Mathematical Modelling and Numerical Analysis Will be set by the publisher
Modélisation Mathématique et Analyse Numérique

A HYBRID HIGH-ORDER METHOD FOR CREEPING FLOWS OF NON-NEWTONIAN FLUIDS

Michele Botti[1], Daniel Castanon Quiroz[2], Daniele A. Di Pietro[2] and André Harnist[2]

Abstract. In this paper, we design and analyze a Hybrid High-Order discretization method for the steady motion of non-Newtonian, incompressible fluids in the Stokes approximation of small velocities. The proposed method has several appealing features including the support of general meshes and high-order, unconditional inf-sup stability, and orders of convergence that match those obtained for scalar Leray–Lions problems. A complete well-posedness and convergence analysis of the method is carried out under new, general assumptions on the strain rate-shear stress law, which encompass several common examples such as the power-law and Carreau–Yasuda models. Numerical examples complete the exposition.

Résumé. Dans cet article, nous développons et analysons une méthode de discrétisation Hybride d'Ordre Élevé pour l'écoulement stationnaire de fluides incompressibles non-newtoniens régis par les équations de Stokes pour des petites vitesses. La méthode proposée présente plusieurs caractéristiques intéressantes, notamment la prise en charge de maillages généraux et d'une stabilité inf-sup d'ordre élevé et inconditionnelle, ainsi que des ordres de convergence qui correspondent à ceux obtenus pour les problèmes scalaires de Leray–Lions. Une analyse complète de l'adéquation et de la convergence de la méthode est effectuée sous de nouvelles hypothèses générales sur la loi de vitesse de déformation-contrainte de cisaillement, qui englobe plusieurs exemples courants tels que la loi en puissance et les modèles de Carreau–Yasuda. Des exemples numériques complètent l'exposition.

1991 Mathematics Subject Classification. 65N08, 65N30, 65N12, 35Q30, 76D05.

The dates will be set by the publisher.

Introduction

In this paper, we design and analyze a Hybrid High-Order (HHO) discretization method for the steady motion of a non-Newtonian, incompressible fluid in the Stokes approximation of small velocities. Notable applications include ice sheet dynamics [32], mantle convection [44], chemical engineering [34], and biological fluids rheology [28,37]. We focus on fluids with shear-rate-dependent viscosity, whose behavior is characterized by a nonlinear strain rate-shear stress function. Physical interpretations and discussions of non-Newtonian fluid

Keywords and phrases: Hybrid High-Order methods, non-Newtonian fluids, power-law, Carreau–Yasuda law, discrete Korn inequality

[1] MOX, Department of Mathematics, Politecnico di Milano, Milano, Italy

[2] IMAG, Univ Montpellier, CNRS, Montpellier, France

models can be found, e.g., in [8, 40]. Typical examples that are frequently used in the applications include the power-law and Carreau–Yasuda model, covered by the present analysis.

The earliest investigations of fluids with shear-dependent viscosity date back to the pioneering work of Ladyzhenskaya [36]. For a detailed mathematical study of the well-posedness and regularity of the continuous problem, see also [3,7,24,39,41] and references therein. Early results on the numerical analysis of non-Newtonian fluid flow problems were given in [2, 30, 42]. Later, these results were improved in [6] and [31] by proving error estimates that are optimal for fluids with shear thinning behavior (described by a power-law exponent $r \leq 2$). In [6], the authors considered a conforming inf-sup stable finite element discretization, while in [31] a low-order scheme with local projection stabilization was proposed. In both works, the use of Orlicz functions is instrumental to unify the treatment of the shear thinning and shear thickening cases (also called pseudoplastic and dilatant, respectively; cf. Example 1.5). More recently, a finite element method based on a four-field formulation of the nonlinear Stokes equations has been analyzed in [43]. Other notable contributions on the numerical approximation of generalized Stokes problems include [25, 32, 33, 35].

The main issues to be accounted for in the numerical solution of non-Newtonian fluid flow problems are the presence of local features emerging from the nonlinear strain rate-shear stress relation, the incompressibility condition leading to indefinite systems, the roughly varying model coefficients, and, possibly, complex geometries requiring unstructured and highly-adapted meshes. The HHO method provides several advantages to deal with the complex nature of the problem, such as the support of general polygonal or polyhedral meshes, the possibility to select the approximation order, and unconditional inf-sup stability. Moreover, HHO schemes can be efficiently implemented thanks to the possibility of statically condensing a large subset of the unknowns for linearized versions of the problem encountered, e.g., when solving the nonlinear system by the Newton method. Hybrid High-Order methods have been successfully applied to the simulation of incompressible flows of Newtonian fluids governed by the Stokes [1] and Navier–Stokes equations [12, 23], possibly driven by large irrotational volumetric forces [15, 22]. Works related to the problem of creeping flows of non-Newtonian fluids are [13] and [17, 18], respectively dealing with nonlinear elasticity and Leray–Lions problems. Going from nonlinear coercive elliptic equations to the nonlinear Stokes system involves additional difficulties arising from the pressure and the divergence constraint. Finally, we mention that HHO methods are members of a wider family of polytopal methods that also includes, e.g., Virtual Element methods (cf., e.g., [4,5] for their application to Newtonian incompressible flows) and can fit within general frameworks for the approximation of nonlinear problems such as the one provided by the Gradient Discretisation Method (see [21, 26]).

The HHO discretization presented in this paper hinges on discontinuous polynomial unknowns on the mesh and on its skeleton, from which discrete differential operators are reconstructed. These operators are used to formulate discrete counterparts of the viscous and pressure-velocity coupling terms. For the former, stability is ensured by a cleverly designed stabilization contribution involving the penalization of boundary differences. We carry out a complete analysis of the proposed method. In particular, under general assumptions on the strain rate-shear stress function, we derive error estimates for the velocity and pressure approximations. The energy-norm error estimate for the velocity given in Theorem 3.7 yields the same convergence orders established in [18, Theorem 3.2] for the scalar Leray–Lions elliptic problem. A key tool in our analysis is provided by Lemma 5.2, in which we prove a generalization of the discrete Korn inequality of [12, Lemma 1] to the non-Hilbertian case. The other main contributions are a novel formulation of the requirements on the strain rate-shear stress function allowing a unified treatment of pseudoplastic and dilatant fluids and the identification of a set of general assumptions on the nonlinear stabilization function ensuring the desired consistency properties along with the well-posedness of the discrete problem.

The rest of the paper is organized as follows. In Section 1 we introduce the strong and weak formulations of the nonlinear Stokes problem and present the assumptions on the strain rate-shear stress function. The discrete setting is established in Section 2, including the definition of the discrete spaces for the velocity and the pressure. The HHO scheme along with the main theoretical results are stated in Section 3, and a numerical validation is provided in Section 4. In Section 5 we prove the discrete counterpart of the Korn inequality needed in the analysis of the method. Section 6 contains the proof of the main results (well-posedness and

error estimates). Finally, in Appendix A we provide a sufficient condition for the strain rate-shear stress law to fulfil the assumptions presented in Section 1. The paper is structured so as to offer two levels of reading. In particular, the reader mainly interested in the formulation of the method and its numerical performance can focus on Section 1–4. The remaining sections cover technical aspects of the analysis, and can be skipped at first reading.

1. Continuous setting

Let $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, denote a bounded, connected, polyhedral open set with Lipschitz boundary $\partial\Omega$. We consider a possibly non-Newtonian fluid occupying Ω and subjected to a volumetric force field $\boldsymbol{f} : \Omega \to \mathbb{R}^d$. Its flow is governed by the generalized Stokes problem, which consists in finding the velocity field $\boldsymbol{u} : \Omega \to \mathbb{R}^d$ and the pressure field $p : \Omega \to \mathbb{R}$ such that

$$-\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}(\cdot, \boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{u}) + \boldsymbol{\nabla} p = \boldsymbol{f} \qquad \text{in } \Omega, \tag{1a}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0 \qquad \text{in } \Omega, \tag{1b}$$

$$\boldsymbol{u} = \boldsymbol{0} \qquad \text{on } \partial\Omega, \tag{1c}$$

$$\int_\Omega p(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} = 0, \tag{1d}$$

where $\boldsymbol{\nabla}\cdot$ denotes the divergence operator applied to vector or tensor fields, $\boldsymbol{\nabla}_{\mathrm{s}}$ is the symmetric part of the gradient operator $\boldsymbol{\nabla}$ applied to vector fields, and, denoting by $\mathbb{R}^{d\times d}_{\mathrm{s}}$ the set of square, symmetric, real-valued $d \times d$ matrices, $\boldsymbol{\sigma} : \Omega \times \mathbb{R}^{d\times d}_{\mathrm{s}} \to \mathbb{R}^{d\times d}_{\mathrm{s}}$ is the strain rate-shear stress law. In what follows, we formulate assumptions on $\boldsymbol{\sigma}$ that encompass common models for non-Newtonian fluids and state a weak formulation for problem (1) that will be used as a starting point for its discretization.

1.1. Strain rate-shear stress law

We define the Frobenius inner product such that, for all $\boldsymbol{\tau} = (\tau_{ij})_{1\le i,j\le d}$ and $\boldsymbol{\eta} = (\eta_{ij})_{1\le i,j\le d}$ in $\mathbb{R}^{d\times d}$, $\boldsymbol{\tau} : \boldsymbol{\eta} \coloneqq \sum_{i,j=1}^d \tau_{ij}\eta_{ij}$, and we denote by $|\boldsymbol{\tau}|_{d\times d} \coloneqq \sqrt{\boldsymbol{\tau} : \boldsymbol{\tau}}$ the corresponding norm.

Assumption 1.1 (Strain rate-shear stress law)**.** Let a real number $r \in (1,\infty)$ be fixed, denote by $r' \coloneqq \frac{r}{r-1} \in (1,\infty)$ the conjugate exponent of r, and define the singular exponent of r by

$$\tilde{r} \coloneqq \min(r, 2) \in (1, 2]. \tag{2}$$

The strain rate-shear stress law satisfies

$$\boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{0}) = \boldsymbol{0} \text{ for almost every } \boldsymbol{x} \in \Omega, \tag{3a}$$

$$\boldsymbol{\sigma} : \Omega \times \mathbb{R}^{d\times d}_{\mathrm{s}} \to \mathbb{R}^{d\times d}_{\mathrm{s}} \text{ is measurable.} \tag{3b}$$

Moreover, there exist real numbers $\sigma_{\mathrm{de}} \in [0,\infty)$ and $\sigma_{\mathrm{hc}}, \sigma_{\mathrm{sm}} \in (0,\infty)$ such that, for all $\boldsymbol{\tau}, \boldsymbol{\eta} \in \mathbb{R}^{d\times d}_{\mathrm{s}}$ and almost every $\boldsymbol{x} \in \Omega$, we have the Hölder continuity property

$$|\boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{\tau}) - \boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{\eta})|_{d\times d} \le \sigma_{\mathrm{hc}} \left(\sigma_{\mathrm{de}}^r + |\boldsymbol{\tau}|_{d\times d}^r + |\boldsymbol{\eta}|_{d\times d}^r\right)^{\frac{r-\tilde{r}}{r}} |\boldsymbol{\tau} - \boldsymbol{\eta}|_{d\times d}^{\tilde{r}-1}, \tag{3c}$$

and the strong monotonicity property

$$(\boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{\tau}) - \boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{\eta})) : (\boldsymbol{\tau} - \boldsymbol{\eta}) \left(\sigma_{\mathrm{de}}^r + |\boldsymbol{\tau}|_{d\times d}^r + |\boldsymbol{\eta}|_{d\times d}^r\right)^{\frac{2-\tilde{r}}{r}} \ge \sigma_{\mathrm{sm}} |\boldsymbol{\tau} - \boldsymbol{\eta}|_{d\times d}^{r+2-\tilde{r}}. \tag{3d}$$

Some remarks are in order.

Remark 1.2 (Residual shear stress)**.** Assumption (3a) can be relaxed by taking $\boldsymbol{\sigma}(\cdot,\mathbf{0}) \in L^{r'}(\Omega,\mathbb{R}^{d\times d}_{\mathrm{s}})$. This modification requires only minor changes in the analysis, not detailed for the sake of conciseness.

Remark 1.3 (Singular exponent)**.** Inequalities (3c)–(3d) can be proved starting from the following assumptions, which correspond to the conditions (74) below characterizing an r-power-framed function: For all $\boldsymbol{\tau},\boldsymbol{\eta} \in \mathbb{R}^{d\times d}_{\mathrm{s}}$ with $\boldsymbol{\tau} \neq \boldsymbol{\eta}$ and almost every $\boldsymbol{x} \in \Omega$,

$$|\boldsymbol{\sigma}(\boldsymbol{x},\boldsymbol{\tau}) - \boldsymbol{\sigma}(\boldsymbol{x},\boldsymbol{\eta})|_{d\times d} \leq \sigma_{\mathrm{hc}} \left(\sigma_{\mathrm{de}}^r + |\boldsymbol{\tau}|_{d\times d}^r + |\boldsymbol{\eta}|_{d\times d}^r\right)^{\frac{r-2}{r}} |\boldsymbol{\tau} - \boldsymbol{\eta}|_{d\times d},$$
$$(\boldsymbol{\sigma}(\boldsymbol{x},\boldsymbol{\tau}) - \boldsymbol{\sigma}(\boldsymbol{x},\boldsymbol{\eta})) : (\boldsymbol{\tau} - \boldsymbol{\eta}) \geq \sigma_{\mathrm{sm}} \left(\sigma_{\mathrm{de}}^r + |\boldsymbol{\tau}|_{d\times d}^r + |\boldsymbol{\eta}|_{d\times d}^r\right)^{\frac{r-2}{r}} |\boldsymbol{\tau} - \boldsymbol{\eta}|_{d\times d}^2.$$

These relations are reminiscent of the ones used in [18] in the context of scalar Leray–Lions problems. The advantage of assumptions (3c)-(3d), expressed in terms of the singular index $\tilde{r}$, is that they enable a unified treatment of the cases $r < 2$ and $r \geq 2$ in the proofs of Lemma 6.3, Theorem 3.6, Lemma 6.5, and Theorem 3.7 below.

Remark 1.4 (Relations between the Hölder and monotonicity constants)**.** Inequalities (3c) and (3d) give

$$\sigma_{\mathrm{sm}} \leq \sigma_{\mathrm{hc}}. \tag{4}$$

Indeed, let $\boldsymbol{\tau} \in \mathbb{R}^{d\times d}_{\mathrm{s}}$ be such that $|\boldsymbol{\tau}|_{d\times d} > 0$. Using the strong monotonicity (3d) (with $\boldsymbol{\eta} = \mathbf{0}$), the Cauchy–Schwarz inequality, and the Hölder continuity (3c) (again with $\boldsymbol{\eta} = \mathbf{0}$), we infer that

$$\sigma_{\mathrm{sm}} \left(\sigma_{\mathrm{de}}^r + |\boldsymbol{\tau}|_{d\times d}^r\right)^{\frac{\tilde{r}-2}{r}} |\boldsymbol{\tau}|_{d\times d}^{r+2-\tilde{r}} \leq \boldsymbol{\sigma}(\cdot,\boldsymbol{\tau}) : \boldsymbol{\tau} \leq |\boldsymbol{\sigma}(\cdot,\boldsymbol{\tau})|_{d\times d}|\boldsymbol{\tau}|_{d\times d} \leq \sigma_{\mathrm{hc}} \left(\sigma_{\mathrm{de}}^r + |\boldsymbol{\tau}|_{d\times d}^r\right)^{\frac{r-\tilde{r}}{r}} |\boldsymbol{\tau}|_{d\times d}^{\tilde{r}}$$

almost everywhere in Ω. Hence, $\frac{\sigma_{\mathrm{sm}}}{\sigma_{\mathrm{hc}}} \leq \left(\frac{\sigma_{\mathrm{de}}^r + |\boldsymbol{\tau}|_{d\times d}^r}{|\boldsymbol{\tau}|_{d\times d}^r}\right)^{\frac{|r-2|}{r}}$. Letting $|\boldsymbol{\tau}|_{d\times d} \to \infty$ gives (4).

Example 1.5 (Carreau–Yasuda fluids)**.** (μ,δ,a,r)-Carreau–Yasuda fluids, introduced in [46] and later generalized in [31, Eq. (1.2)], are fluids for which it holds, for almost every $\boldsymbol{x} \in \Omega$ and all $\boldsymbol{\tau} \in \mathbb{R}^{d\times d}_{\mathrm{s}}$,

$$\boldsymbol{\sigma}(\boldsymbol{x},\boldsymbol{\tau}) = \mu(\boldsymbol{x}) \left(\delta^{a(\boldsymbol{x})} + |\boldsymbol{\tau}|_{d\times d}^{a(\boldsymbol{x})}\right)^{\frac{r-2}{a(\boldsymbol{x})}} \boldsymbol{\tau}, \tag{5}$$

where $\mu : \Omega \to [\mu_-,\mu_+]$ is a measurable function with $\mu_-,\mu_+ \in (0,\infty)$ corresponding to the local flow consistency index, $\delta \in [0,\infty)$ is the degeneracy parameter, $a : \Omega \to [a_-,a_+]$ is a measurable function with $a_-,a_+ \in (0,\infty)$ expressing the local transition flow behavior index, and $r \in (1,\infty)$ is the flow behavior index. The Carreau–Yasuda law is a generalization of the Carreau law (corresponding to $a_- = a_+ = 2$) that takes into account the different local levels of flow behavior in the fluid. The degenerate case $\delta = 0$ corresponds to the power-law model. Non-Newtonian fluids described by constitutive laws with a (μ,δ,a,r)-structure exhibit a different behavior according to the value of r. If $r > 2$, then the fluid shows shear thickening behavior and is called *dilatant*. Examples of dilatant fluids are wet sand and oobleck. The case $r < 2$, on the other hand, corresponds to *pseudoplastic* fluids having shear thinning behavior, such as blood. Finally, if $r = 2$, then the fluid is Newtonian and (1) becomes the classical (linear) Stokes problem. We show in Appendix A that the strain rate-shear stress law (5) is an r-power-framed function with $\sigma_{\mathrm{de}} = \delta$,

$$\sigma_{\mathrm{hc}} = \begin{cases} \frac{\mu_+}{r-1} 2^{\left[-\left(\frac{1}{a_+}-\frac{1}{r}\right)^{\ominus}-1\right](r-2)+\frac{1}{r}} & \text{if } r < 2, \\ \mu_+(r-1)2^{\left(\frac{1}{a_-}-\frac{1}{r}\right)^{\oplus}(r-2)} & \text{if } r \geq 2, \end{cases} \quad \text{and} \quad \sigma_{\mathrm{sm}} = \begin{cases} \mu_-(r-1)2^{\left(\frac{1}{a_-}-\frac{1}{r}\right)^{\oplus}(r-2)} & \text{if } r \leq 2, \\ \frac{\mu_-}{r-1} 2^{\left[-\left(\frac{1}{a_+}-\frac{1}{r}\right)^{\ominus}-1\right](r-2)-1} & \text{if } r > 2, \end{cases}$$

where $\xi^{\oplus} \coloneqq \max(0,\xi)$ and $\xi^{\ominus} \coloneqq -\min(0,\xi)$ denote, respectively, the positive and negative parts of a real number ξ. As a consequence, it matches Assumption 1.1.

