φ by

$$\varphi(t) := \int_0^t \varphi'(s)\, ds \qquad \text{with} \quad \varphi'(t) := (\delta + t)^{p-2} t\,. \tag{2.9}$$

Thus we can replace in the right-hand side of (2.7) the expression $C_i\big(\delta + |\mathbf{A}^{\mathrm{sym}}|\big)^{p-2}$ by $\widetilde{C}_i\, \varphi''\big(|\mathbf{A}^{\mathrm{sym}}|\big)$, $i = 0, 1$. Next, the shifted functions are defined for $t \geq 0$ by

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

For the (p, δ)-structure we have that $\varphi_a(t) \sim (\delta + a + t)^{p-2} t^2$ and also $(\varphi_a)^*(t) \sim ((\delta + a)^{p-1} + t)^{p'-2} t^2$, where the $*$-superscript denotes the complementary function[†]. We will use also Young's inequality: For all $\varepsilon > 0$ there exists $c_\varepsilon > 0$, such that for all $s, t, a \geq 0$ it holds

$$\begin{aligned} ts &\leq \varepsilon\, \varphi_a(t) + c_\varepsilon\, (\varphi_a)^*(s)\,, \\ t\, \varphi_a'(s) + \varphi_a'(t)\, s &\leq \varepsilon\, \varphi_a(t) + c_\varepsilon\, \varphi_a(s). \end{aligned} \tag{2.10}$$

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

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

The main calculations of the paper can be performed by recalling the following lemma, which establishes the connection between $\mathbf{S}$, $\mathbf{F}$, and $\{\varphi_a\}_{a\geq 0}$ (cf. [8, 11]).

Lemma 2.12. *Let $\mathbf{S}$ satisfy Assumption 2.6, let φ be defined in* (2.9)*, and let $\mathbf{F}$ be defined in* (2.11)*. Then*

$$\begin{aligned} \big(\mathbf{S}(\mathbf{P}) - \mathbf{S}(\mathbf{Q})\big) \cdot \big(\mathbf{P} - \mathbf{Q}\big) &\sim \big|\mathbf{F}(\mathbf{P}) - \mathbf{F}(\mathbf{Q})\big|^2 && \text{(2.13a)} \\ &\sim \varphi_{|\mathbf{P}^{\mathrm{sym}}|}(|\mathbf{P}^{\mathrm{sym}} - \mathbf{Q}^{\mathrm{sym}}|) && \text{(2.13b)} \\ &\sim \varphi''\big(|\mathbf{P}^{\mathrm{sym}}| + |\mathbf{Q}^{\mathrm{sym}}|\big)|\mathbf{P}^{\mathrm{sym}} - \mathbf{Q}^{\mathrm{sym}}|^2 && \text{(2.13c)} \end{aligned}$$

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

$$\mathbf{S}(\mathbf{Q}) \cdot \mathbf{Q} \sim |\mathbf{F}(\mathbf{Q})|^2 \sim \varphi(|\mathbf{Q}^{\mathrm{sym}}|). \tag{2.13d}$$

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

Moreover, we observe that

$$|\mathbf{S}(\mathbf{P}) - \mathbf{S}(\mathbf{Q})| \sim \varphi'_{|\mathbf{P}^{\mathrm{sym}}|}\big(|\mathbf{P}^{\mathrm{sym}} - \mathbf{Q}^{\mathrm{sym}}|\big) \qquad \forall\, \mathbf{P}, \mathbf{Q} \in \mathbb{R}^{n\times n}, \tag{2.14}$$

which allows us to introduce a "*Natural distance*" since by the previous lemma we have, for all sufficiently smooth vector fields $\mathbf{u}$ and $\mathbf{w}$,

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

[†]For a N-function ψ the complementary N-function ψ^* is defined by $\psi^*(t) := \sup_{s\geq 0}(st - \psi(s))$.

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

In view of Lemma 2.12 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 (2.14), Young's inequality (2.10), and (2.13).

LEMMA 2.15. *For all $\varepsilon > 0$, there exist a constant $c_\varepsilon > 0$ (depending only on $\varepsilon > 0$ and on the characteristics of $\mathbf{S}$) such that for all sufficiently smooth vector fields $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ we have*

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

2.3. Some technical preliminary results. We recall some regularity results for fluids with shear dependent viscosities (both continuous and time-discrete) and some convergence results we will need in the sequel.

First, we recall that for the continuous problem (NS_p) we have the following existence and uniqueness result for strong solutions (cf. [4, Thm. 5.1]).

THEOREM 2.1. *Let $\mathbf{S}$ satisfy Assumption 2.6 with $\frac{7}{5} < p \le 2$ and $\delta \in [0, \delta_0]$ with $\delta_0 > 0$. Assume that $\mathbf{f} \in L^\infty(I; W^{1,2}(\Omega)) \cap W^{1,2}(I; L^2(\Omega))$ and also $\mathbf{u}_0 \in W^{2,2}_{\mathrm{div}}(\Omega)$, $\langle \mathbf{u}_0, 1\rangle = 0$, and $\operatorname{div} \mathbf{S}(\mathbf{Du}_0) \in L^2(\Omega)$. Then, there exist a time $T' > 0$ and a constant $c_0 > 0$, both depending on $(\delta_0, p, C_0, \mathbf{f}, \mathbf{u}_0, T, \Omega)$ but independent of δ, such that the system (NS_p) has a unique strong solution $\mathbf{u} \in L^p(I'; W^{1,p}_{\mathrm{div}}(\Omega))$, $I' = [0, T']$ such that*

$$\|\mathbf{u}_t\|_{L^\infty(I';L^2(\Omega))} + \|\mathbf{F}(\mathbf{Du})\|_{W^{1,2}(I'\times\Omega)} + \|\mathbf{F}(\mathbf{Du})\|_{L^{2\frac{5p-6}{2-p}}(I';W^{1,2}(\Omega))} \le c_0 . \tag{2.16}$$

In particular this implies, uniformly in $\delta \in [0, \delta_0]$,

$$\mathbf{u} \in L^{\frac{p(5p-6)}{2-p}}(I'; W^{2,\frac{3p}{p+1}}(\Omega)) \cap C(I'; W^{1,r}(\Omega)) \qquad \text{for } 1 \le r < 6(p-1), \tag{2.17a}$$

$$\mathbf{u}_t \in L^\infty(I'; L^2(\Omega)) \cap L^{\frac{p(5p-6)}{(3p-2)(p-1)}}(I'; W^{1,\frac{3p}{p+1}}(\Omega)) . \tag{2.17b}$$

The above theorem, whose proof employs in a substantial manner the hypothesis of space-periodicity, has been used to prove the following optimal convergence result for the numerical error with respect to a semi-implicit time discretization (cf. [3, Thm 1.1, 4.1]).