1.2. **Weak formulation**

From this point on, we omit both the integration variable and the measure from integrals, as they can be in all cases inferred from the context. We define the following velocity and pressure spaces embedding, respectively, the homogeneous boundary condition and the zero-average constraint:

$$\boldsymbol{U} \coloneqq \left\{\boldsymbol{v} \in W^{1,r}(\Omega,\mathbb{R}^d) \ : \ \boldsymbol{v}_{|\partial\Omega} = \boldsymbol{0}\right\}, \qquad P \coloneqq L_0^{r'}(\Omega,\mathbb{R}) \coloneqq \left\{q \in L^{r'}(\Omega,\mathbb{R}) \ : \ \int_\Omega q = 0\right\}.$$

Assuming $\boldsymbol{f} \in L^{r'}(\Omega,\mathbb{R}^d)$, the weak formulation of problem (1) reads: Find $(\boldsymbol{u},p) \in \boldsymbol{U}\times P$ such that

$$a(\boldsymbol{u},\boldsymbol{v}) + b(\boldsymbol{v},p) = \int_\Omega \boldsymbol{f}\cdot\boldsymbol{v} \qquad \forall \boldsymbol{v}\in\boldsymbol{U}, \tag{6a}$$

$$-b(\boldsymbol{u},q) = 0 \qquad \forall q \in P, \tag{6b}$$

where the function $a : \boldsymbol{U}\times\boldsymbol{U}\to\mathbb{R}$ and the bilinear form $b : \boldsymbol{U}\times L^{r'}(\Omega,\mathbb{R})\to\mathbb{R}$ are defined such that, for all $\boldsymbol{v},\boldsymbol{w}\in\boldsymbol{U}$ and all $q\in L^{r'}(\Omega,\mathbb{R})$,

$$a(\boldsymbol{w},\boldsymbol{v}) \coloneqq \int_\Omega \boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w}) : \boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{v}, \qquad b(\boldsymbol{v},q) \coloneqq -\int_\Omega (\boldsymbol{\nabla}\cdot\boldsymbol{v})q. \tag{7}$$

Remark 1.6 (Mass equation)**.** The test space in (6b) can be extended to $L^{r'}(\Omega,\mathbb{R})$ since, for all $\boldsymbol{v}\in\boldsymbol{U}$, the divergence theorem and the fact that $\boldsymbol{v}_{|\partial\Omega}=\boldsymbol{0}$ yield $b(\boldsymbol{v},1) = -\int_\Omega \boldsymbol{\nabla}\cdot\boldsymbol{v} = -\int_{\partial\Omega}\boldsymbol{v}\cdot\boldsymbol{n}_{\partial\Omega} = 0$, with $\boldsymbol{n}_{\partial\Omega}$ denoting the unit vector normal to $\partial\Omega$ and pointing out of Ω.

Remark 1.7 (Well-posedness and a priori estimates)**.** It can be checked that, under Assumption 1.1, the continuous problem (6) admits a unique solution $(\boldsymbol{u},p)\in\boldsymbol{U}\times P$; see, e.g., [31, Section 2.4], where slightly stronger assumptions are considered. For future use, we also note the following a priori bound on the velocity:

$$|\boldsymbol{u}|_{W^{1,r}(\Omega,\mathbb{R}^d)} \le \left(2^{\frac{2-\tilde r}{r}} C_{\mathrm{K}}\sigma_{\mathrm{sm}}^{-1}\|\boldsymbol{f}\|_{L^{r'}(\Omega,\mathbb{R}^d)}\right)^{\frac{1}{r-1}} + \left(2^{\frac{2-\tilde r}{r}} C_{\mathrm{K}}|\Omega|_d^{\frac{2-\tilde r}{r}}\sigma_{\mathrm{de}}^{2-\tilde r}\sigma_{\mathrm{sm}}^{-1}\|\boldsymbol{f}\|_{L^{r'}(\Omega,\mathbb{R}^d)}\right)^{\frac{1}{r+1-\tilde r}}, \tag{8}$$

where $C_{\mathrm{K}}>0$ comes from the Korn inequality given at (33) below. To prove (8), use the strong-monotonicity (3d) of $\boldsymbol{\sigma}$, sum (6a) written for $\boldsymbol{v}=\boldsymbol{u}$ to (6b) written for $q=p$, and use the Hölder inequality together with the Korn inequality (33) to write

$$\begin{aligned}\sigma_{\mathrm{sm}}\left(|\Omega|_d\sigma_{\mathrm{de}}^r + \|\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{u}\|_{L^r(\Omega,\mathbb{R}^{d\times d})}^r\right)^{\frac{\tilde r-2}{r}}\|\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{u}\|_{L^r(\Omega,\mathbb{R}^{d\times d})}^{r+2-\tilde r} &\le a(\boldsymbol{u},\boldsymbol{u})\\ &= \int_\Omega \boldsymbol{f}\cdot\boldsymbol{u} \le C_{\mathrm{K}}\|\boldsymbol{f}\|_{L^{r'}(\Omega,\mathbb{R}^d)}\|\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{u}\|_{L^r(\Omega,\mathbb{R}^{d\times d})},\end{aligned}$$

where $|\Omega|_d$ is the measure of Ω, that is,

$$\mathcal{N} \coloneqq \left(|\Omega|_d\sigma_{\mathrm{de}}^r + \|\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{u}\|_{L^r(\Omega,\mathbb{R}^{d\times d})}^r\right)^{\frac{\tilde r-2}{r}}\|\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{u}\|_{L^r(\Omega,\mathbb{R}^{d\times d})}^{r+1-\tilde r} \le C_{\mathrm{K}}\sigma_{\mathrm{sm}}^{-1}\|\boldsymbol{f}\|_{L^{r'}(\Omega,\mathbb{R}^d)}. \tag{9}$$

Observing that $\|\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{u}\|_{L^r(\Omega,\mathbb{R}^{d\times d})}^{r+1-\tilde r} \le 2^{\frac{2-\tilde r}{r}}\max\left(\|\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{u}\|_{L^r(\Omega,\mathbb{R}^{d\times d})}^r, |\Omega|_d\sigma_{\mathrm{de}}^r\right)^{\frac{2-\tilde r}{r}}\mathcal{N}$, we obtain, enumerating the cases for the maximum and summing the corresponding bounds,

$$\|\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{u}\|_{L^r(\Omega,\mathbb{R}^{d\times d})} \le (2^{\frac{2-\tilde r}{r}}\mathcal{N})^{\frac{1}{r-1}} + \left(2^{\frac{2-\tilde r}{r}}|\Omega|_d^{\frac{2-\tilde r}{r}}\sigma_{\mathrm{de}}^{2-\tilde r}\mathcal{N}\right)^{\frac{1}{r+1-\tilde r}}.$$

Combining this inequality with (9) gives (8).

2. Discrete setting

2.1. Mesh and notation for inequalities up to a multiplicative constant

We define a mesh as a couple $\mathcal{M}_h \coloneqq (\mathcal{T}_h, \mathcal{F}_h)$, where $\mathcal{T}_h$ is a finite collection of polyhedral elements T such that $h = \max_{T\in\mathcal{T}_h} h_T$ with h_T denoting the diameter of T, while $\mathcal{F}_h$ is a finite collection of planar faces F with diameter h_F. Notice that, here and in what follows, we use the three-dimensional nomenclature also when $d = 2$, i.e., we speak of polyhedra and faces rather than polygons and edges. It is assumed henceforth that the mesh $\mathcal{M}_h$ matches the geometrical requirements detailed in [19, Definition 1.7]. In order to have the boundedness property (14) for the interpolator, we additionally assume that the mesh elements are star-shaped with respect to every point of a ball of radius uniformly comparable to the element diameter; see [19, Lemma 7.12] for the Hilbertian case. Boundary faces lying on $\partial\Omega$ and internal faces contained in Ω are collected in the sets $\mathcal{F}_h^{\mathrm{b}}$ and $\mathcal{F}_h^{\mathrm{i}}$, respectively. For every mesh element $T \in \mathcal{T}_h$, we denote by $\mathcal{F}_T$ the subset of $\mathcal{F}_h$ containing the faces that lie on the boundary ∂T of T. For every face $F \in \mathcal{F}_h$, we denote by $\mathcal{T}_F$ the subset of $\mathcal{T}_h$ containing the one (if $F \in \mathcal{F}_h^{\mathrm{b}}$) or two (if $F \in \mathcal{F}_h^{\mathrm{i}}$) elements on whose boundary F lies. Finally, for each mesh element $T \in \mathcal{T}_h$ and face $F \in \mathcal{F}_T$, $\boldsymbol{n}_{TF}$ denotes the (constant) unit vector normal to F pointing out of T.

Our focus is on the h-convergence analysis, so we consider a sequence of refined meshes that is regular in the sense of [19, Definition 1.9] with regularity parameter uniformly bounded away from zero. The mesh regularity assumption implies, in particular, that the diameter of a mesh element and those of its faces are comparable uniformly in h and that the number of faces of one element is bounded above by an integer independent of h.

To avoid the proliferation of generic constants, we write henceforth $a \lesssim b$ (resp., $a \gtrsim b$) for the inequality $a \le Cb$ (resp., $a \ge Cb$) with real number $C > 0$ independent of h, of the constants $\sigma_{\mathrm{de}}, \sigma_{\mathrm{hc}}, \sigma_{\mathrm{sm}}$ in Assumption 1.1, and, for local inequalities, of the mesh element or face on which the inequality holds. We also write $a \simeq b$ to mean $a \lesssim b$ and $b \lesssim a$. The dependencies of the hidden constants are further specified when needed.

2.2. Projectors and broken spaces

Given $X \in \mathcal{T}_h \cup \mathcal{F}_h$ and $l \in \mathbb{N}$, we denote by $\mathbb{P}^l(X, \mathbb{R})$ the space spanned by the restriction to X of scalar-valued, d-variate polynomials of total degree $\le l$. The local L^2-orthogonal projector $\pi_X^l : L^1(X, \mathbb{R}) \to \mathbb{P}^l(X, \mathbb{R})$ is defined such that, for all $v \in L^1(X, \mathbb{R})$,

$$\int_X (\pi_X^l v - v)w = 0 \qquad \forall w \in \mathbb{P}^l(X, \mathbb{R}). \tag{10}$$

When applied to vector-valued fields in $L^1(X, \mathbb{R}^d)$ (resp., tensor-valued fields in $L^1(X, \mathbb{R}^{d\times d})$), the L^2-orthogonal projector mapping on $\mathbb{P}^l(X, \mathbb{R}^d)$ (resp., $\mathbb{P}^l(X, \mathbb{R}^{d\times d})$) acts component-wise and is denoted in boldface font. Let $T \in \mathcal{T}_h$, $n \in [0, l+1]$ and $m \in [0, n]$. The following (n, r, m)-approximation properties of π_T^l hold: For any $v \in W^{n,r}(T, \mathbb{R})$,

$$|v - \pi_T^l v|_{W^{m,r}(T,\mathbb{R})} \lesssim h_T^{n-m} |v|_{W^{n,r}(T,\mathbb{R})}. \tag{11a}$$

The above property will also be used in what follows with r replaced by its conjugate exponent r'. If, additionally, $n \ge 1$, we have the following (n, r')-trace approximation property:

$$\|v - \pi_T^l v\|_{L^{r'}(\partial T,\mathbb{R})} \lesssim h_T^{n-\frac{1}{r'}} |v|_{W^{n,r'}(T,\mathbb{R})}. \tag{11b}$$

The hidden constants in (11) are independent of h and T, but possibly depend on d, the mesh regularity parameter, l, n, and r. The approximation properties (11) are proved for integer n and m in [17, Appendix A.2] (see also [19, Theorem 1.45]), and can be extended to non-integer values using standard interpolation techniques (see, e.g., [38, Theorem 5.1]).

At the global level, for a given integer $l \ge 0$, we define the broken polynomial space $\mathbb{P}^l(\mathcal{T}_h, \mathbb{R})$ spanned by functions in $L^1(\Omega, \mathbb{R})$ whose restriction to each mesh element $T \in \mathcal{T}_h$ lies in $\mathbb{P}^l(T, \mathbb{R})$, and we define the global

L^2-orthogonal projector $\pi_h^l : L^1(\Omega,\mathbb{R}) \to \mathbb{P}^l(\mathcal{T}_h,\mathbb{R})$ such that, for all $v \in L^1(\Omega,\mathbb{R})$ and all $T \in \mathcal{T}_h$,

$$(\pi_h^l v)_{|T} \coloneqq \pi_T^l v_{|T}.$$

Broken polynomial spaces are subspaces of the broken Sobolev spaces

$$W^{n,r}(\mathcal{T}_h,\mathbb{R}) \coloneqq \left\{ v \in L^r(\Omega,\mathbb{R}) \ : \ v_{|T} \in W^{n,r}(T,\mathbb{R}) \quad \forall T \in \mathcal{T}_h \right\}.$$

We define the broken gradient operator $\boldsymbol{\nabla}_h : W^{1,1}(\mathcal{T}_h,\mathbb{R}) \to L^1(\Omega,\mathbb{R}^d)$ such that, for all $v \in W^{1,1}(\mathcal{T}_h,\mathbb{R})$ and all $T \in \mathcal{T}_h$, $(\boldsymbol{\nabla}_h v)_{|T} \coloneqq \boldsymbol{\nabla} v_{|T}$. We define similarly the broken gradient acting on vector fields along with its symmetric part $\boldsymbol{\nabla}_{\mathrm{s},h}$, as well as the broken divergence operator $\boldsymbol{\nabla}_h \cdot$ acting on tensor fields. The global L^2-orthogonal projector $\boldsymbol{\pi}_h^l$ mapping vector-valued fields in $L^1(\Omega,\mathbb{R}^d)$ (resp., tensor-valued fields in $L^1(\Omega,\mathbb{R}^{d\times d})$) on $\mathbb{P}^l(\mathcal{T}_h,\mathbb{R}^d)$ (resp., $\mathbb{P}^l(\mathcal{T}_h,\mathbb{R}^{d\times d})$) is obtained applying π_h^l component-wise.

2.3. **Discrete spaces and norms**

Let an integer $k \geq 1$ be fixed. The HHO space of discrete velocity unknowns is

$$\underline{\boldsymbol{U}}_h^k \coloneqq \left\{ \underline{\boldsymbol{v}}_h = ((\boldsymbol{v}_T)_{T\in\mathcal{T}_h},(\boldsymbol{v}_F)_{F\in\mathcal{F}_h}) \ : \ \boldsymbol{v}_T \in \mathbb{P}^k(T,\mathbb{R}^d) \ \ \forall T \in \mathcal{T}_h \ \text{ and } \ \boldsymbol{v}_F \in \mathbb{P}^k(F,\mathbb{R}^d) \ \ \forall F \in \mathcal{F}_h \right\}.$$

The interpolation operator $\underline{\boldsymbol{I}}_h^k : W^{1,1}(\Omega,\mathbb{R}^d) \to \underline{\boldsymbol{U}}_h^k$ maps a function $\boldsymbol{v} \in W^{1,1}(\Omega,\mathbb{R}^d)$ on the vector of discrete unknowns $\underline{\boldsymbol{I}}_h^k \boldsymbol{v}$ defined as follows:

$$\underline{\boldsymbol{I}}_h^k \boldsymbol{v} \coloneqq ((\boldsymbol{\pi}_T^k \boldsymbol{v}_{|T})_{T\in\mathcal{T}_h},(\boldsymbol{\pi}_F^k \boldsymbol{v}_{|F})_{F\in\mathcal{F}_h}).$$

For all $T \in \mathcal{T}_h$, we denote by $\underline{\boldsymbol{U}}_T^k$ and $\underline{\boldsymbol{I}}_T^k$ the restrictions of $\underline{\boldsymbol{U}}_h^k$ and $\underline{\boldsymbol{I}}_h^k$ to T, respectively and, for all $\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_h^k$, we let $\underline{\boldsymbol{v}}_T \coloneqq (\boldsymbol{v}_T,(\boldsymbol{v}_F)_{F\in\mathcal{F}_T}) \in \underline{\boldsymbol{U}}_T^k$ denote the vector collecting the discrete unknowns attached to T and its faces. Furthermore, for all $\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_h^k$, we define the broken polynomial field $\boldsymbol{v}_h \in \mathbb{P}^k(\mathcal{T}_h,\mathbb{R}^d)$ obtained patching element unknowns, that is,

$$(\boldsymbol{v}_h)_{|T} \coloneqq \boldsymbol{v}_T \qquad \forall T \in \mathcal{T}_h. \tag{12}$$

We define on $\underline{\boldsymbol{U}}_h^k$ the $W^{1,r}(\Omega,\mathbb{R}^d)$-like strain seminorm $\|\cdot\|_{r,h}$ such that, for all $\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_h^k$,

$$\|\underline{\boldsymbol{v}}_h\|_{r,h} \coloneqq \left(\sum_{T\in\mathcal{T}_h} \|\underline{\boldsymbol{v}}_T\|_{r,T}^r \right)^{\frac{1}{r}} \tag{13a}$$

$$\text{with } \|\underline{\boldsymbol{v}}_T\|_{r,T} \coloneqq \left(\|\boldsymbol{\nabla}_{\mathrm{s}} \boldsymbol{v}_T\|_{L^r(T,\mathbb{R}^{d\times d})}^r + \sum_{F\in\mathcal{F}_T} h_F^{1-r} \|\boldsymbol{v}_F - \boldsymbol{v}_T\|_{L^r(F,\mathbb{R}^d)}^r \right)^{\frac{1}{r}} \text{ for all } T \in \mathcal{T}_h. \tag{13b}$$

The following boundedness property for $\underline{\boldsymbol{I}}_T^k$ can be proved adapting the arguments of [19, Proposition 6.24] and requires the star-shaped assumption on the mesh elements: For all $T \in \mathcal{T}_h$ and all $\boldsymbol{v} \in W^{1,r}(T,\mathbb{R}^d)$,

$$\|\underline{\boldsymbol{I}}_T^k \boldsymbol{v}\|_{r,T} \lesssim |\boldsymbol{v}|_{W^{1,r}(T,\mathbb{R}^d)}, \tag{14}$$

where the hidden constant depends only on d, the mesh regularity parameter, r, and k.