THEOREM 2.2. *Let $\mathbf{S}$ satisfy Assumption 2.6 with $p \in \left(\frac{3}{2}, 2\right]$ and $\delta \in [0, \delta_0]$, where $\delta_0 > 0$ is an arbitrary number. Let $\mathbf{f} \in C(I; W^{1,2}(\Omega)) \cap W^{1,2}(I; L^2(\Omega))$, where $I = [0, T]$, for some $T > 0$, and let $\mathbf{u}_0 \in W^{2,2}_{\mathrm{div}}(\Omega)$ with $\operatorname{div} \mathbf{S}(\mathbf{Du}_0) \in L^2(\Omega)$ be given. Let $\mathbf{u}$ be a strong solution of the (continuous) problem (NS_p) satisfying*

$$\|\mathbf{u}_t\|_{L^\infty(I;L^2(\Omega))} + \|\mathbf{F}(\mathbf{Du})\|_{W^{1,2}(I\times\Omega)} + \|\mathbf{F}(\mathbf{Du})\|_{L^{2\frac{5p-6}{2-p}}(I;W^{1,2}(\Omega))} \le c_1. \tag{2.18}$$

Then, there exists $k_0 > 0$ such that for $k \in (0, k_0)$ the unique time-discrete solution $\mathbf{u}^m$ of the semi-implicit time-discrete iterative scheme

$$\begin{aligned} d_t \mathbf{u}^m - \operatorname{div} \mathbf{S}(\mathbf{Du}^m) + [\nabla \mathbf{u}^m]\mathbf{u}^{m-1} + \nabla \pi^m &= \mathbf{f}(t_m) \qquad && \text{in } \Omega, \\ \operatorname{div} \mathbf{u}^m &= 0 && \text{in } \Omega, \end{aligned} \tag{NS_p^k}$$

(endowed with periodic boundary conditions) satisfies the error estimate

$$\max_{0\le m\le M} \|\mathbf{u}(t_m) - \mathbf{u}^m\|_2^2 + k \sum_{m=0}^{M} \|\mathbf{F}(\mathbf{Du}(t_m)) - \mathbf{F}(\mathbf{Du}^m)\|_2^2 \le c\, k^2,$$

where the constants k_0 and c depend on c_1 and on the characteristics of $\mathbf{S}$, but are independent of δ. Moreover, for each $1 \le r < 6(p-1)$, it holds

$$\mathbf{u}^m \in l^{\frac{p(5p-6)}{2-p}}(I^M; W^{2,\frac{3p}{p+1}}(\Omega)) \cap l^\infty(I^M; W^{1,r}(\Omega)), \tag{2.19a}$$

$$d_t\mathbf{u}^m \in l^\infty(I^M; L^2(\Omega)) \cap l^{\frac{p(5p-6)}{(3p-2)(p-1)}}(I^M; W^{1,\frac{3p}{p+1}}(\Omega)). \tag{2.19b}$$

We observe that by parabolic interpolation, (cf. [3, Rem. 2.7]) it also follows that

$$d_t\mathbf{u}^m \in l^{\frac{11p-12}{3(p-1)}}(I^M; L^{\frac{11p-12}{3(p-1)}}(\Omega)),$$

and consequently

$$d_t\mathbf{u}^m \in l^{\frac{p}{p-1}}(I^M; L^{\frac{p}{p-1}}(\Omega)) = l^{p'}(I^M; L^{p'}(\Omega)) \quad \text{if } \; p > \frac{3}{2}. \tag{2.20}$$

The latter property will have a relevant role to estimate in the error equation the term involving the discrete pressure.

One main tool in the sequel will be also the following generalized Gronwall lemma, which is a minor variation of that proved in great detail in [3, Lemma 3.3].

LEMMA 2.3. *Let $1 < p \le 2$ and let be given two non-negative sequences $\{a_m\}_m$ and $\{b_m\}_m$, and two sequences $\{r_m\}_m$ and $\{s_m\}_m$ for which there exists $\gamma_0 > 0$ such that for all $0 < h < 1/\sqrt{\gamma_0}$:*

$$a_0^2 \le \gamma_0\, h^2, \quad b_0^2 \le \gamma_0\, h^2, \quad k\sum_{m=0}^{M} r_m^2 \le \gamma_0\, h^2, \quad \text{and} \quad k\sum_{m=0}^{M} s_m^2 \le \gamma_0\, h^2. \tag{2.21}$$

Further, let there exist constants $\gamma_1, \gamma_2, \gamma_3 > 0$, $\Lambda > 0$, and some $0 < \theta \le 1$ such that for some $\lambda \in [0, \Lambda]$ the following two inequalities are satisfied for all $m \ge 1$:

$$d_t a_m^2 + \gamma_1(\lambda + b_m)^{p-2} b_m^2 \le b_m r_m + \gamma_2 b_{m-1} b_m + s_m^2, \tag{2.22}$$

$$d_t a_m^2 + \gamma_1(\lambda + b_m)^{p-2} b_m^2 \le b_m r_m + \gamma_3 b_m b_{m-1}^{1-\theta} a_m^\theta + s_m^2. \tag{2.23}$$

Then, there exist $\overline{k}, \overline{\gamma_0} > 0$ such that if $h^2 < \overline{\gamma_0}\, k$ and if (2.22), (2.23) hold for $0 < k < \overline{k} \le 1$, then there exist γ_4, $\gamma_5 > 0$, independent of λ, such that

$$\max_{1\le m\le M} b_m \le 1, \tag{2.24}$$

$$\max_{1\le m\le M} a_m^2 + \gamma_1(\lambda+\Lambda)^{p-2} k \sum_{m=1}^{N} b_m^2 \le \gamma_4\, h^2\, exp(2\gamma_5 k\, M). \tag{2.25}$$

Proof. The proof of this result is a simple adaption of that of [3, Lemma 3.3]. Nevertheless we report the main changes needed to accomplish the proof. In particular, we will use it for $a_m := \|\mathbf{u}^m - \mathbf{u}_h^m\|_2$ and $b_m := \|\mathbf{D}\mathbf{u}^m - \mathbf{D}\mathbf{u}_h^m\|_p$.

The proof goes by induction on $1 \le N \le M$. Since in the inequality (2.23) the term b_{m-1} is present and since contrary to Ref. [3] $a_0, b_0 \neq 0$, some care has to be taken to start the induction argument. The most important part of the proof is that of showing that $b_m \le 1$, because then the estimate (2.25) will follow by applying the classical discrete Gronwall lemma. We will use the same argument to check as starting inductive step that (2.25) is satisfied for $N = 1$, as well as to show that if inequality (2.25) is satisfied for a given $N \geq 1$, then holds true also for $N + 1$.

Let us suppose *per absurdum* that $b_N > 1$, while $b_m \leq 1$ for $m < N$. We multiply (2.22) by k and we sum over m, for $m = 1, \dots, N$. It readily follows that:

$$a_N^2 + \gamma_1 k \sum_{m=1}^{N} (\lambda + b_m)^{p-2} b_m^2 \leq$$
$$\leq a_0^2 + \frac{\gamma_1}{2} k \sum_{m=1}^{N} (\lambda + b_m)^{p-2} b_m^2 + \frac{1}{\gamma_1} k \sum_{m=1}^{N} (\lambda + b_m)^{2-p} (r_m^2 + \gamma_2^2 b_{m-1}^2) + k \sum_{m=1}^{N} s_m^2.$$

We absorb the second term from the right-hand side in the left-hand side and we observe that $(\lambda + b_m)^{2-p} \leq (\lambda + b_N)^{2-p} \leq (\lambda + b_N)^{2(2-p)}$, regardless of the value of $\lambda \geq 0$. Neglecting all terms on the left-hand side, except the one with $m = N$, and dividing both sides by $\frac{\gamma_1}{2} k (\lambda + b_N)^{p-2} \neq 0$ we get,

$$b_N^2 \leq \frac{2(\lambda + b_N)^{2(2-p)}}{k\gamma_1} \left[a_0^2 + \frac{1}{\gamma_1} k \sum_{m=1}^{N} (r_m^2 + \gamma_2^2 b_{m-1}^2) + k \sum_{m=1}^{N} s_m^2 \right]. \tag{2.26}$$

Now, if we are dealing with the initial step $N = 1$, we have on the right-hand side of (2.26) a term containing b_0^2 on which we need to show that it satisfies (2.25). The hypothesis (2.21) and the restriction on h imply that $a_0 \leq \gamma_0 h^2$ and $b_0 \leq 1$. We also need to satisfy the same estimate (2.25) when $m = 0$, namely:

$$\gamma_1 (\lambda + \Lambda)^{p-2} k \, b_0^2 \leq \gamma_4 \, h^2 \exp(2\gamma_5 k \, M).$$

Since $k \leq 1$, and given $\gamma_0 > 0$, it is enough to choose $\gamma_4 > 0$ large enough such that the following inequality is satisfied

$$\gamma_0 \leq \min \left\{ 1, \frac{(2\Lambda)^{2-p}}{\gamma_1} \right\} \gamma_4 \exp(2\gamma_5 k \, M). \tag{2.27}$$

Observe that this choice is always possible since $p \leq 2$.