The discrete velocity and pressure are sought in the following spaces, which embed, respectively, the homogeneous boundary condition for the velocity and the zero-average constraint for the pressure:

$$\underline{\boldsymbol{U}}_{h,0}^k \coloneqq \left\{ \underline{\boldsymbol{v}}_h = ((\boldsymbol{v}_T)_{T\in\mathcal{T}_h},(\boldsymbol{v}_F)_{F\in\mathcal{F}_h}) \in \underline{\boldsymbol{U}}_h^k \ : \ \boldsymbol{v}_F = \boldsymbol{0} \quad \forall F \in \mathcal{F}_h^{\mathrm{b}} \right\}, \quad P_h^k \coloneqq \mathbb{P}^k(\mathcal{T}_h,\mathbb{R}) \cap P.$$

By the discrete Korn inequality proved in Lemma 5.2 below, $\|\cdot\|_{r,h}$ is a norm on $\underline{\boldsymbol{U}}_{h,0}^k$ (the proof is obtained reasoning as in [19, Corollary 2.16]).

3. HHO scheme

In this section, after introducing the discrete counterparts of the viscous and pressure-velocity coupling terms, we state the discrete problem along with the main results.

3.1. Viscous term

3.1.1. *Local symmetric gradient reconstruction*

For all $T \in \mathcal{T}_h$, we define the local symmetric gradient reconstruction $\mathbf{G}_{\mathrm{s},T}^k : \underline{\boldsymbol{U}}_T^k \to \mathbb{P}^k(T, \mathbb{R}_{\mathrm{s}}^{d\times d})$ such that, for all $\underline{\boldsymbol{v}}_T \in \underline{\boldsymbol{U}}_T^k$,

$$\int_T \mathbf{G}_{\mathrm{s},T}^k \underline{\boldsymbol{v}}_T : \boldsymbol{\tau} = \int_T \boldsymbol{\nabla}_{\mathrm{s}} \boldsymbol{v}_T : \boldsymbol{\tau} + \sum_{F\in\mathcal{F}_T} \int_F (\boldsymbol{v}_F - \boldsymbol{v}_T)\cdot(\boldsymbol{\tau}\boldsymbol{n}_{TF}) \qquad \forall \boldsymbol{\tau} \in \mathbb{P}^k(T, \mathbb{R}_{\mathrm{s}}^{d\times d}). \tag{15}$$

This symmetric gradient reconstruction, originally introduced in [13, Section 4.2], is designed so that the following relation holds (see, e.g., [14, Proposition 5] or [19, Section 7.2.5]): For all $\boldsymbol{v} \in W^{1,1}(T, \mathbb{R}^d)$,

$$\mathbf{G}_{\mathrm{s},T}^k(\boldsymbol{I}_T^k \boldsymbol{v}) = \boldsymbol{\pi}_T^k(\boldsymbol{\nabla}_{\mathrm{s}} \boldsymbol{v}). \tag{16}$$

The global symmetric gradient reconstruction $\mathbf{G}_{\mathrm{s},h}^k : \underline{\boldsymbol{U}}_h^k \to \mathbb{P}^k(\mathcal{T}_h, \mathbb{R}_{\mathrm{s}}^{d\times d})$ is obtained patching the local contributions, that is, for all $\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_h^k$, we set

$$(\mathbf{G}_{\mathrm{s},h}^k \underline{\boldsymbol{v}}_h)_{|_T} \coloneqq \mathbf{G}_{\mathrm{s},T}^k \underline{\boldsymbol{v}}_T \qquad \forall T \in \mathcal{T}_h. \tag{17}$$

3.1.2. *Discrete viscous function*

The discrete counterpart of the function a defined by (7) is $\mathrm{a}_h : \underline{\boldsymbol{U}}_h^k \times \underline{\boldsymbol{U}}_h^k \to \mathbb{R}$ such that, for all $\underline{\boldsymbol{v}}_h, \underline{\boldsymbol{w}}_h \in \underline{\boldsymbol{U}}_h^k$,

$$\mathrm{a}_h(\underline{\boldsymbol{w}}_h, \underline{\boldsymbol{v}}_h) \coloneqq \int_\Omega \boldsymbol{\sigma}(\cdot, \mathbf{G}_{\mathrm{s},h}^k \underline{\boldsymbol{w}}_h) : \mathbf{G}_{\mathrm{s},h}^k \underline{\boldsymbol{v}}_h + \gamma \mathrm{s}_h(\underline{\boldsymbol{w}}_h, \underline{\boldsymbol{v}}_h). \tag{18}$$

In the above definition, recalling (4), γ is a stabilization parameter such that

$$\gamma \in [\sigma_{\mathrm{sm}}, \sigma_{\mathrm{hc}}], \tag{19}$$

while the stabilization function $\mathrm{s}_h : \underline{\boldsymbol{U}}_h^k \times \underline{\boldsymbol{U}}_h^k \to \mathbb{R}$ is such that, for all $\underline{\boldsymbol{v}}_h, \underline{\boldsymbol{w}}_h \in \underline{\boldsymbol{U}}_h^k$,

$$\mathrm{s}_h(\underline{\boldsymbol{w}}_h, \underline{\boldsymbol{v}}_h) \coloneqq \sum_{T\in\mathcal{T}_h} \mathrm{s}_T(\underline{\boldsymbol{w}}_T, \underline{\boldsymbol{v}}_T), \tag{20}$$

where the local contributions are assumed to satisfy the following assumption.

Assumption 3.1 (Local stabilization function)**.** For all $T \in \mathcal{T}_h$, the local stabilization function $\mathrm{s}_T : \underline{\boldsymbol{U}}_T^k \times \underline{\boldsymbol{U}}_T^k \to \mathbb{R}$ is linear in its second argument and satisfies the following properties, with hidden constants independent of both h and T:

(1) *Stability and boundedness.* Recalling the definition (13b) of the local $\|\cdot\|_{r,T}$-seminorm, for all $\underline{\boldsymbol{v}}_T \in \underline{\boldsymbol{U}}_T^k$ it holds:

$$\|\mathbf{G}_{\mathrm{s},T}^k \underline{\boldsymbol{v}}_T\|_{L^r(T,\mathbb{R}^{d\times d})}^r + \mathrm{s}_T(\underline{\boldsymbol{v}}_T, \underline{\boldsymbol{v}}_T) \simeq \|\underline{\boldsymbol{v}}_T\|_{r,T}^r. \tag{21}$$

(2) *Polynomial consistency.* For all $\boldsymbol{w} \in \mathbb{P}^{k+1}(T,\mathbb{R}^d)$ and all $\underline{\boldsymbol{v}}_T \in \underline{\boldsymbol{U}}_T^k$,

$$\mathrm{s}_T(\underline{\boldsymbol{I}}_T^k \boldsymbol{w}, \underline{\boldsymbol{v}}_T) = 0. \tag{22}$$

(3) *Hölder continuity.* For all $\underline{\boldsymbol{u}}_T, \underline{\boldsymbol{v}}_T, \underline{\boldsymbol{w}}_T \in \underline{\boldsymbol{U}}_T^k$, it holds, setting $\underline{\boldsymbol{e}}_T \coloneqq \underline{\boldsymbol{u}}_T - \underline{\boldsymbol{w}}_T$,

$$|\mathrm{s}_T(\underline{\boldsymbol{u}}_T, \underline{\boldsymbol{v}}_T) - \mathrm{s}_T(\underline{\boldsymbol{w}}_T, \underline{\boldsymbol{v}}_T)| \lesssim \left(\mathrm{s}_T(\underline{\boldsymbol{u}}_T, \underline{\boldsymbol{u}}_T) + \mathrm{s}_T(\underline{\boldsymbol{w}}_T, \underline{\boldsymbol{w}}_T)\right)^{\frac{r-\tilde{r}}{r}} \mathrm{s}_T(\underline{\boldsymbol{e}}_T, \underline{\boldsymbol{e}}_T)^{\frac{\tilde{r}-1}{r}} \mathrm{s}_T(\underline{\boldsymbol{v}}_T, \underline{\boldsymbol{v}}_T)^{\frac{1}{r}}. \tag{23}$$

(4) *Strong monotonicity.* For all $\underline{\boldsymbol{u}}_T, \underline{\boldsymbol{w}}_T \in \underline{\boldsymbol{U}}_T^k$, it holds, setting again $\underline{\boldsymbol{e}}_T \coloneqq \underline{\boldsymbol{u}}_T - \underline{\boldsymbol{w}}_T$,

$$\left(\mathrm{s}_T(\underline{\boldsymbol{u}}_T, \underline{\boldsymbol{e}}_T) - \mathrm{s}_T(\underline{\boldsymbol{w}}_T, \underline{\boldsymbol{e}}_T)\right)\left(\mathrm{s}_T(\underline{\boldsymbol{u}}_T, \underline{\boldsymbol{u}}_T) + \mathrm{s}_T(\underline{\boldsymbol{w}}_T, \underline{\boldsymbol{w}}_T)\right)^{\frac{2-\tilde{r}}{r}} \gtrsim \mathrm{s}_T(\underline{\boldsymbol{e}}_T, \underline{\boldsymbol{e}}_T)^{\frac{r+2-\tilde{r}}{r}}. \tag{24}$$

Remark 3.2 (Comparison with the linear case)**.** If $r = 2$, s_T can be any symmetric bilinear form satisfying (21)–(22). Indeed, property (23) coincides in this case with the Cauchy–Schwarz inequality, while, by linearity of s_T, property (24) holds with the equal sign.

3.1.3. *An example of viscous stabilization function*

Taking inspiration from the scalar case (cf., e.g., [17, Eq. (4.11c)]), a local stabilization function that matches Assumption 3.1 can be obtained setting, for all $\underline{\boldsymbol{v}}_T, \underline{\boldsymbol{w}}_T \in \underline{\boldsymbol{U}}_T^k$,

$$\mathrm{s}_T(\underline{\boldsymbol{w}}_T, \underline{\boldsymbol{v}}_T) \coloneqq \int_{\partial T} |\boldsymbol{\Delta}_{\partial T}^k \underline{\boldsymbol{w}}_T|^{r-2} \boldsymbol{\Delta}_{\partial T}^k \underline{\boldsymbol{w}}_T \cdot \boldsymbol{\Delta}_{\partial T}^k \underline{\boldsymbol{v}}_T, \tag{25}$$

where, denoting by $\mathbb{P}^k(\mathcal{F}_T,\mathbb{R}^d)$ the space of vector-valued broken polynomials of total degree $\le k$ on $\mathcal{F}_T$, the boundary residual operator $\boldsymbol{\Delta}_{\partial T}^k : \underline{\boldsymbol{U}}_T^k \to \mathbb{P}^k(\mathcal{F}_T,\mathbb{R}^d)$ is such that, for all $\underline{\boldsymbol{v}}_T \in \underline{\boldsymbol{U}}_T^k$,

$$(\boldsymbol{\Delta}_{\partial T}^k \underline{\boldsymbol{v}}_T)_{|F} \coloneqq h_F^{-\frac{1}{r'}} \left(\boldsymbol{\pi}_F^k(\mathbf{r}_T^{k+1}\underline{\boldsymbol{v}}_T - \boldsymbol{v}_F) - \boldsymbol{\pi}_T^k(\mathbf{r}_T^{k+1}\underline{\boldsymbol{v}}_T - \boldsymbol{v}_T)\right) \qquad \forall F \in \mathcal{F}_T,$$

with velocity reconstruction $\mathbf{r}_T^{k+1} : \underline{\boldsymbol{U}}_T^k \to \mathbb{P}^{k+1}(T,\mathbb{R}^d)$ such that

$$\int_T (\boldsymbol{\nabla}_{\mathrm{s}} \mathbf{r}_T^{k+1}\underline{\boldsymbol{v}}_T - \mathbf{G}_{\mathrm{s},T}^k \underline{\boldsymbol{v}}_T) : \boldsymbol{\nabla}_{\mathrm{s}} \boldsymbol{w} = 0 \qquad \forall \boldsymbol{w} \in \mathbb{P}^{k+1}(T,\mathbb{R}^d),$$
$$\int_T \mathbf{r}_T^{k+1}\underline{\boldsymbol{v}}_T = \int_T \boldsymbol{v}_T, \text{ and } \int_T \boldsymbol{\nabla}_{\mathrm{ss}} \mathbf{r}_T^{k+1}\underline{\boldsymbol{v}}_T = \frac{1}{2}\sum_{F\in\mathcal{F}_T}\int_F (\boldsymbol{v}_F \otimes \boldsymbol{n}_{TF} - \boldsymbol{n}_{TF} \otimes \boldsymbol{v}_F).$$

Above, $\boldsymbol{\nabla}_{\mathrm{ss}}$ denotes the skew-symmetric part of the gradient operator $\boldsymbol{\nabla}$ applied to vector fields and $\otimes$ is the tensor product such that, for all $\boldsymbol{x} = (x_i)_{1\le i\le d}$ and $\boldsymbol{y} = (y_i)_{1\le i\le d}$ in $\mathbb{R}^d$, $\boldsymbol{x}\otimes\boldsymbol{y} \coloneqq (x_i y_j)_{1\le i,j\le d} \in \mathbb{R}^{d\times d}$.

Lemma 3.3 (Stabilization function (25))**.** *The local stabilization function defined by* (25) *satisfies Assumption 3.1.*

Proof. The proof of (21) for $r = 2$ is given in [13, Eq. (25)]. The result can be generalized to $r \neq 2$ using the same arguments of [17, Lemma 5.2]. Property (22) is an immediate consequence of the fact that $\boldsymbol{\Delta}_{\partial T}^k(\underline{\boldsymbol{I}}_T^k \boldsymbol{w}) = \mathbf{0}$ for any $\boldsymbol{w} \in \mathbb{P}^{k+1}(T,\mathbb{R}^d)$, which can be proved reasoning as in [19, Proposition 2.6].

Let us prove (23). First, we remark that, since the function $\alpha \mapsto \alpha^{r-2}$ verifies the conditions in (73b) below, we can apply Theorem A.1 to infer that the function $\mathbb{R}^d \ni \boldsymbol{x} \mapsto |\boldsymbol{x}|^{r-2}\boldsymbol{x}$ satisfies for all $\boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^d$,

$$\left||\boldsymbol{x}|^{r-2}\boldsymbol{x} - |\boldsymbol{y}|^{r-2}\boldsymbol{y}\right| \lesssim \left(|\boldsymbol{x}|^r + |\boldsymbol{y}|^r\right)^{\frac{r-\tilde{r}}{r}} |\boldsymbol{x}-\boldsymbol{y}|^{\tilde{r}-1}, \tag{26a}$$

$$\left(|\boldsymbol{x}|^{r-2}\boldsymbol{x} - |\boldsymbol{y}|^{r-2}\boldsymbol{y}\right)\cdot(\boldsymbol{x}-\boldsymbol{y})\left(|\boldsymbol{x}|^r + |\boldsymbol{y}|^r\right)^{\frac{2-\tilde{r}}{r}} \gtrsim |\boldsymbol{x}-\boldsymbol{y}|^{r+2-\tilde{r}}. \tag{26b}$$

Recalling (25), we can write

$$\begin{aligned}|\mathrm{s}_T(\underline{\boldsymbol{u}}_T,\underline{\boldsymbol{v}}_T)-\mathrm{s}_T(\underline{\boldsymbol{w}}_T,\underline{\boldsymbol{v}}_T)| &\le \int_{\partial T}\left||\boldsymbol{\Delta}^k_{\partial T}\underline{\boldsymbol{u}}_T|^{r-2}\boldsymbol{\Delta}^k_{\partial T}\underline{\boldsymbol{u}}_T-|\boldsymbol{\Delta}^k_{\partial T}\underline{\boldsymbol{w}}_T|^{r-2}\boldsymbol{\Delta}^k_{\partial T}\underline{\boldsymbol{w}}_T\right||\boldsymbol{\Delta}^k_{\partial T}\underline{\boldsymbol{v}}_T|\\ &\lesssim \int_{\partial T}\left(|\boldsymbol{\Delta}^k_{\partial T}\underline{\boldsymbol{u}}_T|^r+|\boldsymbol{\Delta}^k_{\partial T}\underline{\boldsymbol{w}}_T|^r\right)^{\frac{r-\tilde r}{r}}|\boldsymbol{\Delta}^k_{\partial T}\underline{\boldsymbol{e}}_T|^{\tilde r-1}|\boldsymbol{\Delta}^k_{\partial T}\underline{\boldsymbol{v}}_T|\\ &\le (\mathrm{s}_T(\underline{\boldsymbol{u}}_T,\underline{\boldsymbol{u}}_T)+\mathrm{s}_T(\underline{\boldsymbol{w}}_T,\underline{\boldsymbol{w}}_T))^{\frac{r-\tilde r}{r}}\,\mathrm{s}_T(\underline{\boldsymbol{e}}_T,\underline{\boldsymbol{e}}_T)^{\frac{\tilde r-1}{r}}\mathrm{s}_T(\underline{\boldsymbol{v}}_T,\underline{\boldsymbol{v}}_T)^{\frac1r},\end{aligned}$$

where we have used (26a) to pass to the second line and the $(1;\frac{r}{r-\tilde r},\frac{r}{\tilde r-1},r)$-Hölder inequality to conclude.

Moving to (24), (26b) and the $(1;\frac{r+2-\tilde r}{2-\tilde r},\frac{r+2-\tilde r}{r})$-Hölder inequality yield

$$\begin{aligned}&\mathrm{s}_T(\underline{\boldsymbol{e}}_T,\underline{\boldsymbol{e}}_T)\\ &\quad=\int_{\partial T}|\boldsymbol{\Delta}^k_{\partial T}\underline{\boldsymbol{u}}_T-\boldsymbol{\Delta}^k_{\partial T}\underline{\boldsymbol{w}}_T|^r\\ &\quad\lesssim\int_{\partial T}\left(|\boldsymbol{\Delta}^k_{\partial T}\underline{\boldsymbol{u}}_T|^r+|\boldsymbol{\Delta}^k_{\partial T}\underline{\boldsymbol{w}}_T|^r\right)^{\frac{2-\tilde r}{r+2-\tilde r}}\left[\left(|\boldsymbol{\Delta}^k_{\partial T}\underline{\boldsymbol{u}}_T|^{r-2}\boldsymbol{\Delta}^k_{\partial T}\underline{\boldsymbol{u}}_T-|\boldsymbol{\Delta}^k_{\partial T}\underline{\boldsymbol{w}}_T|^{r-2}\boldsymbol{\Delta}^k_{\partial T}\underline{\boldsymbol{w}}_T\right)\cdot\boldsymbol{\Delta}^k_{\partial T}\underline{\boldsymbol{e}}_T\right]^{\frac{r}{r+2-\tilde r}}\\ &\quad\le(\mathrm{s}_T(\underline{\boldsymbol{u}}_T,\underline{\boldsymbol{u}}_T)+\mathrm{s}_T(\underline{\boldsymbol{w}}_T,\underline{\boldsymbol{w}}_T))^{\frac{2-\tilde r}{r+2-\tilde r}}\,(\mathrm{s}_T(\underline{\boldsymbol{u}}_T,\underline{\boldsymbol{e}}_T)-\mathrm{s}_T(\underline{\boldsymbol{w}}_T,\underline{\boldsymbol{e}}_T))^{\frac{r}{r+2-\tilde r}}.\quad\square\end{aligned}$$

3.2. **Pressure-velocity coupling**

For all $T\in\mathcal{T}_h$, we define the local divergence reconstruction $\mathrm{D}^k_T:\underline{\boldsymbol{U}}^k_T\to\mathbb{P}^k(T,\mathbb{R})$ by setting, for all $\underline{\boldsymbol{v}}_T\in\underline{\boldsymbol{U}}^k_T$, $\mathrm{D}^k_T\underline{\boldsymbol{v}}_T\coloneqq\mathrm{tr}(\mathbf{G}^k_{\mathrm{s},T}\underline{\boldsymbol{v}}_T)$. We have the following characterization of D^k_T: For all $\underline{\boldsymbol{v}}_T\in\underline{\boldsymbol{U}}^k_T$,

$$\int_T\mathrm{D}^k_T\underline{\boldsymbol{v}}_T\;q=\int_T(\boldsymbol{\nabla}\cdot\boldsymbol{v}_T)\;q+\sum_{F\in\mathcal{F}_T}\int_F(\boldsymbol{v}_F-\boldsymbol{v}_T)\cdot\boldsymbol{n}_{TF}\;q\qquad\forall q\in\mathbb{P}^k(T,\mathbb{R}),\tag{27}$$

as can be checked writing (15) for $\boldsymbol{\tau}=q\mathrm{I}_d$. Taking the trace of (16), it is inferred that, for all $T\in\mathcal{T}_h$ and all $\boldsymbol{v}\in W^{1,1}(T,\mathbb{R}^d)$, $\mathrm{D}^k_T(\underline{\boldsymbol{I}}^k_T\boldsymbol{v})=\pi^k_T(\boldsymbol{\nabla}\cdot\boldsymbol{v})$. The pressure-velocity coupling is realized by the bilinear form $\mathrm{b}_h:\underline{\boldsymbol{U}}^k_h\times\mathbb{P}^k(\mathcal{T}_h,\mathbb{R})\to\mathbb{R}$ such that, for all $(\underline{\boldsymbol{v}}_h,q_h)\in\underline{\boldsymbol{U}}^k_h\times\mathbb{P}^k(\mathcal{T}_h,\mathbb{R})$, setting $q_T\coloneqq(q_h)_{|T}$ for all $T\in\mathcal{T}_h$,

$$\mathrm{b}_h(\underline{\boldsymbol{v}}_h,q_h)\coloneqq-\sum_{T\in\mathcal{T}_h}\int_T\mathrm{D}^k_T\underline{\boldsymbol{v}}_T\;q_T.\tag{28}$$

3.3. **Discrete problem and main results**

The discrete problem reads: Find $(\underline{\boldsymbol{u}}_h,p_h)\in\underline{\boldsymbol{U}}^k_{h,0}\times P^k_h$ such that

$$\mathrm{a}_h(\underline{\boldsymbol{u}}_h,\underline{\boldsymbol{v}}_h)+\mathrm{b}_h(\underline{\boldsymbol{v}}_h,p_h)=\int_\Omega\boldsymbol{f}\cdot\boldsymbol{v}_h\qquad\forall\underline{\boldsymbol{v}}_h\in\underline{\boldsymbol{U}}^k_{h,0},\tag{29a}$$

$$-\mathrm{b}_h(\underline{\boldsymbol{u}}_h,q_h)=0\qquad\forall q_h\in P^k_h.\tag{29b}$$

Remark 3.4 (Discrete mass equation)**.** The space of test functions in (29b) can be extended to $\mathbb{P}^k(\mathcal{T}_h,\mathbb{R})$ since, for all $\underline{\boldsymbol{v}}_h\in\underline{\boldsymbol{U}}^k_{h,0}$, the divergence theorem together with the fact that $\boldsymbol{v}_F=\boldsymbol{0}$ for all $F\in\mathcal{F}^{\mathrm{b}}_h$ and $\sum_{T\in\mathcal{T}_F}\int_F\boldsymbol{v}_F\cdot\boldsymbol{n}_{TF}=0$ for all $F\in\mathcal{F}^{\mathrm{i}}_h$, yield

$$\mathrm{b}_h(\underline{\boldsymbol{v}}_h,1)=-\sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T}\int_F\boldsymbol{v}_F\cdot\boldsymbol{n}_{TF}=-\sum_{F\in\mathcal{F}^{\mathrm{i}}_h}\sum_{T\in\mathcal{T}_F}\int_F\boldsymbol{v}_F\cdot\boldsymbol{n}_{TF}=0.$$

Remark 3.5 (Efficient implementation)**.** When solving the system of nonlinear algebraic equations corresponding to (29) by, e.g., the Newton algorithm, all element-based velocity unknowns and all but one pressure unknown per element can be locally eliminated at each iteration by static condensation. As all the computations are local, this procedure is an embarrassingly parallel task which can fully benefit from multi-thread and multi-processor architectures. This implementation strategy has been described for the linear Stokes problem in [22, Section 6.2]. After further eliminating the boundary unknowns by strongly enforcing the boundary condition (1c), we end up solving, at each iteration of the nonlinear solver, a linear system of size $d\mathrm{card}(\mathcal{F}_h^{\mathrm{i}})\binom{k+d-1}{d-1} + \mathrm{card}(\mathcal{T}_h)$. Concerning the interplay between the static condensation strategy and the performance of p-multilevel linear solvers, we refer to [11].

In what follows, we state the main results for the HHO scheme (29). The proofs are postponed to Section 6.

Theorem 3.6 (Well-posedness)**.** *There exists a unique solution $(\underline{\boldsymbol{u}}_h, p_h) \in \underline{\boldsymbol{U}}_{h,0}^k \times P_h^k$ to the discrete problem (29). Additionally, the following a priori bounds hold:*

$$\|\underline{\boldsymbol{u}}_h\|_{r,h} \lesssim \left(\sigma_{\mathrm{sm}}^{-1}\|\boldsymbol{f}\|_{L^{r'}(\Omega,\mathbb{R}^d)}\right)^{\frac{1}{r-1}} + \left(\sigma_{\mathrm{de}}^{2-\tilde{r}}\sigma_{\mathrm{sm}}^{-1}\|\boldsymbol{f}\|_{L^{r'}(\Omega,\mathbb{R}^d)}\right)^{\frac{1}{r+1-\tilde{r}}}, \tag{30a}$$

$$\|p_h\|_{L^{r'}(\Omega,\mathbb{R})} \lesssim \sigma_{\mathrm{hc}}\left(\sigma_{\mathrm{sm}}^{-1}\|\boldsymbol{f}\|_{L^{r'}(\Omega,\mathbb{R}^d)} + \sigma_{\mathrm{de}}^{|r-2|(\tilde{r}-1)}\left(\sigma_{\mathrm{sm}}^{-1}\|\boldsymbol{f}\|_{L^{r'}(\Omega,\mathbb{R}^d)}\right)^{\frac{\tilde{r}-1}{r+1-\tilde{r}}}\right). \tag{30b}$$

Proof. See Section 6.2. □

Theorem 3.7 (Error estimate)**.** *Let $(\boldsymbol{u}, p) \in \boldsymbol{U} \times P$ and $(\underline{\boldsymbol{u}}_h, p_h) \in \underline{\boldsymbol{U}}_{h,0}^k \times P_h^k$ solve (6) and (29), respectively. Assume the additional regularity $\boldsymbol{u} \in W^{k+2,r}(\mathcal{T}_h,\mathbb{R}^d)$, $\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{u}) \in W^{1,r'}(\Omega,\mathbb{R}^{d\times d}_{\mathrm{s}}) \cap W^{(k+1)(\tilde{r}-1),r'}(\mathcal{T}_h,\mathbb{R}^{d\times d}_{\mathrm{s}})$, and $p \in W^{1,r'}(\Omega,\mathbb{R}) \cap W^{(k+1)(\tilde{r}-1),r'}(\mathcal{T}_h,\mathbb{R})$. Then, under Assumptions 1.1 and 3.1,*

$$\|\underline{\boldsymbol{u}}_h - \underline{\boldsymbol{I}}_h^k\boldsymbol{u}\|_{r,h} \lesssim h^{\frac{(k+1)(\tilde{r}-1)}{r+1-\tilde{r}}}\left(\sigma_{\mathrm{sm}}^{-1}\mathcal{N}_{\boldsymbol{f}}^{2-\tilde{r}}\mathcal{N}_{\boldsymbol{\sigma},\boldsymbol{u},p}\right)^{\frac{1}{r+1-\tilde{r}}}, \tag{31a}$$

$$\|p_h - \pi_h^k p\|_{L^{r'}(\Omega,\mathbb{R})} \lesssim h^{(k+1)(\tilde{r}-1)}\mathcal{N}_{\boldsymbol{\sigma},\boldsymbol{u},p} + h^{\frac{(k+1)(\tilde{r}-1)^2}{r+1-\tilde{r}}}\sigma_{\mathrm{hc}}\mathcal{N}_{\boldsymbol{f}}^{|r-2|(\tilde{r}-1)}\left(\sigma_{\mathrm{sm}}^{-1}\mathcal{N}_{\boldsymbol{\sigma},\boldsymbol{u},p}\right)^{\frac{\tilde{r}-1}{r+1-\tilde{r}}}, \tag{31b}$$

where we have set, for the sake of brevity,

$$\begin{aligned}\mathcal{N}_{\boldsymbol{\sigma},\boldsymbol{u},p} &\coloneqq \sigma_{\mathrm{hc}}\left(\sigma_{\mathrm{de}}^{r} + |\boldsymbol{u}|_{W^{1,r}(\Omega,\mathbb{R}^d)}^{r}\right)^{\frac{r-\tilde{r}}{r}}|\boldsymbol{u}|_{W^{k+2,r}(\mathcal{T}_h,\mathbb{R}^d)}^{\tilde{r}-1}\\ &\quad + |\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{u})|_{W^{(k+1)(\tilde{r}-1),r'}(\mathcal{T}_h,\mathbb{R}^{d\times d})} + |p|_{W^{(k+1)(\tilde{r}-1),r'}(\mathcal{T}_h,\mathbb{R})},\\ \mathcal{N}_{\boldsymbol{f}} &\coloneqq \sigma_{\mathrm{de}} + \left(\sigma_{\mathrm{sm}}^{-1}\|\boldsymbol{f}\|_{L^{r'}(\Omega,\mathbb{R}^d)}\right)^{\frac{1}{r-1}} + \left(\sigma_{\mathrm{de}}^{2-\tilde{r}}\sigma_{\mathrm{sm}}^{-1}\|\boldsymbol{f}\|_{L^{r'}(\Omega,\mathbb{R}^d)}\right)^{\frac{1}{r+1-\tilde{r}}}.\end{aligned}$$

Proof. See Section 6.3. □

Remark 3.8 (Orders of convergence)**.** From (31), neglecting higher-order terms, we infer asymptotic convergence rates of $\mathcal{O}_{\mathrm{vel}}^k \coloneqq \frac{(k+1)(\tilde{r}-1)}{r+1-\tilde{r}}$ for the velocity and $\mathcal{O}_{\mathrm{pre}}^k \coloneqq \frac{(k+1)(\tilde{r}-1)^2}{r+1-\tilde{r}}$ for the pressure, that is,

$$\mathcal{O}_{\mathrm{vel}}^k = \begin{cases}(k+1)(r-1) & \text{if } r<2,\\ \frac{k+1}{r-1} & \text{if } r\geq 2,\end{cases} \quad\text{and}\quad \mathcal{O}_{\mathrm{pre}}^k = \begin{cases}(k+1)(r-1)^2 & \text{if } r<2,\\ \frac{k+1}{r-1} & \text{if } r\geq 2.\end{cases} \tag{32}$$

Notice that, owing to the presence of higher-order terms in the right-hand sides of (31), higher convergence rates may be observed before attaining the asymptotic ones; see Section 4. The asymptotic order of convergence for the velocity coincides with the one proved in [18, Theorem 3.2] for HHO discretizations of scalar Leray–Lions problems. We refer to [20] for recent improvements on these estimates depending on the degeneracy of the problem.

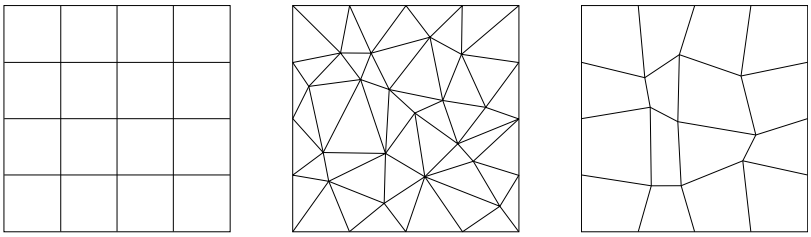

FIGURE 1. Coarsest Cartesian, distorted triangular, and distorted Cartesian meshes used in Section 4.

4. Numerical examples

In this section, we evaluate the numerical performance of the HHO method on a complete panel of numerical test cases. We focus on the $(\mu, 0, 1, r)$-Carreau–Yasuda law (5) (corresponding to the power-law model) with values of the exponent r ranging from 1.25 to 2.75. Our implementation relies on the SpaFEDTe library (cf. `https://spafedte.github.io`).

4.1. **Trigonometric solution**

We begin by considering a manufactured solution to problem (1) in order to assess the convergence of the method. We take $\Omega = (0,1)^2$ and exact velocity $\boldsymbol{u}$ and pressure p given by, respectively,

$$\boldsymbol{u}(x_1,x_2) = \left(\sin\left(\tfrac{\pi}{2}x_1\right)\cos\left(\tfrac{\pi}{2}x_2\right), -\cos\left(\tfrac{\pi}{2}x_1\right)\sin\left(\tfrac{\pi}{2}x_2\right)\right), \quad p(x_1,x_2) = \sin\left(\tfrac{\pi}{2}x_1\right)\sin\left(\tfrac{\pi}{2}x_2\right) - \tfrac{4}{\pi^2}.$$

The volumetric load $\boldsymbol{f}$ and the Dirichlet boundary condition are inferred from the exact solution. Considering $\mu = 1$ and $r \in \{1.5, 1.75, \dots, 2.75\}$, this solution matches the assumptions required in Theorem 3.7 for $k = 1$, except the case $r = 1.5$ for which $\boldsymbol{\sigma}(\cdot, \boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{u}) \notin W^{1,r'}(\Omega, \mathbb{R}^{d\times d}_{\mathrm{s}})$. We consider the HHO scheme for $k = 1$ on three mesh families, namely Cartesian orthogonal, distorted triangular, and distorted Cartesian; see Figure 1. Overall, the results are in agreement with the theoretical predictions, and in some cases the expected asymptotic orders of convergence are exceeded. Specifically, for $r \neq 2$, the convergence rates computed on the last refinement surpass in some cases the theoretical ones. As noticed in Remark 3.8, this suggests that the asymptotic order is still not attained. A similar phenomenon has been observed on certain meshes for the p-Laplace problem; see [18, Section 3.5.2] and [21, Section 3.7]. In some cases, we observe a better convergence for the velocity on distorted triangular meshes than on Cartesian meshes. This phenomenon possibly results from the combination of two factors: on one hand, the improved robustness of HHO methods with respect to elongated elements when compared to classical discretization methods; on the other hand, the fact that unstructured triangular meshes have more elements than Cartesian meshes for a given meshsize and lack privileged directions, which reduces mesh bias. Further investigation is postponed to a future work.