On the other hand, in the calculations with $N > 1$ we can simply use (2.25) (which starts at $N = 1$ if (2.27) is satisfied) as inductive assumption to estimate the right-hand side of (2.26).

As a result of the choice of γ_4, in both cases it follows with the same algebraic manipulations of [3, Lemma 3.3] that we can bound the right-hand side of (2.26) as follows:

$$1 < b_N^{2(p-1)} \leq \frac{h^2}{k} \frac{2\,(1 + \Lambda)^{2(2-p)}}{\gamma_1} \left[\gamma_0 + \frac{\gamma_0}{\gamma_1} + \frac{\gamma_2^2 \gamma_4}{\gamma_1 (\lambda + \Lambda)^{p-2}} \exp(2\gamma_5 k \, N) + \gamma_0 \right].$$

This gives a contradiction, provided that

$$\frac{h^2}{k} \leq \frac{\gamma_1}{2\,(1 + \Lambda)^{2(2-p)}} \left[\gamma_0 + \frac{\gamma_0}{\gamma_1} + \frac{\gamma_2^2 \gamma_4}{\gamma_1 (\lambda + \Lambda)^{p-2}} \exp(2\gamma_5 k \, N) + \gamma_0 \right]^{-1} := \overline{\gamma_0}.$$

This finally proves that $b_N \leq 1$. Then the rest of the proof goes exactly as in the cited lemma, with a further application of the standard discrete Gronwall lemma. It is in this last step that one has to assume the limitation $k < \overline{k} := \min\{1, (2\gamma_5)^{-1}\}$. □

Inspecting the proof it is clear that we have also the following result, where the doubling of the constant on the right-hand side comes from having the same estimate separately for a_0, b_0 and for a_m, b_m, for $m > 1$.

COROLLARY 2.28. *Let the same hypotheses of Lemma 2.3 be satisfied, then in addition we have that*

$$\max_{0\le m\le M} b_m \le 1, \tag{2.29}$$

$$\max_{0\le m\le M} a_m^2 + \gamma_1(\lambda+\Lambda)^{p-2} k \sum_{m=0}^{M} b_m^2 \le 2\gamma_4\, h^2\, exp(2\gamma_5 k\, M). \tag{2.30}$$

2.4. Numerical algorithms. We write explicitly the numerical algorithms we will consider and state some basic existence results for the space-time-discrete solutions.

Given a net I^M, a triangulation $\mathcal{T}_h$ of Ω, and conforming spaces V_h, Q_h, (recall notation from Sec. 2.1) for the space-time-discrete problem, we use the following algorithm:

Algorithm (space-time-discrete, Euler semi-implicit) Let $\mathbf{u}_h^0 = \Pi_h^{\text{div}}\mathbf{u}_0$. Then, for $m \ge 1$ and $\mathbf{u}_h^{m-1} \in V_h$ given from the previous time-step, compute the iterate $(\mathbf{u}_h^m, \pi_h^m) \in V_h \times Q_h$ such that for all $\boldsymbol{\xi}_h \in V_h$, and $\eta_h \in Q_h$

$$\begin{aligned}\langle d_t\mathbf{u}_h^m, \boldsymbol{\xi}_h\rangle_h + \langle \mathbf{S}(\mathbf{D}\mathbf{u}_h^m), \mathbf{D}\boldsymbol{\xi}_h\rangle_h + b(\mathbf{u}_h^{m-1}, \mathbf{u}_h^m, \boldsymbol{\xi}_h) - \langle \operatorname{div}\boldsymbol{\xi}_h, \pi_h^m\rangle_h &= \langle \mathbf{f}(t_m), \boldsymbol{\xi}_h\rangle_h,\\ \langle \operatorname{div}\mathbf{u}_h^m, \eta_h\rangle_h &= 0.\end{aligned} \tag{Q_h^m}$$

We also observe that the (space-continuous) time-discrete scheme (NS_p^k) from Theorem 2.2 can be formulated in a weak form as follows: Let be given $\mathbf{u}^0 = \mathbf{u}_0$, $m \ge 1$, and $\mathbf{u}^{m-1} \in V$ evaluated from the previous time-step, compute the iterate $(\mathbf{u}^m, \pi^m) \in V \times Q$ such that for all $\boldsymbol{\xi} \in V$, and $\eta \in Q$

$$\begin{aligned}\langle d_t\mathbf{u}^m, \boldsymbol{\xi}\rangle + \langle \mathbf{S}(\mathbf{D}\mathbf{u}^m), \mathbf{D}\boldsymbol{\xi}\rangle + b(\mathbf{u}^{m-1}, \mathbf{u}^m, \boldsymbol{\xi}) - \langle \operatorname{div}\boldsymbol{\xi}, \pi^m\rangle &= \langle \mathbf{f}(t_m), \boldsymbol{\xi}\rangle,\\ \langle \operatorname{div}\mathbf{u}^m, \eta\rangle &= 0.\end{aligned} \tag{Q^m}$$

The existence of a solution $(\mathbf{u}^m, \pi^m)$ and its uniqueness follow from Thm. 2.2, concerning strong solutions $\mathbf{u}^m \in V(0)$ of (NS_p^k). This solution is a fortiori also a weak solution of the following problem: Find $\mathbf{u}^m \in V(0)$ such that for all $\boldsymbol{\xi} \in V(0)$

$$\langle d_t\mathbf{u}^m, \boldsymbol{\xi}\rangle + \langle \mathbf{S}(\mathbf{D}\mathbf{u}^m), \mathbf{D}\boldsymbol{\xi}\rangle + b(\mathbf{u}^{m-1}, \mathbf{u}^m, \boldsymbol{\xi}) = \langle \mathbf{f}(t_m), \boldsymbol{\xi}\rangle, \tag{P^m}$$

where $V(0) := \{\mathbf{w} \in V : \ \langle \operatorname{div}\mathbf{w}, \eta\rangle = 0, \ \forall\, \eta \in Y\}$. The existence of the associated pressure $\pi^m \in Q$ follows then from the DeRham theorem and the inf-sup condition.