4.2. **Lid-driven cavity flow**

We next consider the lid-driven cavity flow, a well-known problem in fluid mechanics. The domain is the unit square $\Omega = (0,1)^2$, and we enforce a unit tangential velocity $\boldsymbol{u} = (1,0)$ on the top edge (of equation $x_2 = 1$) and wall boundary conditions on the other edges. This boundary condition is incompatible with the formulation (6), even generalized to non-homogeneous boundary conditions, since $\boldsymbol{u} \notin W^{1,r}(\Omega, \mathbb{R}^d)$. However, this is a very classical test that demonstrates the quality of the method. We consider a low Reynolds number $\mathrm{Re} \coloneqq \frac{2}{\mu} = 1$. For $r \in \{1.25, 2, 2.75\}$, we solve the discrete problem on Cartesian and distorted triangular meshes (cf. Figure 1) of approximate size 128×128 for $k = 1$, and 16×16 for $k = 5$. This choice is meant to compare the low-order version of the method on a fine mesh with the high-order version on a very coarse one. The corresponding total number of degrees of freedom is: 130048 for the fine Cartesian mesh with $k = 1$; 5760 for the coarse Cartesian

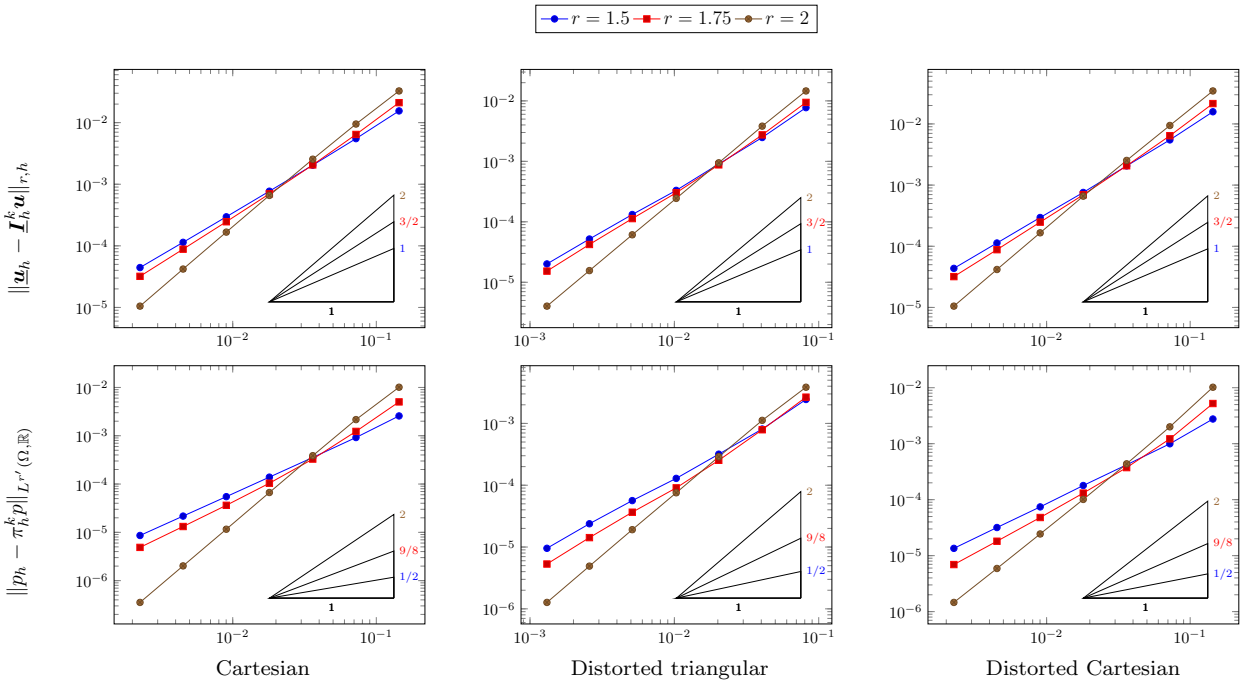

FIGURE 2. Numerical results for the test case of Section 4. The slopes indicate the order of convergence expected from Theorem 3.7, i.e. $\mathcal{O}^1_{\mathrm{vel}} = 2(r-1)$ and $\mathcal{O}^1_{\mathrm{pre}} = 2(r-1)^2$ for $r \in \{1.5, 1.75, 2\}$.

mesh with $k = 5$; 298676 for the fine triangular mesh with $k = 1$; and 14196 for the coarse triangular mesh with $k = 5$. In the left column of Figure 4 we display the velocity magnitude, while in the right column we plot the horizontal component u_1 of the velocity along the vertical centreline $x_1 = \frac{1}{2}$ (resp., vertical component u_2 along the horizontal centreline $x_2 = \frac{1}{2}$). The lines corresponding to $k = 1$ on the fine mesh and to $k = 5$ on the coarse mesh are perfectly superimposed, regardless of the mesh family and of the value of r. This shows that, despite the lack of regularity of the exact solution, high-order versions of the scheme on very coarse meshes deliver similar results as low-order versions on very fine grids. Furthermore, we observe significant differences in the behavior of the flow according to r, coherent with the expected physical behavior. In particular, the viscous effects increase with r, as reflected by the size of the central vortex.

5. Discrete Korn inequality

We prove in this section a discrete counterpart of the following Korn inequality (see [29, Theorem 1]) that will be needed in the analysis: There is $C_{\mathrm{K}} > 0$ depending only on Ω and r such that for all $\boldsymbol{v} \in \boldsymbol{U}$,

$$\|\boldsymbol{v}\|_{W^{1,r}(\Omega,\mathbb{R}^d)} \le C_{\mathrm{K}} \|\boldsymbol{\nabla}_{\mathrm{s}} \boldsymbol{v}\|_{L^r(\Omega,\mathbb{R}^{d\times d})}. \tag{33}$$

We start by recalling the following preliminary result concerning the node-averaging interpolator (sometimes called Oswald interpolator). Let $\mathfrak{T}_h$ be a matching simplicial submesh of $\mathcal{M}_h$ in the sense of [19, Definition 1.8]. The node-averaging operator $\boldsymbol{I}^k_{\mathrm{av},h} : \mathbb{P}^k(\mathcal{T}_h,\mathbb{R}^d) \to \mathbb{P}^k(\mathfrak{T}_h,\mathbb{R}^d) \cap W^{1,r}(\Omega,\mathbb{R}^d)$ is such that, for all $\boldsymbol{v}_h \in \mathbb{P}^k(\mathcal{T}_h,\mathbb{R}^d)$

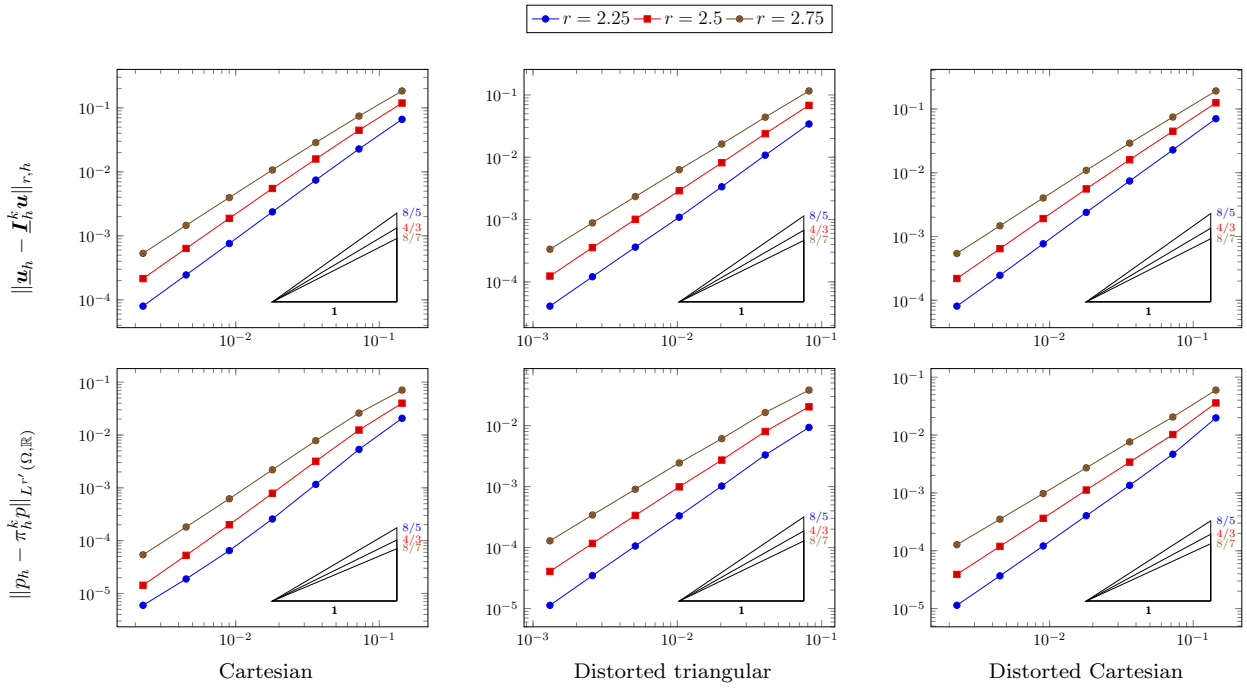

FIGURE 3. Numerical results for the test case of Section 4.1. The slopes indicate the order of convergence expected from Theorem 3.7, i.e. $\mathcal{O}^1_{\mathrm{vel}} = \mathcal{O}^1_{\mathrm{pre}} = \frac{2}{r-1}$ for $r \in \{2.25, 2.5, 2.75\}$.

and all Lagrange node V of $\mathfrak{T}_h$, denoting by $\mathfrak{T}_V$ the set of simplices sharing V,

$$(\boldsymbol{I}^k_{\mathrm{av},h}\boldsymbol{v}_h)(V) \coloneqq \begin{cases} \frac{1}{\mathrm{card}(\mathfrak{T}_V)} \sum_{\boldsymbol{\tau}\in\mathfrak{T}_V} \boldsymbol{v}_{h|\boldsymbol{\tau}}(V) & \text{if } V \in \Omega, \\ \boldsymbol{0} & \text{if } V \in \partial\Omega. \end{cases}$$

For all $F \in \mathcal{F}_h^{\mathrm{i}}$, denote by $T_1, T_2 \in \mathcal{T}_h$ the elements sharing F, taken in an arbitrary but fixed order. We define the jump operator such that, for any function $\boldsymbol{v} \in W^{1,1}(\mathcal{T}_h, \mathbb{R}^d)$, $[\boldsymbol{v}]_F \coloneqq (\boldsymbol{v}_{|T_1})_{|F} - (\boldsymbol{v}_{|T_2})_{|F}$. This definition is extended to boundary faces $F \in \mathcal{F}_h^{\mathrm{b}}$ by setting $[\boldsymbol{v}]_F \coloneqq \boldsymbol{v}_{|F}$.

Proposition 5.1 (Boundedness of the node-averaging operator)**.** *For all $\boldsymbol{v}_h \in \mathbb{P}^k(\mathcal{T}_h, \mathbb{R}^d)$, it holds*

$$|\boldsymbol{v}_h - \boldsymbol{I}^k_{\mathrm{av},h}\boldsymbol{v}_h|^r_{W^{1,r}(\mathcal{T}_h,\mathbb{R}^d)} \lesssim \sum_{F\in\mathcal{F}_h} h_F^{1-r} \|[\boldsymbol{v}_h]_F\|^r_{L^r(F,\mathbb{R}^d)}. \tag{34}$$

Proof. Combining [19, Eq. (4.13)] (which corresponds to (34) for $r = 2$) with the local Lebesgue embeddings of [19, Lemma 1.25] (see also [17, Lemma 5.1]) gives, for any $T \in \mathcal{T}_h$,

$$\|\boldsymbol{v}_h - \boldsymbol{I}^k_{\mathrm{av},h}\boldsymbol{v}_h\|^r_{L^r(T,\mathbb{R}^d)} \lesssim \sum_{F\in\mathcal{F}_{\mathcal{V},T}} h_F \|[\boldsymbol{v}_h]_F\|^r_{L^r(F,\mathbb{R}^d)}, \tag{35}$$

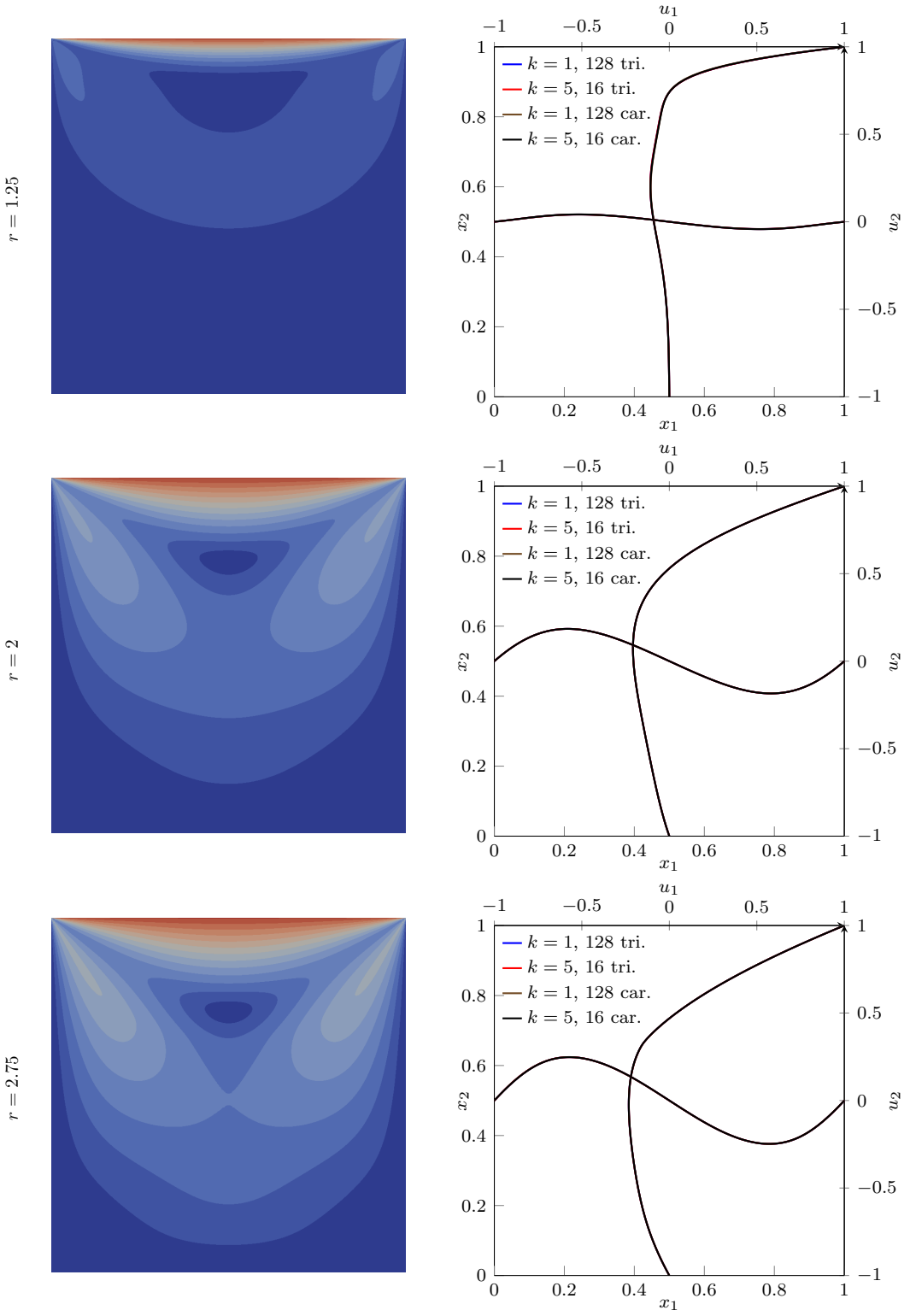

FIGURE 4. Numerical results for the test case of Section 4.2. *Left:* velocity magnitude contours (15 equispaced values in the range $[0, 1]$). Computations on a Cartesian mesh of size 128×128 with $k = 5$. *Right:* horizontal component u_1 of the velocity along the vertical centreline $x_1 = \frac{1}{2}$ and vertical component u_2 of the velocity along the horizontal centreline $x_2 = \frac{1}{2}$.

where $\mathcal{F}_{\mathcal{V},T}$ collects the faces whose closure has non-empty intersection with $\overline{T}$. Using the local inverse inequality of [19, Lemma 1.28] (see also [17, Eq. (A.1)]), we can write

$$\begin{aligned}|\boldsymbol{v}_h - \boldsymbol{I}_{\mathrm{av},h}^k \boldsymbol{v}_h|_{W^{1,r}(\mathcal{T}_h,\mathbb{R}^d)}^r &\lesssim \sum_{T\in\mathcal{T}_h} h_T^{-r} \|\boldsymbol{v}_h - \boldsymbol{I}_{\mathrm{av},h}^k \boldsymbol{v}_h\|_{L^r(T,\mathbb{R}^d)}^r \\ &\lesssim \sum_{T\in\mathcal{T}_h} \sum_{F\in\mathcal{F}_{\mathcal{V},T}} h_F^{1-r} \|[\boldsymbol{v}_h]_F\|_{L^r(F,\mathbb{R}^d)}^r \\ &\lesssim \sum_{F\in\mathcal{F}_h} \sum_{T\in\mathcal{T}_{\mathcal{V},F}} h_F^{1-r} \|[\boldsymbol{v}_h]_F\|_{L^r(F,\mathbb{R}^d)}^r \\ &\le \max_{F\in\mathcal{F}_h} \mathrm{card}(\mathcal{T}_{\mathcal{V},F}) \sum_{F\in\mathcal{F}_h} h_F^{1-r} \|[\boldsymbol{v}_h]_F\|_{L^r(F,\mathbb{R}^d)}^r,\end{aligned}$$

where we have used the fact that $h_T^{-r} \le h_F^{-r}$ along with inequality (35) to pass to the second line, and we have exchanged the sums after setting $\mathcal{T}_{\mathcal{V},F} \coloneqq \{T \in \mathcal{T}_h : \overline{F} \cap \overline{T} \neq \emptyset\}$ for all $F \in \mathcal{F}_h$ to pass to the third line. Observing that $\max_{F\in\mathcal{F}_h} \mathrm{card}(\mathcal{T}_{\mathcal{V},F}) \lesssim 1$ (since, for any $F \in \mathcal{F}_h$, $\mathrm{card}(\mathcal{T}_{\mathcal{V},F})$ is bounded by the left-hand side of [19, Eq. (4.23)] written for any $T \in \mathcal{T}_h$ to which F belongs), (34) follows. □

The following inequalities between sums of powers will be often used in what follows without necessarily recalling this fact explicitly each time. Let an integer $n \ge 1$ and a real number $m \in (0,\infty)$ be given. Then, for all $a_1,\dots,a_n \in (0,\infty)$, we have

$$n^{-(m-1)^\ominus} \sum_{i=1}^n a_i^m \le \left(\sum_{i=1}^n a_i\right)^m \le n^{(m-1)^\oplus} \sum_{i=1}^n a_i^m. \tag{36}$$

If $m = 1$, then (36) holds with the equal sign. If $m < 1$, [45, Eqs. (5) and (3)] with $\alpha = 1$ and $\beta = m$ give $n^{m-1}\sum_{i=1}^n a_i^m \le (\sum_{i=1}^n a_i)^m \le \sum_{i=1}^n a_i^m$. If, on the other hand, $m > 1$, [45, Eqs. (3) and (5)] with $\alpha = m$ and $\beta = 1$ give $\sum_{i=1}^n a_i^m \le (\sum_{i=1}^n a_i)^m \le n^{m-1}\sum_{i=1}^n a_i^m$. Gathering the above cases yields (36).

Lemma 5.2. *(Discrete Korn inequality) We have, for all $\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_{h,0}^k$, recalling the notation* (12)*,*

$$\|\boldsymbol{v}_h\|_{L^r(\Omega,\mathbb{R}^d)}^r + |\boldsymbol{v}_h|_{W^{1,r}(\mathcal{T}_h,\mathbb{R}^d)}^r \lesssim \|\underline{\boldsymbol{v}}_h\|_{r,h}^r. \tag{37}$$

Proof. Let $\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_{h,0}^k$. Using a triangle inequality followed by (36), we can write

$$\begin{aligned}|\boldsymbol{v}_h|_{W^{1,r}(\mathcal{T}_h,\mathbb{R}^d)}^r &\lesssim |\boldsymbol{I}_{\mathrm{av},h}^k \boldsymbol{v}_h|_{W^{1,r}(\mathcal{T}_h,\mathbb{R}^d)}^r + |\boldsymbol{v}_h - \boldsymbol{I}_{\mathrm{av},h}^k \boldsymbol{v}_h|_{W^{1,r}(\mathcal{T}_h,\mathbb{R}^d)}^r \\ &\lesssim \|\boldsymbol{\nabla}_{\mathrm{s}}(\boldsymbol{I}_{\mathrm{av},h}^k \boldsymbol{v}_h)\|_{L^r(\Omega,\mathbb{R}^{d\times d})}^r + |\boldsymbol{v}_h - \boldsymbol{I}_{\mathrm{av},h}^k \boldsymbol{v}_h|_{W^{1,r}(\mathcal{T}_h,\mathbb{R}^d)}^r \\ &\lesssim \|\boldsymbol{\nabla}_{\mathrm{s},h} \boldsymbol{v}_h\|_{L^r(\Omega,\mathbb{R}^{d\times d})}^r + |\boldsymbol{v}_h - \boldsymbol{I}_{\mathrm{av},h}^k \boldsymbol{v}_h|_{W^{1,r}(\mathcal{T}_h,\mathbb{R}^d)}^r \\ &\lesssim \|\boldsymbol{\nabla}_{\mathrm{s},h} \boldsymbol{v}_h\|_{L^r(\Omega,\mathbb{R}^{d\times d})}^r + \sum_{F\in\mathcal{F}_h} h_F^{1-r} \|[\boldsymbol{v}_h]_F\|_{L^r(F,\mathbb{R}^d)}^r,\end{aligned}$$

where we have used the continuous Korn inequality (33) to pass to the second line, we have inserted $\pm\boldsymbol{\nabla}_{\mathrm{s},h}\boldsymbol{v}_h$ into the first norm and used a triangle inequality followed by (36) to pass to the third line, and we have invoked the bound (34) to conclude. Observing that, for any $F \in \mathcal{F}_h$, $|[\boldsymbol{v}_h]_F| \le \sum_{T\in\mathcal{T}_F} |\boldsymbol{v}_F - \boldsymbol{v}_T|$ by a triangle inequality, and using (36), we can continue writing

$$|\boldsymbol{v}_h|_{W^{1,r}(\mathcal{T}_h,\mathbb{R}^d)}^r \lesssim \|\boldsymbol{\nabla}_{\mathrm{s},h} \boldsymbol{v}_h\|_{L^r(\Omega,\mathbb{R}^{d\times d})}^r + \sum_{F\in\mathcal{F}_h} \sum_{T\in\mathcal{T}_F} h_F^{1-r} \|\boldsymbol{v}_F - \boldsymbol{v}_T\|_{L^r(F,\mathbb{R}^d)}^r = \|\underline{\boldsymbol{v}}_h\|_{r,h}^r,$$

where we have exchanged the sums over faces and elements and recalled definition (13a) to conclude. This proves the bound for the second term in the left-hand side of (37). Combining this result with the global discrete Sobolev embeddings of [17, Proposition 5.4] yields the bound for the first term in (37). □

6. Well-posedness and convergence analysis

In this section, after studying the stabilization function s_h, we prove the main results stated in Section 3.3.

6.1. Properties of the stabilization function

Lemma 6.1 (Consistency of s_T)**.** *For any $T \in \mathcal{T}_h$ and any s_T satisfying Assumption 3.1, it holds, for all $\boldsymbol{w} \in W^{k+2,r}(T,\mathbb{R}^d)$ and all $\underline{\boldsymbol{v}}_T \in \underline{\boldsymbol{U}}_T^k$,*

$$|\mathrm{s}_T(\underline{\boldsymbol{I}}_T^k \boldsymbol{w}, \underline{\boldsymbol{v}}_T)| \lesssim h_T^{(k+1)(\tilde{r}-1)} |\boldsymbol{w}|^{r-\tilde{r}}_{W^{1,r}(T,\mathbb{R}^d)} |\boldsymbol{w}|^{\tilde{r}-1}_{W^{k+2,r}(T,\mathbb{R}^d)} \|\underline{\boldsymbol{v}}_T\|_{r,T}, \tag{38}$$

where the hidden constant is independent of h, T, and $\boldsymbol{w}$.

Proof. The proof adapts the arguments of [19, Propositon 2.14]. Using the polynomial consistency property (22), we can write

$$\begin{aligned}
|\mathrm{s}_T(\underline{\boldsymbol{I}}_T^k \boldsymbol{w}, \underline{\boldsymbol{v}}_T)| &= |\mathrm{s}_T(\underline{\boldsymbol{I}}_T^k \boldsymbol{w}, \underline{\boldsymbol{v}}_T) - \mathrm{s}_T(\underline{\boldsymbol{I}}_T^k(\boldsymbol{\pi}_T^{k+1}\boldsymbol{w}), \underline{\boldsymbol{v}}_T)| \\
&\lesssim \mathrm{s}_T(\underline{\boldsymbol{I}}_T^k \boldsymbol{w}, \underline{\boldsymbol{I}}_T^k \boldsymbol{w})^{\frac{r-\tilde{r}}{r}} \mathrm{s}_T(\underline{\boldsymbol{I}}_T^k(\boldsymbol{w} - \boldsymbol{\pi}_T^{k+1}\boldsymbol{w}), \underline{\boldsymbol{I}}_T^k(\boldsymbol{w} - \boldsymbol{\pi}_T^{k+1}\boldsymbol{w}))^{\frac{\tilde{r}-1}{r}} \mathrm{s}_T(\underline{\boldsymbol{v}}_T, \underline{\boldsymbol{v}}_T)^{\frac{1}{r}} \\
&\lesssim \|\underline{\boldsymbol{I}}_T^k \boldsymbol{w}\|_{r,T}^{r-\tilde{r}} \|\underline{\boldsymbol{I}}_T^k(\boldsymbol{w} - \boldsymbol{\pi}_T^{k+1}\boldsymbol{w})\|_{r,T}^{\tilde{r}-1} \|\underline{\boldsymbol{v}}_T\|_{r,T} \\
&\lesssim |\boldsymbol{w}|^{r-\tilde{r}}_{W^{1,r}(T,\mathbb{R}^d)} |\boldsymbol{w} - \boldsymbol{\pi}_T^{k+1}\boldsymbol{w}|^{\tilde{r}-1}_{W^{1,r}(T,\mathbb{R}^d)} \|\underline{\boldsymbol{v}}_T\|_{r,T} \\
&\lesssim h_T^{(k+1)(\tilde{r}-1)} |\boldsymbol{w}|^{r-\tilde{r}}_{W^{1,r}(T,\mathbb{R}^d)} |\boldsymbol{w}|^{\tilde{r}-1}_{W^{k+2,r}(T,\mathbb{R}^d)} \|\underline{\boldsymbol{v}}_T\|_{r,T},
\end{aligned}$$

where we have used the Hölder continuity (23) and observed that, by the polynomial consistency property (22), $\mathrm{s}_T(\underline{\boldsymbol{I}}_T^k(\boldsymbol{\pi}_T^{k+1}\boldsymbol{w}), \underline{\boldsymbol{I}}_T^k(\boldsymbol{\pi}_T^{k+1}\boldsymbol{w})) = 0$ to pass to the second line, we have used the boundedness property (21) to pass to the third line, the boundedness (14) of $\underline{\boldsymbol{I}}_T^k$ to pass to the fourth line, and the $(k+2,r,1)$-approximation property (11a) of $\boldsymbol{\pi}_T^{k+1}$ to conclude. □

In what follows, we will need generalized versions of the continuous and discrete Hölder inequalities, recalled hereafter for the sake of convenience. Let $X \subset \mathbb{R}^d$ be measurable, $n \in \mathbb{N}^*$, and let $t, p_1, \dots, p_n \in (0,\infty]$ be such that $\sum_{i=1}^n \frac{1}{p_i} = \frac{1}{t}$. The continuous $(t; p_1, \dots, p_n)$-Hölder inequality reads: For any $(f_1, \dots, f_n) \in \bigtimes_{i=1}^n L^{p_i}(X,\mathbb{R})$,

$$\left\| \prod_{i=1}^n f_i \right\|_{L^t(X,\mathbb{R})} \le \prod_{i=1}^n \|f_i\|_{L^{p_i}(X,\mathbb{R})}. \tag{39}$$

Let $m \in \mathbb{N}^*$. For all $f : \{1,\dots,m\} \to \mathbb{R}$ and all $q \in [1,\infty)$, setting $\|f\|_q \coloneqq (\sum_{i=1}^m |f(i)|^q)^{\frac{1}{q}}$, and $\|f\|_\infty \coloneqq \max_{1\le i\le m} |f(i)|$, the discrete $(t; p_1, \dots, p_n)$-Hölder inequality reads: For any $f_1, \dots, f_n : \{1,\dots,m\} \to \mathbb{R}$,

$$\left\| \prod_{i=1}^n f_i \right\|_t \le \prod_{i=1}^n \|f_i\|_{p_i}. \tag{40}$$

Proposition 6.2 (Properties of s_h)**.** *Let s_h be given by (20) with, for all $T \in \mathcal{T}_h$, s_T satisfying Assumption 3.1. Then it holds, for all $\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_h^k$,*

$$\|\mathbf{G}_{\mathrm{s},h}^k \underline{\boldsymbol{v}}_h\|^r_{L^r(\Omega,\mathbb{R}^{d\times d})} + \mathrm{s}_h(\underline{\boldsymbol{v}}_h, \underline{\boldsymbol{v}}_h) \simeq \|\underline{\boldsymbol{v}}_h\|^r_{r,h}. \tag{41a}$$

Furthermore, for all $\underline{\boldsymbol{u}}_h, \underline{\boldsymbol{v}}_h, \underline{\boldsymbol{w}}_h \in \underline{\boldsymbol{U}}_h^k$ *it holds, setting* $\underline{\boldsymbol{e}}_h \coloneqq \underline{\boldsymbol{u}}_h - \underline{\boldsymbol{w}}_h$,

$$|\mathrm{s}_h(\underline{\boldsymbol{u}}_h, \underline{\boldsymbol{v}}_h) - \mathrm{s}_h(\underline{\boldsymbol{w}}_h, \underline{\boldsymbol{v}}_h)| \lesssim (\mathrm{s}_h(\underline{\boldsymbol{u}}_h, \underline{\boldsymbol{u}}_h) + \mathrm{s}_h(\underline{\boldsymbol{w}}_h, \underline{\boldsymbol{w}}_h))^{\frac{r-\tilde{r}}{r}} \mathrm{s}_h(\underline{\boldsymbol{e}}_h, \underline{\boldsymbol{e}}_h)^{\frac{\tilde{r}-1}{r}} \mathrm{s}_h(\underline{\boldsymbol{v}}_h, \underline{\boldsymbol{v}}_h)^{\frac{1}{r}}, \tag{41b}$$

$$(\mathrm{s}_h(\underline{\boldsymbol{u}}_h, \underline{\boldsymbol{e}}_h) - \mathrm{s}_h(\underline{\boldsymbol{w}}_h, \underline{\boldsymbol{e}}_h)) (\mathrm{s}_h(\underline{\boldsymbol{u}}_h, \underline{\boldsymbol{u}}_h) + \mathrm{s}_h(\underline{\boldsymbol{w}}_h, \underline{\boldsymbol{w}}_h))^{\frac{2-\tilde{r}}{r}} \gtrsim \mathrm{s}_h(\underline{\boldsymbol{e}}_h, \underline{\boldsymbol{e}}_h)^{\frac{r+2-\tilde{r}}{r}}. \tag{41c}$$

Finally, for any $\boldsymbol{w} \in \boldsymbol{U} \cap W^{k+2,r}(\mathcal{T}_h, \mathbb{R}^d)$, *it holds*

$$\sup_{\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_{h,0}^k, \|\underline{\boldsymbol{v}}_h\|_{r,h}=1} \mathrm{s}_h(\boldsymbol{I}_h^k \boldsymbol{w}, \underline{\boldsymbol{v}}_h) \lesssim h^{(k+1)(\tilde{r}-1)} |\boldsymbol{w}|_{W^{1,r}(\Omega,\mathbb{R}^d)}^{r-\tilde{r}} |\boldsymbol{w}|_{W^{k+2,r}(\mathcal{T}_h,\mathbb{R}^d)}^{\tilde{r}-1}. \tag{42}$$

Above, the hidden constants are independent of h *and of the arguments of* s_h.

Proof. For the sake of conciseness, we only sketch the proof and leave the details to the reader. Summing (21) over $T \in \mathcal{T}_h$ immediately yields (41a). The Hölder continuity property (41b) follows applying to the quantity in the left-hand side triangle inequalities, using (23), and concluding with a discrete $(1; \frac{r}{r-\tilde{r}}, \frac{r}{\tilde{r}-1}, r)$-Hölder inequality. Moving to (41c), starting from $|\mathrm{s}_h(\underline{\boldsymbol{e}}_h, \underline{\boldsymbol{e}}_h)|$, we use (24) and apply a discrete $(1; \frac{r+2-\tilde{r}}{2-\tilde{r}}, \frac{r+2-\tilde{r}}{r})$-Hölder inequality to conclude. Finally, to prove (42) we start from $\mathrm{s}_h(\boldsymbol{I}_h^k \boldsymbol{w}, \underline{\boldsymbol{v}}_h)$, expand this quantity according to (20), use, for all $T \in \mathcal{T}_h$, the local consistency property (38) together with $h_T \le h$, invoke the discrete $(1; \frac{r}{r-\tilde{r}}, \frac{r}{\tilde{r}-1}, r)$-Hölder inequality, and pass to the supremum to conclude. □

6.2. **Well-posedness**

In this section, after proving Hölder continuity and strong monotonicity properties for the discrete viscous function a_h and the inf-sup stability of the pressure-velocity coupling bilinear form b_h, we prove Theorem 3.6.