The situation for the space-time-discrete problem is similar: The existence of the solution $(\mathbf{u}_h^m, \pi_h^m)$ can be inferred in the following way. First, for $V_h(0) = \{\mathbf{w}_h \in V_h : \langle \operatorname{div}\mathbf{w}_h, \eta_h\rangle = 0, \ \forall\, \eta_h \in Y_h\}$ consider the following algorithm, given $\mathbf{u}^{m-1} \in V_h(0)$ find $\mathbf{u}^m \in V_h(0)$ such that for all $\boldsymbol{\xi}_h \in V(0)$

$$\langle d_t\mathbf{u}_h^m, \boldsymbol{\xi}_h\rangle_h + \langle \mathbf{S}(\mathbf{D}\mathbf{u}_h^m), \mathbf{D}\boldsymbol{\xi}_h\rangle_h + b(\mathbf{u}_h^{m-1}, \mathbf{u}_h^m, \boldsymbol{\xi}_h) = \langle \mathbf{f}(t_m), \boldsymbol{\xi}_h\rangle_h\,. \tag{P_h^m}$$

The existence of a weak solution for (P_h^m) follows directly by applying the Brouwer's theorem, see also [10, Lemma 7.1]. Uniqueness follows from the semi-implicit expression for the convective term and from the monotonicity of $\mathbf{S}$. Moreover, the following energy estimate holds true:

$$\max_{0\le m\le M}\|\mathbf{u}_h^m\|_2^2 + k\sum_{m=0}^{M}\|\mathbf{D}\mathbf{u}_h^m\|_p^p \le C(\mathbf{u}_0,\mathbf{f}),$$

which is obtained by using $\mathbf{u}_h^m \in V_h(0)$ as test function. Coming back to Problem (Q_h^m), the existence of the associated pressure $\pi_h^m \in Q_h$ such that $(\mathbf{u}_h^m, \pi_h^m)$ is a solution of (Q_h^m) is derived from the previous result of existence of a solution for (P_h^m) and the inf-sup condition. See also [2, Lemma 4.1] for such inequality in the setting of Orlicz spaces.

The main result of this paper is the following error estimate.

THEOREM 2.31. *Let* $\mathbf{S}$ *satisfy Assumption 2.6 with* $p \in \left(\frac{3}{2}, 2\right]$ *and* $\delta \in [0, \delta_0]$, *where* $\delta_0 > 0$ *is an arbitrary number. Let* $\mathbf{f} \in C(I; W^{1,2}(\Omega)) \cap W^{1,2}(I; L^2(\Omega))$, *where* $I = [0, T]$, *for some* $T > 0$, *and let* $\mathbf{u}_0 \in W^{2,2}_{\text{div}}(\Omega)$ *with* $\operatorname{div}\mathbf{S}(\mathbf{D}\mathbf{u}_0) \in L^2(\Omega)$ *be given. Let* $\mathbf{u}$ *be a strong solution of the (continuous) problem* (NS_p) *satisfying*

$$\|\mathbf{u}_t\|_{L^\infty(I;L^2(\Omega))} + \|\mathbf{F}(\mathbf{D}\mathbf{u})\|_{W^{1,2}(I\times\Omega)} + \|\mathbf{F}(\mathbf{D}\mathbf{u})\|_{L^{2\frac{5p-6}{2-p}}(I;W^{1,2}(\Omega))} \le c_2. \tag{2.32}$$

Let $\mathcal{T}_h$ *be a triangulation as introduced in Sec. 2 and let* $(\mathbf{u}_h^m, \pi_h^m)$ *be the unique solution of the space-time-discrete problem* (Q_h^m) *corresponding to the data* $(\Pi_h^{\text{div}}\mathbf{u}_0, \mathbf{f})$. *Then, there exists a time-step* $k_1 > 0$ *and a mesh-size* $h_1 > 0$ *such that, if* $\max\{h^{\frac{3p-2}{2}}, h^2\} \le c_3\, k$ *for some* $c_3 > 0$, *for all* $k \in (0, k_1)$ *and for all* $h \in (0, h_1)$, *then the following error estimate holds true:*

$$\max_{0\le m\le M}\|\mathbf{u}(t_m) - \mathbf{u}_h^m\|_2^2 + k\sum_{m=0}^{M}\|\mathbf{F}(\mathbf{D}\mathbf{u}(t_m)) - \mathbf{F}(\mathbf{D}\mathbf{u}_h^m)\|_2^2 \le c_4\,(h^2 + k^2).$$

The constants k_1, h_1, c_3 *and* c_4 *depend only on* c_2, *the characteristics of* $\mathbf{S}$, *and* $|\Omega|$, *but they are independent of* $\delta \in [0, \delta_0]$.

REMARK 2.33. A explained in [3], in the space periodic setting we are able to obtain (2.32) starting from the assumptions on the data of the problem, at least in a small time interval $[0, T']$, see Thm. 2.1. On the other hand, the analysis performed below is correct also in the Dirichlet case, provided one can show the regularities (2.32) and (2.19) for the continuous and time-discrete problem, respectively.

In the case of Dirichlet data, the semigroup approach of Bothe and Prüss [5] proves the existence and uniqueness of a strong solution in a small interval $[0, T']$ for $p \ge 1$, under the hypothesis of smooth data and $\delta > 0$. Note that their analysis does not ensure the regularity of the time derivative stated in (2.32). However, one can easily prove this property based on the results proved in [5] by standard techniques.

Unfortunately the needed space regularity is still open even for steady problems with (p, δ)-structure in the Dirichlet case, and this would be the basis for the regularity of the time-discrete problem. For partial results in the steady case, see for instance Beirão da Veiga [1].

We also wish to point out that one of the main difficulties in the Dirichlet case is that of having estimate independent of δ, which is one of the key points in our analysis also of the degenerate problem.

3. Proof of the main result. The proof of Thm. 2.31 is obtained by splitting the numerical error as follows:

$$\mathbf{u}(t_m) - \mathbf{u}_h^m = \mathbf{u}(t_m) - \mathbf{u}^m + \mathbf{u}^m - \mathbf{u}_h^m =: \boldsymbol{\varepsilon}^m + \mathbf{e}^m,$$

For the error $\boldsymbol{\varepsilon}^m$ Thm. 2.2 ensures

$$\max_{0\le m\le M} \| \mathbf{u}(t_m) - \mathbf{u}^m\|_2^2 + k \sum_{m=0}^{M} \|\mathbf{F}(\mathbf{D}\mathbf{u}(t_m)) - \mathbf{F}(\mathbf{D}\mathbf{u}^m)\|_2^2 \le c\, k^2.$$

Hence, we need to focus only on the second part of the error, namely $\mathbf{e}^m$.

The main error estimate is obtained by taking the difference between the equation satisfied by $\mathbf{u}^m$ and that for $\mathbf{u}_h^m$, and using as test function $\boldsymbol{\xi}_h \in V_h \subset V$. In this way we obtain the following error equation for all $\boldsymbol{\xi}_h \in V_h$

$$\begin{aligned}\langle d_t \mathbf{e}^m, \boldsymbol{\xi}_h\rangle_h + \langle \mathbf{S}(\mathbf{D}\mathbf{u}^m) - \mathbf{S}(\mathbf{D}\mathbf{u}_h^m), \mathbf{D}\boldsymbol{\xi}_h\rangle_h + b(\mathbf{u}^{m-1}, \mathbf{u}^m, \boldsymbol{\xi}_h)&\\ - b(\mathbf{u}_h^{m-1}, \mathbf{u}_h^m, \boldsymbol{\xi}_h) - \langle \operatorname{div} \boldsymbol{\xi}_h, \pi^m - \pi_h^m\rangle_h = 0.&\end{aligned} \tag{3.1}$$

Clearly, a "natural" test function $\boldsymbol{\xi}_h$ to get the error estimate would be $\mathbf{e}^m := \mathbf{u}^m - \mathbf{u}_h^m$, which cannot be used, since it is not a discrete functions, that is $\mathbf{e}^m \notin V_h$. The error estimate is then obtained by using as test function the projection $\boldsymbol{\xi}_h := \Pi_h^{\operatorname{div}} \mathbf{e}^m \in V_h$ and treating the various terms arising from the following identity:

$$\begin{aligned}\Pi_h^{\operatorname{div}} \mathbf{e}^m = \Pi_h^{\operatorname{div}}(\mathbf{u}^m - \mathbf{u}_h^m) = \Pi_h^{\operatorname{div}} \mathbf{u}^m - \mathbf{u}_h^m &= \Pi_h^{\operatorname{div}} \mathbf{u}^m - \mathbf{u}^m + \mathbf{u}^m - \mathbf{u}_h^m\\ &=: \mathbf{R}_h^m + \mathbf{e}^m,\end{aligned}$$

where we used that $\Pi_h^{\operatorname{div}} = \operatorname{id}$ on V_h. Let us start from the first term from the left-hand side of the error equation, that one concerning the discrete time-derivative. We have the following result