6.2.1. *Hölder continuity and strong monotonicity of the viscous function*

Lemma 6.3 (Hölder continuity and strong monotonicity of a_h)**.** *For all* $\underline{\boldsymbol{u}}_h, \underline{\boldsymbol{v}}_h, \underline{\boldsymbol{w}}_h \in \underline{\boldsymbol{U}}_h^k$, *setting* $\underline{\boldsymbol{e}}_h \coloneqq \underline{\boldsymbol{u}}_h - \underline{\boldsymbol{w}}_h$, *it holds*

$$|\mathrm{a}_h(\underline{\boldsymbol{u}}_h, \underline{\boldsymbol{v}}_h) - \mathrm{a}_h(\underline{\boldsymbol{w}}_h, \underline{\boldsymbol{v}}_h)| \lesssim \sigma_{\mathrm{hc}} \left(\sigma_{\mathrm{de}}^r + \|\underline{\boldsymbol{u}}_h\|_{r,h}^r + \|\underline{\boldsymbol{w}}_h\|_{r,h}^r\right)^{\frac{r-\tilde{r}}{r}} \|\underline{\boldsymbol{e}}_h\|_{r,h}^{\tilde{r}-1} \|\underline{\boldsymbol{v}}_h\|_{r,h}, \tag{43a}$$

$$(\mathrm{a}_h(\underline{\boldsymbol{u}}_h, \underline{\boldsymbol{e}}_h) - \mathrm{a}_h(\underline{\boldsymbol{w}}_h, \underline{\boldsymbol{e}}_h)) \left(\sigma_{\mathrm{de}}^r + \|\underline{\boldsymbol{u}}_h\|_{r,h}^r + \|\underline{\boldsymbol{w}}_h\|_{r,h}^r\right)^{\frac{2-\tilde{r}}{r}} \gtrsim \sigma_{\mathrm{sm}} \|\underline{\boldsymbol{e}}_h\|_{r,h}^{r+2-\tilde{r}}. \tag{43b}$$

Proof. (i) *Hölder continuity.* Using a Cauchy–Schwarz inequality followed by the Hölder continuity (3c) of $\boldsymbol{\sigma}$, we can write

$$\begin{aligned}
&\left| \int_\Omega \left(\boldsymbol{\sigma}(\cdot, \mathbf{G}_{\mathrm{s},h}^k \underline{\boldsymbol{u}}_h) - \boldsymbol{\sigma}(\cdot, \mathbf{G}_{\mathrm{s},h}^k \underline{\boldsymbol{w}}_h)\right) : \mathbf{G}_{\mathrm{s},h}^k \underline{\boldsymbol{v}}_h \right| \\
&\quad \le \sigma_{\mathrm{hc}} \int_\Omega \left(\sigma_{\mathrm{de}}^r + |\mathbf{G}_{\mathrm{s},h}^k \underline{\boldsymbol{u}}_h|_{d\times d}^r + |\mathbf{G}_{\mathrm{s},h}^k \underline{\boldsymbol{w}}_h|_{d\times d}^r\right)^{\frac{r-\tilde{r}}{r}} |\mathbf{G}_{\mathrm{s},h}^k \underline{\boldsymbol{e}}_h|_{d\times d}^{\tilde{r}-1} |\mathbf{G}_{\mathrm{s},h}^k \underline{\boldsymbol{v}}_h|_{d\times d} \\
&\quad \lesssim \sigma_{\mathrm{hc}} \left(|\Omega|_d \sigma_{\mathrm{de}}^r + \|\mathbf{G}_{\mathrm{s},h}^k \underline{\boldsymbol{u}}_h\|_{L^r(\Omega,\mathbb{R}^{d\times d})}^r + \|\mathbf{G}_{\mathrm{s},h}^k \underline{\boldsymbol{w}}_h\|_{L^r(\Omega,\mathbb{R}^{d\times d})}^r\right)^{\frac{r-\tilde{r}}{r}} \\
&\qquad \times \|\mathbf{G}_{\mathrm{s},h}^k \underline{\boldsymbol{e}}_h\|_{L^r(\Omega,\mathbb{R}^{d\times d})}^{\tilde{r}-1} \|\mathbf{G}_{\mathrm{s},h}^k \underline{\boldsymbol{v}}_h\|_{L^r(\Omega,\mathbb{R}^{d\times d})} \\
&\quad \lesssim \sigma_{\mathrm{hc}} \left(\sigma_{\mathrm{de}}^r + \|\underline{\boldsymbol{u}}_h\|_{r,h}^r + \|\underline{\boldsymbol{w}}_h\|_{r,h}^r\right)^{\frac{r-\tilde{r}}{r}} \|\underline{\boldsymbol{e}}_h\|_{r,h}^{\tilde{r}-1} \|\underline{\boldsymbol{v}}_h\|_{r,h},
\end{aligned} \tag{44}$$

where we have used the $(1; \frac{r}{r-\tilde{r}}, \frac{r}{\tilde{r}-1}, r)$-Hölder inequality (39) in the second bound and the global seminorm equivalence (41a) together with the fact that $|\Omega|_d \lesssim 1$ (since Ω is bounded) to conclude. For the stabilization term, combining the Hölder continuity (41b) of s_h and the seminorm equivalence (41a) readily gives

$$|\mathrm{s}_h(\underline{\boldsymbol{u}}_h, \underline{\boldsymbol{v}}_h) - \mathrm{s}_h(\underline{\boldsymbol{w}}_h, \underline{\boldsymbol{v}}_h)| \lesssim \left(\sigma_{\mathrm{de}}^r + \|\underline{\boldsymbol{u}}_h\|_{r,h}^r + \|\underline{\boldsymbol{w}}_h\|_{r,h}^r\right)^{\frac{r-\tilde{r}}{r}} \|\underline{\boldsymbol{e}}_h\|_{r,h}^{\tilde{r}-1} \|\underline{\boldsymbol{v}}_h\|_{r,h}, \tag{45}$$

where we have additionally noticed that $\sigma_{\rm de}^r \geq 0$ to add this term to the quantity inside parentheses. Using the definition (18) of a_h, a triangle inequality followed by (44) and (45), and recalling that $\gamma \leq \sigma_{\rm hc}$ (cf. (19)), (43a) follows.

(ii) *Strong monotonicity.* Using the strong monotonicity (3d) of $\boldsymbol{\sigma}$ and the $(1; \frac{r+2-\tilde{r}}{2-\tilde{r}}, \frac{r+2-\tilde{r}}{r})$-Hölder inequality (39), we get

$$
\begin{aligned}
&\sigma_{\rm sm}^{\frac{r}{r+2-\tilde{r}}} \|\mathbf{G}_{{\rm s},h}^k \underline{\boldsymbol{e}}_h\|_{L^r(\Omega,\mathbb{R}^{d\times d})}^r \\
&\leq \int_\Omega \Big(\sigma_{\rm de}^r + |\mathbf{G}_{{\rm s},h}^k \underline{\boldsymbol{u}}_h|_{d\times d}^r + |\mathbf{G}_{{\rm s},h}^k \underline{\boldsymbol{w}}_h|_{d\times d}^r\Big)^{\frac{2-\tilde{r}}{r+2-\tilde{r}}} \Big(\big(\boldsymbol{\sigma}(\cdot, \mathbf{G}_{{\rm s},h}^k \underline{\boldsymbol{u}}_h) - \boldsymbol{\sigma}(\cdot, \mathbf{G}_{{\rm s},h}^k \underline{\boldsymbol{w}}_h)\big) : \mathbf{G}_{{\rm s},h}^k \underline{\boldsymbol{e}}_h\Big)^{\frac{r}{r+2-\tilde{r}}} \\
&\lesssim \Big(\sigma_{\rm de}^r + \|\mathbf{G}_{{\rm s},h}^k \underline{\boldsymbol{u}}_h\|_{L^r(\Omega,\mathbb{R}^{d\times d})}^r + \|\mathbf{G}_{{\rm s},h}^k \underline{\boldsymbol{w}}_h\|_{L^r(\Omega,\mathbb{R}^{d\times d})}^r\Big)^{\frac{2-\tilde{r}}{r+2-\tilde{r}}} \\
&\qquad \times \left(\int_\Omega \Big(\boldsymbol{\sigma}(\cdot, \mathbf{G}_{{\rm s},h}^k \underline{\boldsymbol{u}}_h) - \boldsymbol{\sigma}(\cdot, \mathbf{G}_{{\rm s},h}^k \underline{\boldsymbol{w}}_h)\Big) : \mathbf{G}_{{\rm s},h}^k \underline{\boldsymbol{e}}_h\right)^{\frac{r}{r+2-\tilde{r}}} \\
&\lesssim \big(\sigma_{\rm de}^r + \|\underline{\boldsymbol{u}}_h\|_{r,h}^r + \|\underline{\boldsymbol{w}}_h\|_{r,h}^r\big)^{\frac{2-\tilde{r}}{r+2-\tilde{r}}} \left(\int_\Omega \Big(\boldsymbol{\sigma}(\cdot, \mathbf{G}_{{\rm s},h}^k \underline{\boldsymbol{u}}_h) - \boldsymbol{\sigma}(\cdot, \mathbf{G}_{{\rm s},h}^k \underline{\boldsymbol{w}}_h)\Big) : \mathbf{G}_{{\rm s},h}^k \underline{\boldsymbol{e}}_h\right)^{\frac{r}{r+2-\tilde{r}}},
\end{aligned}
\tag{46}
$$

where the conclusion follows from the global seminorm equivalence (41a). Additionally, using the strong monotonicity (41c) of s_h together with the fact that $\sigma_{\rm sm} \leq \gamma$ (cf. (19)) and invoking again the seminorm equivalence (41a), we readily obtain

$$
\sigma_{\rm sm}^{\frac{r}{r+2-\tilde{r}}} \mathrm{s}_h(\underline{\boldsymbol{e}}_h, \underline{\boldsymbol{e}}_h) \lesssim \big(\sigma_{\rm de}^r + \|\underline{\boldsymbol{u}}_h\|_{r,h}^r + \|\underline{\boldsymbol{w}}_h\|_{r,h}^r\big)^{\frac{2-\tilde{r}}{r+2-\tilde{r}}} \big(\gamma \mathrm{s}_h(\underline{\boldsymbol{u}}_h, \underline{\boldsymbol{e}}_h) - \gamma \mathrm{s}_h(\underline{\boldsymbol{w}}_h, \underline{\boldsymbol{e}}_h)\big)^{\frac{r}{r+2-\tilde{r}}}. \tag{47}
$$

Finally, combining again the norm equivalence (41a) with (46) and (47), and using (36) yields

$$
\sigma_{\rm sm}^{\frac{r}{r+2-\tilde{r}}} \|\underline{\boldsymbol{e}}_h\|_{r,h}^r \lesssim \big(\sigma_{\rm de}^r + \|\underline{\boldsymbol{u}}_h\|_{r,h}^r + \|\underline{\boldsymbol{w}}_h\|_{r,h}^r\big)^{\frac{2-\tilde{r}}{r+2-\tilde{r}}} \big(\mathrm{a}_h(\underline{\boldsymbol{u}}_h, \underline{\boldsymbol{e}}_h) - \mathrm{a}_h(\underline{\boldsymbol{w}}_h, \underline{\boldsymbol{e}}_h)\big)^{\frac{r}{r+2-\tilde{r}}}.
$$

Raising this inequality to the power $\frac{r-2-\tilde{r}}{r}$ yields (43b). □

6.2.2. *Stability of the pressure-velocity coupling*

Lemma 6.4 (Inf-sup stability of b_h)**.** *It holds, for all $q_h \in P_h^k$,*

$$
\|q_h\|_{L^{r'}(\Omega,\mathbb{R})} \lesssim \sup_{\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_{h,0}^k, \|\underline{\boldsymbol{v}}_h\|_{r,h}=1} \mathrm{b}_h(\underline{\boldsymbol{v}}_h, q_h), \tag{48}
$$

with hidden constant depending only on d, k, r, Ω, and the mesh regularity parameter.

Proof. The proof follows the classical Fortin argument (cf., e.g., [9, Section 8.4]), adapted here to the non-Hilbertian setting.

(i) *Fortin operator.* We need to prove that the following properties hold for any $\boldsymbol{v} \in W^{1,r}(\Omega,\mathbb{R}^d)$:

$$
\|\underline{\boldsymbol{I}}_h^k \boldsymbol{v}\|_{r,h} \lesssim |\boldsymbol{v}|_{W^{1,r}(\Omega,\mathbb{R}^d)}, \tag{49a}
$$

$$
\mathrm{b}_h(\underline{\boldsymbol{I}}_h^k \boldsymbol{v}, q_h) = b(\boldsymbol{v}, q_h) \qquad \forall q_h \in \mathbb{P}^k(\mathcal{T}_h, \mathbb{R}). \tag{49b}
$$

Property (49a) is obtained by raising both sides of (14) to the power r, summing over $T \in \mathcal{T}_h$, then taking the rth root of the resulting inequality. The proof of (49b) is given, e.g., in [19, Lemma 8.12].

(ii) *Inf-sup condition on* b_h. Let $q_h \in P_h^k$ and set $c_h \coloneqq \int_\Omega |q_h|^{r'-2} q_h$. Using the triangle and Hölder inequalities,

we get

$$\||q_h|^{r'-2}q_h - c_h\|_{L^r(\Omega,\mathbb{R})} \le \|q_h\|_{L^{r'}(\Omega,\mathbb{R})}^{r'-1} + |c_h||\Omega|_d^{\frac{1}{r}} \le (1+|\Omega|_d)\,\|q_h\|_{L^{r'}(\Omega,\mathbb{R})}^{r'-1} \lesssim \|q_h\|_{L^{r'}(\Omega,\mathbb{R})}^{r'-1}, \tag{50}$$

where we have used the fact that $|c_h| \le \|q_h\|_{L^{r'}(\Omega,\mathbb{R})}^{r'-1}|\Omega|_d^{\frac{1}{r'}}$ along with $\frac{1}{r}+\frac{1}{r'}=1$ in the second bound and the fact that $|\Omega|_d \lesssim 1$ to conclude. Thus, using the surjectivity of the continuous divergence operator $\boldsymbol{\nabla}\cdot : \boldsymbol{U} \to L_0^r(\Omega,\mathbb{R}) \coloneqq \{q \in L^r(\Omega,\mathbb{R}) : \int_\Omega q = 0\}$, (c.f. [27] and also [10, Theorem 1]), we infer that there exists $\boldsymbol{v}_{q_h} \in \boldsymbol{U}$ such that

$$-\boldsymbol{\nabla}\cdot\boldsymbol{v}_{q_h} = |q_h|^{r'-2}q_h - c_h \quad \text{and} \quad |\boldsymbol{v}_{q_h}|_{W^{1,r}(\Omega,\mathbb{R}^d)} \lesssim \||q_h|^{r'-2}q_h - c_h\|_{L^r(\Omega,\mathbb{R})}. \tag{51}$$

Denote by \$ the supremum in (48). Using the fact that q_h has zero mean value over Ω, the equality in (51) together with the definition (7) of b, and the second Fortin property (49b), we have

$$\|q_h\|_{L^{r'}(\Omega,\mathbb{R})}^{r'} = \int_\Omega \big(|q_h|^{r'-2}q_h - c_h\big)q_h = b(\boldsymbol{v}_{q_h}, q_h) = \mathrm{b}_h(\underline{\boldsymbol{I}}_h^k\boldsymbol{v}_{q_h}, q_h) \le \$\|\underline{\boldsymbol{I}}_h^k\boldsymbol{v}_{q_h}\|_{r,h} \lesssim \$\|q_h\|_{L^{r'}(\Omega,\mathbb{R})}^{r'-1},$$

where, to conclude, we have used (49a) followed by (51) and (50). Simplifying yields (48). □

6.2.3. *Proof of Theorem 3.6*

Proof of Theorem 3.6. (i) *Existence.* Denote by $P_h^{k,*}$ the dual space of P_h^k and let $B_h : \underline{\boldsymbol{U}}_{h,0}^k \to P_h^{k,*}$ be such that, for all $\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_{h,0}^k$,

$$\langle B_h\underline{\boldsymbol{v}}_h, q_h\rangle \coloneqq -\mathrm{b}_h(\underline{\boldsymbol{v}}_h, q_h) \qquad \forall q_h \in P_h^k.$$

Here and in what follows, $\langle\cdot,\cdot\rangle$ denotes the appropriate duality pairing as inferred from its arguments. Define the following subspace of $\underline{\boldsymbol{U}}_{h,0}^k$ spanned by vectors of discrete unknowns with zero discrete divergence:

$$\underline{\boldsymbol{W}}_h^k \coloneqq \mathrm{Ker}(B_h) = \left\{\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_{h,0}^k : \mathrm{b}_h(\underline{\boldsymbol{v}}_h, q_h) = 0 \quad \forall q_h \in P_h^k\right\}, \tag{52}$$

and consider the following problem: Find $\underline{\boldsymbol{u}}_h \in \underline{\boldsymbol{W}}_h^k$ such that

$$\mathrm{a}_h(\underline{\boldsymbol{u}}_h, \underline{\boldsymbol{v}}_h) = \int_\Omega \boldsymbol{f}\cdot\boldsymbol{v}_h \qquad \forall \underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{W}}_h^k. \tag{53}$$

Existence of a solution to this problem for a fixed h can be proved adapting the arguments of [17, Theorem 4.5]. Specifically, equip $\underline{\boldsymbol{W}}_h^k$ with an inner product $(\cdot,\cdot)_{\boldsymbol{W},h}$ (which need not be further specified), denote by $\|\cdot\|_{\boldsymbol{W},h}$ the induced norm, and let $\boldsymbol{\Phi}_h : \underline{\boldsymbol{W}}_h^k \to \underline{\boldsymbol{W}}_h^k$ be such that, for all $\underline{\boldsymbol{w}}_h \in \underline{\boldsymbol{W}}_h^k$, $(\boldsymbol{\Phi}_h(\underline{\boldsymbol{w}}_h), \underline{\boldsymbol{v}}_h)_{\boldsymbol{W},h} = \mathrm{a}_h(\underline{\boldsymbol{w}}_h, \underline{\boldsymbol{v}}_h)$ for all $\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{W}}_h^k$. The strong monotonicity (43b) of a_h yields, for any $\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{W}}_h^k$ such that $\|\underline{\boldsymbol{v}}_h\|_{r,h} \ge \sigma_{\mathrm{de}}$,

$$(\boldsymbol{\Phi}_h(\underline{\boldsymbol{v}}_h), \underline{\boldsymbol{v}}_h)_{\boldsymbol{W},h} \ge \sigma_{\mathrm{sm}}(\sigma_{\mathrm{de}}^r + \|\underline{\boldsymbol{v}}_h\|_{r,h}^r)^{\frac{\tilde r-2}{r}}\|\underline{\boldsymbol{v}}_h\|_{r,h}^{r+2-\tilde r} \gtrsim \sigma_{\mathrm{sm}}\|\underline{\boldsymbol{v}}_h\|_{r,h}^r \ge C^r\sigma_{\mathrm{sm}}\|\underline{\boldsymbol{v}}_h\|_{\boldsymbol{W},h}^r,$$

where C denotes the constant (possibly depending on h) in the equivalence of the norms $\|\cdot\|_{r,h}$ and $\|\cdot\|_{\boldsymbol{W},h}$ (which holds since $\underline{\boldsymbol{W}}_h^k$ is finite-dimensional). This shows that $\boldsymbol{\Phi}_h$ is coercive hence, by [16, Theorem 3.3], surjective. Let now $\underline{\boldsymbol{w}}_h \in \underline{\boldsymbol{W}}_h^k$ be such that $(\underline{\boldsymbol{w}}_h, \underline{\boldsymbol{v}}_h)_{\boldsymbol{W},h} = \int_\Omega \boldsymbol{f}\cdot\boldsymbol{v}_h$ for all $\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{W}}_h^k$. By the surjectivity of $\boldsymbol{\Phi}_h$, there exists $\underline{\boldsymbol{u}}_h \in \underline{\boldsymbol{W}}_h^k$ such that $\boldsymbol{\Phi}_h(\underline{\boldsymbol{u}}_h) = \underline{\boldsymbol{w}}_h$ which, by definition of $\underline{\boldsymbol{w}}_h$ and $\boldsymbol{\Phi}_h$, is a solution to the discrete problem (53).