LEMMA 3.1. *The following estimate holds true*

$$\frac{1}{2} d_t \|\mathbf{e}^m\|_2^2 + \frac{k}{4}\|d_t \mathbf{e}^m\|_2^2 - \frac{1}{k}\|\mathbf{R}_h^m\|_2^2 \le \langle d_t \mathbf{e}^m, \Pi_h^{\operatorname{div}} \mathbf{e}^m\rangle_h.$$

Proof. By standard manipulations of the discrete time-derivative we get

$$\begin{aligned}\langle d_t \mathbf{e}^m, \Pi_h^{\operatorname{div}} \mathbf{e}^m\rangle_h &= \langle d_t \mathbf{e}^m, \mathbf{e}^m\rangle_h + \langle d_t \mathbf{e}^m, \mathbf{R}_h^m\rangle_h\\ &= \frac{1}{2} d_t \|\mathbf{e}^m\|_2^2 + \frac{k}{2}\|d_t \mathbf{e}^m\|_2^2 + \langle d_t \mathbf{e}^m, \mathbf{R}_h^m\rangle_h.\end{aligned}$$

Observe now that, by Young's inequality, we have

$$\big|\langle d_t \mathbf{e}^m, \mathbf{R}_h^m\rangle\big| \le \frac{k}{4}\|d_t \mathbf{e}^m\|_2^2 + \frac{1}{k}\|\mathbf{R}_h^m\|_2^2,$$

hence the statement. ☐

Next, we treat the second term from the left-hand side of the error equation, that one related with the extra stress-tensor.

LEMMA 3.2. *There exists $c > 0$, independent of h and δ, such that*

$$\begin{aligned}c\left(\|\mathbf{F}(\mathbf{D}\mathbf{u}^m) - \mathbf{F}(\mathbf{D}\mathbf{u}_h^m)\|_2^2 - \|\mathbf{F}(\mathbf{D}\mathbf{u}^m) - \mathbf{F}(\mathbf{D}\Pi_h^{\operatorname{div}} \mathbf{u}^m)\|_2^2\right)&\\ \le \langle \mathbf{S}(\mathbf{D}\mathbf{u}^m) - \mathbf{S}(\mathbf{D}\mathbf{u}_h^m), \mathbf{D}(\Pi_h^{\operatorname{div}} \mathbf{e}^m)\rangle_h.&\end{aligned}$$

Proof. From standard properties concerning the structure of $\mathbf{S}$, as recalled in Sec. 2.2, we get

$$
\begin{aligned}
&\langle \mathbf{S}(\mathbf{D}\mathbf{u}^m) - \mathbf{S}(\mathbf{D}\mathbf{u}_h^m), \mathbf{D}(\Pi_h^{\text{div}}\mathbf{u}^m - \mathbf{u}_h^m)\rangle_h \\
&= \langle \mathbf{S}(\mathbf{D}\mathbf{u}^m) - \mathbf{S}(\mathbf{D}\mathbf{u}_h^m), \mathbf{D}\mathbf{u}^m - \mathbf{D}\mathbf{u}_h^m\rangle_h + \langle \mathbf{S}(\mathbf{D}\mathbf{u}^m) - \mathbf{S}(\mathbf{D}\mathbf{u}_h^m), \mathbf{D}\Pi_h^{\text{div}}\mathbf{u}^m - \mathbf{D}\mathbf{u}^m\rangle_h \\
&= \|\mathbf{F}(\mathbf{D}\mathbf{u}^m) - \mathbf{F}(\mathbf{D}\mathbf{u}_h^m)\|_2^2 + \langle \mathbf{S}(\mathbf{D}\mathbf{u}^m) - \mathbf{S}(\mathbf{D}\mathbf{u}_h^m), \mathbf{D}(\Pi_h^{\text{div}}\mathbf{u}^m - \mathbf{u}^m)\rangle_h .
\end{aligned}
$$

The latter term on the right-hand side can be estimated with the help of Lemma 2.15 as follows

$$
\begin{aligned}
&\big|\langle \mathbf{S}(\mathbf{D}\mathbf{u}^m) - \mathbf{S}(\mathbf{D}\mathbf{u}_h^m), \mathbf{D}(\Pi_h^{\text{div}}\mathbf{u}^m - \mathbf{u}^m)\rangle_h\big| \\
&\qquad \le \varepsilon\|\mathbf{F}(\mathbf{D}\mathbf{u}^m) - \mathbf{F}(\mathbf{D}\mathbf{u}_h^m)\|_2^2 + c_\varepsilon\|\mathbf{F}(\mathbf{D}\mathbf{u}^m) - \mathbf{F}(\mathbf{D}\Pi_h^{\text{div}}\mathbf{u}^m)\|_2^2,
\end{aligned}
$$

ending the proof. □

Some care is needed also to handle the two terms coming from the convective term, which are estimated in the following lemma, by using the regularity results for the solution $\mathbf{u}^m$ of the time-discrete problem.

Lemma 3.3. *There exist $c > 0$ and $\theta \in]0, 1[$, not depending on h and δ, such that*

$$
\begin{aligned}
&\big|b(\mathbf{u}^{m-1}, \mathbf{u}^m, \Pi_h^{\text{div}}\mathbf{e}^m) - b(\mathbf{u}_h^{m-1}, \mathbf{u}_h^m, \Pi_h^{\text{div}}\mathbf{e}^m)\big| \\
&\qquad \le c\,\big(\|\nabla\mathbf{R}_h^m\|_{\frac{3p}{3+1}}\|\mathbf{D}\mathbf{e}^m\|_p + \|\mathbf{e}^{m-1}\|_2^\theta\|\mathbf{D}\mathbf{e}^{m-1}\|_p^{1-\theta}\|\mathbf{D}\mathbf{e}^m\|_p\big).
\end{aligned} \tag{3.2}
$$

Proof. By adding and subtracting $b(\mathbf{u}^{m-1}, \Pi_h^{\text{div}}\mathbf{u}^m, \Pi_h^{\text{div}}\mathbf{e}^m)$ and also in a second step $b(\mathbf{u}_h^{m-1}, \Pi_h^{\text{div}}\mathbf{u}^m, \Pi_h^{\text{div}}\mathbf{e}^m)$ and, by observing that $b(\mathbf{u}_h^{m-1}, \Pi_h^{\text{div}}\mathbf{e}^m, \Pi_h^{\text{div}}\mathbf{e}^m) = 0$, we get

$$
\begin{aligned}
&b(\mathbf{u}^{m-1}, \mathbf{u}^m, \Pi_h^{\text{div}}\mathbf{e}^m) - b(\mathbf{u}_h^{m-1}, \mathbf{u}_h^m, \Pi_h^{\text{div}}\mathbf{e}^m) \\
&= b(\mathbf{u}^{m-1}, \mathbf{u}^m - \Pi_h^{\text{div}}\mathbf{u}^m, \Pi_h^{\text{div}}\mathbf{e}^m) + b(\mathbf{u}_h^{m-1}, \Pi_h^{\text{div}}(\mathbf{u}^m - \mathbf{u}_h^m), \Pi_h^{\text{div}}\mathbf{e}^m) \\
&\quad + b(\mathbf{u}^{m-1} - \mathbf{u}_h^{m-1}, \Pi_h^{\text{div}}\mathbf{u}^m, \Pi_h^{\text{div}}\mathbf{e}^m) \\
&= b(\mathbf{u}^{m-1}, \mathbf{R}_h^m, \Pi_h^{\text{div}}\mathbf{e}^m) + b(\mathbf{e}^{m-1}, \Pi_h^{\text{div}}\mathbf{u}^m, \Pi_h^{\text{div}}\mathbf{e}^m) \\
&=: I_1 + I_2.
\end{aligned}
$$

Since $\operatorname{div}\mathbf{u}^{m-1} = 0$ (in the continuous sense) the first term is estimated as follows, by using Hölder inequality

$$
\begin{aligned}
I_1 &= b(\mathbf{u}^{m-1}, \mathbf{R}_h^m, \Pi_h^{\text{div}}\mathbf{e}^m) = \langle[\nabla\mathbf{R}_h^m]\mathbf{u}^{m-1}, \Pi_h^{\text{div}}\mathbf{e}^m\rangle_h \\
&\le \|\mathbf{u}^{m-1}\|_{\frac{3p}{3p-4}}\|\nabla\mathbf{R}_h^m\|_{\frac{3p}{p+1}}\|\Pi_h^{\text{div}}\mathbf{e}^m\|_{\frac{3p}{3-p}}
\end{aligned}
$$

provided that $p > \frac{4}{3}$. By a Sobolev embedding theorem, the Korn's inequality (valid in the case of functions vanishing at the boundary or with zero mean value), and by the continuity of the interpolation operator Π_h^{div} (cf. Prop. 2.3, Rem. 2.5) we can write

$$
\|\Pi_h^{\text{div}}\mathbf{e}^m\|_{\frac{3p}{3-p}} \le c\,\|\nabla\Pi_h^{\text{div}}\mathbf{e}^m\|_p \le c\,\|\mathbf{D}\mathbf{e}^m\|_p.
$$

Thus we arrive at $I_1 \le c\,\|\mathbf{u}^{m-1}\|_\infty\|\nabla\mathbf{R}_h^m\|_{\frac{3p}{3+1}}\|\mathbf{D}\mathbf{e}^m\|_p$ and now observe that, by using regularity (2.19) of the solution $\mathbf{u}^m$, this term is bounded by the first one from the right-hand side of (3.2).

Concerning I_2, by using the definition of $b(\,.\,,\,.\,,\,.\,)$, we split it as follows

$$I_2 = I_{2,1} + I_{2,2} := \frac{1}{2}\langle[\nabla\Pi_h^{\text{div}}\mathbf{e}^m]\mathbf{e}^{m-1}, \Pi_h^{\text{div}}\mathbf{u}^m\rangle_h - \frac{1}{2}\langle[\nabla\Pi_h^{\text{div}}\mathbf{u}^m]\mathbf{e}^{m-1}, \Pi_h^{\text{div}}\mathbf{e}^m\rangle_h.$$

We estimate $I_{2,1}$ with the Hölder inequality:

$$I_{2,1} \le c\,\|\Pi_h^{\text{div}}\mathbf{u}^m\|_\alpha\|\mathbf{e}^{m-1}\|_{s_1}\|\nabla\Pi_h^{\text{div}}\mathbf{e}^m\|_p,$$

for some $s_1 \in (p', p^*) = \big(\frac{p}{p-1}, \frac{3p}{3-p}\big)$ and $\alpha = \frac{ps_1}{ps_1 - s_1 - p} < \infty$. We have that $2 \le p'$, hence, by standard convex interpolation with $\theta \in (0,1)$ such that $\frac{1}{s_1} = \frac{\theta}{2} + \frac{(1-\theta)}{p^*}$, by the properties of Π_h^{div} (cf. Prop. 2.3, Rem. 2.5), by Korn's inequality, and since (2.19) implies $\|\mathbf{u}^m\|_\alpha \le c\,\|\mathbf{u}^m\|_\infty \in l^\infty(I_M)$, we obtain that

$$\begin{aligned}
I_{2,1} &\le c\,\|\mathbf{u}^m\|_\alpha\|\mathbf{e}^{m-1}\|_2^\theta\|\mathbf{D}\mathbf{e}^{m-1}\|_p^{1-\theta}\|\nabla\Pi_h^{\text{div}}\mathbf{e}^m\|_p \\
&\le c\,\|\mathbf{e}^{m-1}\|_2^\theta\|\mathbf{D}\mathbf{e}^{m-1}\|_p^{1-\theta}\|\mathbf{D}\mathbf{e}^m\|_p.
\end{aligned}$$

For the term $I_{2,2}$ we have, by Hölder inequality

$$I_{2,2} \le \frac{1}{2}\|\nabla\Pi_h^{\text{div}}\mathbf{u}^m\|_r\|\mathbf{e}^{m-1}\|_{s_2}\|\Pi_h^{\text{div}}\mathbf{e}^m\|_{\frac{3p}{3-p}},$$

for some $1 < r < 6(p-1)$ and $s_2 = \frac{rp^*}{rp^* - r - p^*}$. A straightforward computation shows that for any $\frac{3}{2} < p \le 2$ one can choose r close enough to $6(p-1)$ in such a way that $s_2 < \frac{p}{p-1} < s_1$. Hence, by using again the properties of the interpolation operator (cf. Prop. 2.3, Rem. 2.5), since by (2.19) we have that $\|\nabla\mathbf{u}^m\|_r \in l^\infty(I_M)$ for all $r < 6(p-1)$, by Hölder and Korn's inequality, and by the embedding $L^{s_1}(\Omega) \subset L^{s_2}(\Omega)$, we get

$$I_{2,2} \le c\,\|\mathbf{e}^{m-1}\|_{s_1}\|\mathbf{D}\mathbf{e}^m\|_p.$$

Thus the right-hand can be estimated as $I_{2,1}$, which completes the proof. ∎

Now we need to estimate the term involving the pressure which can be handled by using the same approach as in [2]. Note that the regularity for the gradient of the pressure represents an outstanding open problem, with only partial results. In fact, at present, for the time evolution (either continuous or discrete) there are only results which exclude the degenerate case $\delta = 0$, see [3, 4]. Let us now show how the last term in (3.1) is estimated only in terms of the external force and of the velocity, by using once again the equations, as done in [2].

Lemma 3.4. *For each $\varepsilon > 0$ there exists $c_\varepsilon > 0$, not depending on h and δ, such that*

$$\begin{aligned}
&\big|\langle \operatorname{div}\Pi_h^{\text{div}}\mathbf{e}^m, \pi^m - \pi_h^m\rangle_h\big| \\
&\quad\le c\sum_K\int_K \big(\varphi_{|\mathbf{D}\mathbf{u}^m|}\big)^*\big(h|\mathbf{f}(t_m)| + h|d_t\mathbf{u}^m| + h|\mathbf{u}^{m-1}|\,|\nabla\mathbf{u}^m|\big)\,dx \\
&\qquad + c\sum_K\int_{S_K}|\mathbf{F}(\mathbf{D}\mathbf{u}^m) - \langle\mathbf{F}(\mathbf{D}\mathbf{u}^m)\rangle_{S_K}|^2\,dx \\
&\qquad + \varepsilon\Big(\|\mathbf{F}(\mathbf{D}\mathbf{u}^m) - \mathbf{F}(\mathbf{D}\Pi_h^{\text{div}}\mathbf{u}^m)\|_2^2 + \|\mathbf{F}(\mathbf{D}\mathbf{u}^m) - \mathbf{F}(\mathbf{D}\mathbf{u}_h^m)\|_2^2\Big).
\end{aligned}$$

Proof. We start by observing that for all $\boldsymbol{\xi}_h \in V_h(0)$ we have

$$\langle \operatorname{div} \boldsymbol{\xi}_h, \pi^m - \pi_h^m \rangle_h = \langle \operatorname{div} \boldsymbol{\xi}_h, \pi^m - \eta_h^m \rangle_h \qquad \forall\, \eta_h^m \in Y_h.$$

Then, if we use (in the same way as in [2, Lemma 3.1]) the divergence-preserving projection operator Π_h^{div}, we can estimate the term involving the pressure in the error equation as follows: For each $\eta_h^m \in Y_h$ it holds

$$\begin{aligned}
&|\langle \operatorname{div} \Pi_h^{\text{div}} \mathbf{e}^m, \pi^m - \pi_h^m \rangle_h| = |\langle \operatorname{div}(\Pi_h^{\text{div}} \mathbf{u}^m - \mathbf{u}_h^m), \pi^m - \eta_h^m \rangle_h| \\
&\quad \le \int_\Omega |\mathbf{D}\Pi_h^{\text{div}} \mathbf{u}^m - \mathbf{D}\mathbf{u}^m + \mathbf{D}\mathbf{u}^m - \mathbf{D}\mathbf{u}_h^m|\, |\pi^m - \eta_h^m|\, dx \\
&\quad \le \varepsilon \int_\Omega \varphi_{|\mathbf{D}\mathbf{u}^m|}(|\mathbf{D}\Pi_h^{\text{div}} \mathbf{u}^m - \mathbf{D}\mathbf{u}^m|) + \varphi_{|\mathbf{D}\mathbf{u}^m|}(|\mathbf{D}\mathbf{u}^m - \mathbf{D}\mathbf{u}_h^m|)\, dx \\
&\qquad\quad + c_\varepsilon \int_\Omega (\varphi_{|\mathbf{D}\mathbf{u}^m|})^*(|\pi^m - \eta_h^m|)\, dx \\
&\quad \le \varepsilon\, c \Big(\|\mathbf{F}(\mathbf{D}\mathbf{u}^m) - \mathbf{F}(\mathbf{D}\Pi_h^{\text{div}} \mathbf{u}^m)\|_2^2 + \|\mathbf{F}(\mathbf{D}\mathbf{u}^m) - \mathbf{F}(\mathbf{D}\mathbf{u}_h^m)\|_2^2 \Big) \\
&\qquad\quad + c_\varepsilon \int_\Omega (\varphi_{|\mathbf{D}\mathbf{u}^m|})^*(|\pi^m - \eta_h^m|)\, dx.
\end{aligned}$$

In particular we can choose $\eta_h^m = \Pi_h^Y \pi^m$. By using also Assumption 2.2 the latter term is estimated by using the same techniques as in [2, Lemma 6.4] as follows:

$$\begin{aligned}
\int_K (\varphi_{|\mathbf{D}\mathbf{v}|})^*(|\pi^m - \Pi_h^Y \pi^m|)\, dx &\le c \int_K \big(\varphi_{|\mathbf{D}\mathbf{v}|}\big)^*(h|\mathbf{f}(t_m)| + h|d_t \mathbf{u}^m| + h|\mathbf{u}^{m-1}|\, |\nabla \mathbf{u}^m|)\, dx \\
&\quad + c \int_{S_K} |\mathbf{F}(\mathbf{D}\mathbf{v}) - \langle \mathbf{F}(\mathbf{D}\mathbf{v}) \rangle_{S_K}|^2\, dx.
\end{aligned}$$

Finally, summing over $K \in \mathcal{T}_h$ we get the assertion. ☐

By collecting the above results we can now prove the main result of the paper.

Proof. [*of Theorem 2.31*] By gathering the results from Lemmas 3.1-3.4 we get the following discrete inequality: exists $c > 0$, independent of δ and h, and $\theta \in (0, 1)$ such that

$$\begin{aligned}
d_t \|\mathbf{e}^m\|_2^2 + k\|d_t \mathbf{e}^m\|_2^2 &+ \|\mathbf{F}(\mathbf{D}\mathbf{u}^m) - \mathbf{F}(\mathbf{D}\mathbf{u}_h^m)\|_2^2 + (\delta + \|\mathbf{D}\mathbf{e}^m\|_p)^{p-2} \|\mathbf{D}\mathbf{e}^m\|_p^2 \\
&\le c \Big[\frac{1}{k} \|\mathbf{R}_h^m\|_2^2 + \|\mathbf{F}(\mathbf{D}\mathbf{u}^m) - \mathbf{F}(\mathbf{D}\Pi_h^{\text{div}} \mathbf{u}^m)\|_2^2 + \|\nabla \mathbf{R}_h^m\|_{\frac{3p}{3+1}} \|\mathbf{D}\mathbf{e}^m\|_p \\
&\quad + \|\mathbf{e}^{m-1}\|_2^\theta \|\mathbf{D}\mathbf{e}^{m-1}\|_p^{1-\theta} \|\mathbf{D}\mathbf{e}^m\|_p \\
&\quad + \sum_K \int_K \big(\varphi_{|\mathbf{D}\mathbf{v}|}\big)^*(h|\mathbf{f}(t_m)| + h|d_t \mathbf{u}^m| + h|\mathbf{u}^{m-1}|\, |\nabla \mathbf{u}^m|)\, dx \\
&\quad + \sum_K \int_{S_K} |\mathbf{F}(\mathbf{D}\mathbf{u}^m) - \langle \mathbf{F}(\mathbf{D}\mathbf{u}^m) \rangle_{S_K}|^2\, dx \Big],
\end{aligned}$$

and by using a Sobolev embedding we can also obtain the following bound

$$\|\mathbf{e}^{m-1}\|_2^\theta \|\mathbf{D}\mathbf{e}^{m-1}\|_p^{1-\theta} \|\mathbf{D}\mathbf{e}^m\|_p \le c\, \|\mathbf{D}\mathbf{e}^{m-1}\|_p \|\mathbf{D}\mathbf{e}^m\|_p.$$

With this observation and by setting

$$
\begin{aligned}
a_m &:= \|\mathbf{e}^m\|_2,\\
b_m &:= \|\mathbf{D}\mathbf{e}^m\|_p,\\
r_m &:= \|\nabla \mathbf{R}_h^m\|_{\frac{3p}{p+1}},\\
s_m^2 &:= \|\mathbf{F}(\mathbf{D}\mathbf{u}^m) - \mathbf{F}(\mathbf{D}\Pi_h^{\mathrm{div}}\mathbf{u}^m)\|_2^2 + \sum_K \int_{S_K} \left|\mathbf{F}(\mathbf{D}\mathbf{u}^m) - \langle \mathbf{F}(\mathbf{D}\mathbf{u}^m)\rangle_{S_K}\right|^2 dx\\
&\qquad + \sum_K \int_K \big(\varphi_{|\mathbf{D}\mathbf{v}|}\big)^*(h|\mathbf{f}(t_m)| + h|d_t\mathbf{u}^m| + h|\mathbf{u}^{m-1}|\,|\nabla\mathbf{u}^m|)\,dx + \frac{\|\mathbf{R}_h^m\|_2^2}{k},
\end{aligned}
$$

we have that the two inequalities (2.22), (2.23) are satisfied. Hence, in order to apply Lemma 2.3, we need just to verify the hypotheses on the initial values a_0, b_0 and on r_m and s_m.

To this end, first we observe that $\mathbf{e}^0 = \mathbf{u}_0 - \Pi_h^{\mathrm{div}}\mathbf{u}_0$. By using the assumption $\mathbf{u}_0 \in W^{2,2}_{\mathrm{div}}$, by the properties of the interpolation operator Π_h^{div}, and due to $p \le 2$ we obtain:

$$
\|\mathbf{e}^0\|_2 \le c\,h^2 \qquad \text{and} \qquad \|\mathbf{D}\mathbf{e}^0\|_p \le c\,h \tag{3.3}
$$

We now check the hypotheses needed on r_m and we observe, that if $\mathbf{u}^m \in W^{2,\frac{3p}{p+1}}(\Omega)$, then

$$
\|\nabla \mathbf{R}_h^m\|_{\frac{3p}{p+1}} \le c\,h\,\|\nabla^2\mathbf{u}^m\|_{\frac{3p}{p+1}},
$$

by the properties of the interpolation operator (cf. Prop. 2.3, Rem. 2.5). Hence, under the assumptions of regularity of $\mathbf{u}^m$, we also obtain that

$$
k\sum_{m=0}^{M} \|\nabla \mathbf{R}_h^m\|^2_{\frac{3p}{p+1}} \le c\,h^2,
$$

for some constant c independent of δ and h.

Let us now consider s_m and we recall that if $\mathbf{F}(\mathbf{D}\mathbf{u}^m) \in W^{1,2}(\Omega)$, then uniformly with respect to $K \in \mathcal{T}_h$ (cf. [2, Thm 3.7,Thm 5.1])

$$
\begin{aligned}
\|\mathbf{F}(\mathbf{D}\mathbf{u}^m) - \mathbf{F}(\mathbf{D}\Pi_h^{\mathrm{div}}\mathbf{u}^m)\|_2^2 &\le \sum_K \int_{S_K} \left|\mathbf{F}(\mathbf{D}\mathbf{u}^m) - \langle \mathbf{F}(\mathbf{D}\mathbf{u}^m)\rangle_{S_K}\right|^2 dx\\
&\le c\,h^2\|\nabla \mathbf{F}(\mathbf{D}\mathbf{u}^m)\|_2^2.
\end{aligned}
$$

We now estimate the third term in the definition of s_m^2 by defining the following non-negative sequence $\{g^m\}_m$

$$
g^m := |\mathbf{f}(t_m)| + |d_t\mathbf{u}^m| + |\mathbf{u}^{m-1}|\,|\nabla\mathbf{u}^m|.
$$

By Young's inequality and by using the following inequality for φ defined in (2.9)

$$
(\varphi_a)^*(\kappa\,t) \le c\,\kappa^2\,(\varphi_a)^*(t)
$$

valid for $\kappa \in [0, \kappa_0]$ and $p \le 2$ with a constant c independent of δ, a, and t (cf. [2]), we have

$$\sum_K \int_{S_K} \big(\varphi_{|\mathbf{D}\mathbf{u}^m|}\big)^*(h\, g^m)\, dx \le c\, h^2 \sum_K \int_{S_K} \big(\varphi_{|\mathbf{D}\mathbf{u}^m|}\big)^*(g^m)\, dx$$
$$\le c\, h^2 \sum_K \int_{S_K} \varphi(|\mathbf{D}\mathbf{u}^m|) + \varphi^*(g^m)\, dx.$$

Pointing out that

$$k \sum_{m=0}^{M} \int_{\Omega} \varphi^*(g^m)\, dx \le k \sum_{m=0}^{M} \|g^m\|_{p'}^{p'} + (\delta|\Omega|)^{p'},$$

we need just to check that $g^m \in l^{p'}(I^M)$. This follows by interpolation from Thm. 2.2, and especially from the observation in (2.20)

To conclude we need also to estimate the term $k^{-1}\|\mathbf{R}_h^m\|_2^2$. There is another (we also have one in the discrete Gronwall Lemma 2.3) h-k coupling that enters the proof at this point. In fact, by Sobolev embedding, the standard properties of interpolation operators in Sobolev space (see e.g. [6, Thm. 3.1.6]), and the assumptions on $\mathcal{T}_h$ we get $\|\mathbf{R}_h^m\|_2 \le c\, h^{\frac{5p-2}{2p}} \|\nabla^2 \mathbf{u}^m\|_{\frac{3p}{p+1}}$. Then, by using the regularity on $\mathbf{u}^m$ from Thm. 2.2 we obtain

$$k \sum_{m=0}^{M} \frac{\|\mathbf{R}_h^m\|_2^2}{k} \le \frac{h^{\frac{5p-2}{p}}}{k} k \sum_{m=0}^{M} \|\nabla^2 \mathbf{u}^m\|_{\frac{3p}{p+1}}^2 \le c \frac{h^{\frac{5p-2}{p}}}{k},$$

and if $h^{\frac{3p-2}{p}} \le c\, k$, then

$$k \sum_{m=0}^{M} \frac{\|\mathbf{R}_h^m\|_2^2}{k} \le c\, h^2$$

This coupling between k and h derives from the natural regularity of the problem, which is at the moment at disposal under rather general assumptions on the data. We believe that this condition, appearing also in simpler parabolic problems with p-structure [7], is only of technical character.

Then, by collecting all the previous estimate, we obtain that all the hypotheses of Lemma 2.3 are satisfied, hence we end the proof. ◻

Acknowledgment. L.Berselli would like to thank the SFB/TR 71 "Geometric Partial Differential Equations" for the hospitality during his stays In Freiburg. L.Diening and M.Růžička have been supported by by the project C2 of the SFB/TR 71 "Geometric Partial Differential Equations".

REFERENCES

[1] H. Beirão da Veiga, *On the global regularity of shear thinning flows in smooth domains*, J. Math. Anal. Appl., 349 (2009), pp. 335–360.

[2] L. Belenki, L. C. Berselli, L. Diening, and M. Růžička, *On the Finite Element approximation of p-Stokes systems*, SIAM J. Numer. Anal., 50 (2012), pp. 373–397.

[3] L. C. Berselli, L. Diening, and M. Růžička, *Optimal estimates for a semi implicit Euler scheme for incompressible fluids with shear dependent viscosities*, SIAM J. Numer. Anal., 47 (2009), pp. 2177–2202.

[4] L. C. Berselli, L. Diening, and M. Růžička, *Existence of strong solutions for incompressible fluids with shear dependent viscosities*, J. Math. Fluid Mech., 12 (2010), pp. 101–132.

[5] D. Bothe and J. Prüss, *L^p-theory for a class of Non–Newtonian fluids*, SIAM J. Math. Anal., 39 (2007), pp. 379–421.

[6] P. G. Ciarlet, *The finite element method for elliptic problems*, North-Holland Publishing Co., Amsterdam, 1978. Studies in Mathematics and its Applications, Vol. 4.

[7] L. Diening, C. Ebmeyer, and M. Růžička, *Optimal convergence for the implicit space-time discretization of parabolic systems with p-structure*, SIAM J. Numer. Anal., 45 (2007), pp. 457–472.

[8] L. Diening and F. Ettwein, *Fractional estimates for non–differentiable elliptic systems with general growth*, Forum Mathematicum, 20 (2008), pp. 523–556.

[9] L. Diening and C. Kreuzer, *Linear convergence of an adaptive finite element method for the p-Laplacian equation*, SIAM J. Numer. Anal., 46 (2008), pp. 614–638.

[10] A. Prohl and M. Růžička, *On fully implicit space-time discretization for motions of incompressible fluids with shear dependent viscosities: The case $p \leq 2$*, SIAM J. Num. Anal., 39 (2001), pp. 214–249.

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