The proof of existence now continues as in the linear case; see, e.g., [9, Theorem 4.2.1]. Denote by $\underline{\boldsymbol{U}}_{h,0}^{k,*}$ the dual space of $\underline{\boldsymbol{U}}_{h,0}^k$ and consider the linear mapping $\ell_h \in \underline{\boldsymbol{U}}_{h,0}^{k,*}$ such that, for all $\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_{h,0}^k$,

$$\langle \ell_h, \underline{\boldsymbol{v}}_h\rangle \coloneqq \int_\Omega \boldsymbol{f}\cdot\boldsymbol{v}_h - \mathrm{a}_h(\underline{\boldsymbol{u}}_h, \underline{\boldsymbol{v}}_h).$$

Thanks to (53), ℓ_h vanishes identically for every $\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{W}}_h^k$, that is to say, ℓ_h lies in the polar space of $\underline{\boldsymbol{W}}_h^k$ which, denoting by $B_h^* : P_h^k \to \underline{\boldsymbol{U}}_{h,0}^{k,*}$ the adjoint operator of B_h, coincides in our case with $\mathrm{Im}(B_h^*)$ (see, e.g., [9, Theorem 4.14]). Hence, $\ell_h \in \mathrm{Im}(B_h^*)$, and there exists therefore a $p_h \in P_h^k$ such that $B_h^* p_h = \ell_h$. This means that, for all $\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_{h,0}^k$,

$$\mathrm{b}_h(\underline{\boldsymbol{v}}_h, p_h) = \langle B_h^* p_h, \underline{\boldsymbol{v}}_h\rangle = \langle \ell_h, \underline{\boldsymbol{v}}_h\rangle = \int_\Omega \boldsymbol{f}\cdot\boldsymbol{v}_h - \mathrm{a}_h(\underline{\boldsymbol{u}}_h, \underline{\boldsymbol{v}}_h),$$

i.e., the $(\underline{\boldsymbol{u}}_h, p_h)$ satisfies the discrete momentum equation (29a). On the other hand, since $\underline{\boldsymbol{u}}_h \in \underline{\boldsymbol{W}}_h^k$, we also have, by the definition (52) of $\underline{\boldsymbol{W}}_h^k$, $\mathrm{b}_h(\underline{\boldsymbol{u}}_h, q_h) = 0$ for all $q_h \in P_h^k$, which shows that the discrete mass equation (29b) is also verified. In conclusion, $(\underline{\boldsymbol{u}}_h, p_h) \in \underline{\boldsymbol{U}}_{h,0}^k \times P_h^k$ solves (29).

(ii) *Uniqueness.* We start by proving uniqueness for the velocity. Let $(\underline{\boldsymbol{u}}_h, p_h), (\underline{\boldsymbol{u}}_h', p_h') \in \underline{\boldsymbol{U}}_{h,0}^k \times P_h^k$ be two solutions of (29). Making $\underline{\boldsymbol{v}}_h = \underline{\boldsymbol{u}}_h - \underline{\boldsymbol{u}}_h'$ in (29a) written first for $(\underline{\boldsymbol{u}}_h, p_h)$ then for $(\underline{\boldsymbol{u}}_h', p_h')$, then taking the difference and observing that $\mathrm{b}_h(\underline{\boldsymbol{u}}_h - \underline{\boldsymbol{u}}_h', p_h) = \mathrm{b}_h(\underline{\boldsymbol{u}}_h - \underline{\boldsymbol{u}}_h', p_h') = 0$ by (29b), we infer that

$$\mathrm{a}_h(\underline{\boldsymbol{u}}_h, \underline{\boldsymbol{u}}_h - \underline{\boldsymbol{u}}_h') - \mathrm{a}_h(\underline{\boldsymbol{u}}_h', \underline{\boldsymbol{u}}_h - \underline{\boldsymbol{u}}_h') = 0.$$

Thus, the strong monotonicity (43b) of a_h yields $\|\underline{\boldsymbol{u}}_h - \underline{\boldsymbol{u}}_h'\|_{r,h} = 0$, which implies $\underline{\boldsymbol{u}}_h = \underline{\boldsymbol{u}}_h'$ since $\|\cdot\|_{r,h}$ is a norm on $\underline{\boldsymbol{U}}_{h,0}^k$. Moreover, using the inf-sup stability (48) of b_h and (29a) written first for $\underline{\boldsymbol{u}}_h$ then for $\underline{\boldsymbol{u}}_h'$, we get

$$\begin{aligned}\|p_h - p_h'\|_{L^{r'}(\Omega,\mathbb{R})} &\lesssim \sup_{\underline{\boldsymbol{v}}_h\in\underline{\boldsymbol{U}}_{h,0}^k, \|\underline{\boldsymbol{v}}_h\|_{r,h}=1} \mathrm{b}_h(\underline{\boldsymbol{v}}_h, p_h - p_h')\\ &= \sup_{\underline{\boldsymbol{v}}_h\in\underline{\boldsymbol{U}}_{h,0}^k, \|\underline{\boldsymbol{v}}_h\|_{r,h}=1} (\mathrm{a}_h(\underline{\boldsymbol{u}}_h', \underline{\boldsymbol{v}}_h) - \mathrm{a}_h(\underline{\boldsymbol{u}}_h, \underline{\boldsymbol{v}}_h)) = 0,\end{aligned}$$

hence $p_h = p_h'$.

(iii) *A priori estimates.* Using the strong monotonicity (43b) of a_h (with $\underline{\boldsymbol{w}}_h = \underline{\boldsymbol{0}}$), equation (29a) together with (29b), and the Hölder inequality together with the discrete Korn inequality (37), we obtain

$$\sigma_{\mathrm{sm}}\big(\sigma_{\mathrm{de}}^r + \|\underline{\boldsymbol{u}}_h\|_{r,h}^r\big)^{\frac{\tilde r-2}{r}}\|\underline{\boldsymbol{u}}_h\|_{r,h}^{r+2-\tilde r} \lesssim \mathrm{a}_h(\underline{\boldsymbol{u}}_h, \underline{\boldsymbol{u}}_h) = \int_\Omega \boldsymbol{f}\cdot\boldsymbol{u}_h \lesssim \|\boldsymbol{f}\|_{L^{r'}(\Omega,\mathbb{R}^d)}\|\underline{\boldsymbol{u}}_h\|_{r,h}. \tag{54}$$

We then conclude as in the continuous case to infer (30a) (see Remark 1.7). To prove the bound (30b) on the pressure, we use the inf-sup stability (48) of b_h to write

$$\begin{aligned}\|p_h\|_{L^{r'}(\Omega,\mathbb{R})} &\lesssim \sup_{\underline{\boldsymbol{v}}_h\in\underline{\boldsymbol{U}}_{h,0}^k, \|\underline{\boldsymbol{v}}_h\|_{r,h}=1} \mathrm{b}_h(\underline{\boldsymbol{v}}_h, p_h)\\ &= \sup_{\underline{\boldsymbol{v}}_h\in\underline{\boldsymbol{U}}_{h,0}^k, \|\underline{\boldsymbol{v}}_h\|_{r,h}=1} \left(\int_\Omega \boldsymbol{f}\cdot\boldsymbol{v}_h - \mathrm{a}_h(\underline{\boldsymbol{u}}_h, \underline{\boldsymbol{v}}_h)\right)\\ &\lesssim \|\boldsymbol{f}\|_{L^{r'}(\Omega,\mathbb{R}^d)} + \sigma_{\mathrm{hc}}(\sigma_{\mathrm{de}}^r + \|\underline{\boldsymbol{u}}_h\|_{r,h}^r)^{\frac{r-\tilde r}{r}}\|\underline{\boldsymbol{u}}_h\|_{r,h}^{\tilde r-1}\\ &\lesssim \sigma_{\mathrm{hc}}\left(\sigma_{\mathrm{sm}}^{-1}\|\boldsymbol{f}\|_{L^{r'}(\Omega,\mathbb{R}^d)} + \sigma_{\mathrm{de}}^{|r-2|(\tilde r-1)}\left(\sigma_{\mathrm{sm}}^{-1}\|\boldsymbol{f}\|_{L^{r'}(\Omega,\mathbb{R}^d)}\right)^{\frac{\tilde r-1}{r+1-\tilde r}}\right),\end{aligned}$$

where we have used the discrete momentum equation (29a) to pass to the second line, the Hölder and discrete Korn (37) inequalities together with the Hölder continuity (43a) of a_h to pass to the third line, and the a priori bound (30a) on the velocity together with $\frac{\sigma_{\mathrm{hc}}}{\sigma_{\mathrm{sm}}} \geq 1$ (see (4)) to conclude. □

6.3. **Error estimate**

In this section, after studying the consistency of the viscous and pressure-velocity coupling terms, we prove Theorem 3.7.

6.3.1. *Consistency of the viscous function*

Lemma 6.5 (Consistency of a_h)**.** *Let $\boldsymbol{w} \in \boldsymbol{U} \cap W^{k+2,r}(\mathcal{T}_h,\mathbb{R}^d)$ be such that $\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w}) \in W^{1,r'}(\Omega,\mathbb{R}^{d\times d}_{\mathrm{s}}) \cap W^{(k+1)(\tilde{r}-1),r'}(\mathcal{T}_h,\mathbb{R}^{d\times d}_{\mathrm{s}})$. Define the viscous consistency error linear form $\mathcal{E}_{\mathrm{a},h}(\boldsymbol{w};\cdot) : \underline{\boldsymbol{U}}_h^k \to \mathbb{R}$ such that, for all $\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_h^k$,*

$$\mathcal{E}_{\mathrm{a},h}(\boldsymbol{w};\underline{\boldsymbol{v}}_h) \coloneqq \int_\Omega (\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w}))\cdot\boldsymbol{v}_h + \mathrm{a}_h(\underline{\boldsymbol{I}}_h^k\boldsymbol{w},\underline{\boldsymbol{v}}_h). \tag{55}$$

Then, under Assumptions 1.1 and 3.1, we have

$$\sup_{\underline{\boldsymbol{v}}_h\in\underline{\boldsymbol{U}}_{h,0}^k,\|\underline{\boldsymbol{v}}_h\|_{r,h}=1} \mathcal{E}_{\mathrm{a},h}(\boldsymbol{w};\underline{\boldsymbol{v}}_h) \lesssim h^{(k+1)(\tilde{r}-1)}\Bigg[\sigma_{\mathrm{hc}}\left(\sigma_{\mathrm{de}}^r + |\boldsymbol{w}|^r_{W^{1,r}(\Omega,\mathbb{R}^d)}\right)^{\frac{r-\tilde{r}}{r}} |\boldsymbol{w}|^{\tilde{r}-1}_{W^{k+2,r}(\mathcal{T}_h,\mathbb{R}^d)} + |\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w})|_{W^{(k+1)(\tilde{r}-1),r'}(\mathcal{T}_h,\mathbb{R}^{d\times d})}\Bigg]. \tag{56}$$

Proof. Let $\underline{\hat{\boldsymbol{w}}}_h \coloneqq \underline{\boldsymbol{I}}_h^k\boldsymbol{w}$ and $\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_{h,0}^k$. Expanding a_h according to its definition (18) in the expression (55) of $\mathcal{E}_{\mathrm{a},h}$, inserting $\pm\left(\int_\Omega \boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w}) : \mathbf{G}_{\mathrm{s},h}^k\underline{\boldsymbol{v}}_h + \int_\Omega \boldsymbol{\pi}_h^k\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w}) : \mathbf{G}_{\mathrm{s},h}^k\underline{\boldsymbol{v}}_h\right)$, and rearranging, we obtain

$$\begin{aligned}\mathcal{E}_{\mathrm{a},h}(\boldsymbol{w};\underline{\boldsymbol{v}}_h) = {}& \underbrace{\int_\Omega (\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w}))\cdot\boldsymbol{v}_h + \int_\Omega \boldsymbol{\pi}_h^k\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w}) : \mathbf{G}_{\mathrm{s},h}^k\underline{\boldsymbol{v}}_h}_{\mathcal{T}_1} + \cancel{\int_\Omega \left(\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w}) - \boldsymbol{\pi}_h^k\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w})\right) : \mathbf{G}_{\mathrm{s},h}^k\underline{\boldsymbol{v}}_h} \\ & + \underbrace{\int_\Omega \left(\boldsymbol{\sigma}(\cdot,\mathbf{G}_{\mathrm{s},h}^k\underline{\hat{\boldsymbol{w}}}_h) - \boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w})\right) : \mathbf{G}_{\mathrm{s},h}^k\underline{\boldsymbol{v}}_h}_{\mathcal{T}_2} + \underbrace{\gamma\mathrm{s}_h(\underline{\hat{\boldsymbol{w}}}_h,\underline{\boldsymbol{v}}_h)}_{\mathcal{T}_3},\end{aligned} \tag{57}$$

where have used the definition (10) of $\boldsymbol{\pi}_h^k$ together with the fact that $\mathbf{G}_{\mathrm{s},h}^k\underline{\boldsymbol{v}}_h \in \mathbb{P}^k(\mathcal{T}_h,\mathbb{R}^{d\times d}_{\mathrm{s}})$ in the cancellation. We proceed to estimate the terms in the right-hand side. For the first term, we start by noticing that

$$\sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T}\int_F \boldsymbol{v}_F\cdot(\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w})\boldsymbol{n}_{TF}) = 0 \tag{58}$$

as a consequence of the continuity of the normal trace of $\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w})$ together with the single-valuedness of $\boldsymbol{v}_F$ across each interface $F \in \mathcal{F}_h^{\mathrm{i}}$ and of the fact that $\boldsymbol{v}_F = \mathbf{0}$ for every boundary face $F \in \mathcal{F}_h^{\mathrm{b}}$. Using an element by element integration by parts on the first term of $\mathcal{T}_1$ along with the definitions (17) of $\mathbf{G}_{\mathrm{s},h}^k$ and (15) of $\mathbf{G}_{\mathrm{s},T}^k$, we can write

$$\begin{aligned}\mathcal{T}_1 = {}& \cancel{\int_\Omega \left(\boldsymbol{\pi}_h^k\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w}) - \boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w})\right) : \boldsymbol{\nabla}_{\mathrm{s},h}\boldsymbol{v}_h} \\ & + \sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T}\left(\int_F(\boldsymbol{v}_F-\boldsymbol{v}_T)\cdot(\boldsymbol{\pi}_T^k\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w}))\boldsymbol{n}_{TF} + \int_F \boldsymbol{v}_T\cdot(\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w})\boldsymbol{n}_{TF})\right) \\ = {}& \sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T}\int_F(\boldsymbol{v}_F-\boldsymbol{v}_T)\cdot\left(\boldsymbol{\pi}_T^k\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w}) - \boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w})\right)\boldsymbol{n}_{TF},\end{aligned}$$

where we have used the definition (10) of $\boldsymbol{\pi}_h^k$ together with the fact that $\boldsymbol{\nabla}_{\mathrm{s},h}\boldsymbol{v}_h \in \mathbb{P}^{k-1}(\mathcal{T}_h,\mathbb{R}^{d\times d}_{\mathrm{s}}) \subset \mathbb{P}^k(\mathcal{T}_h,\mathbb{R}^{d\times d}_{\mathrm{s}})$ to cancel the term in the first line, and we have inserted (58) and rearranged to conclude. Therefore, applying

the Hölder inequality together with the bound $h_F \le h_T$, we infer

$$\begin{aligned}|\mathcal{T}_1| &\le \left(\sum_{T\in\mathcal{T}_h} h_T\|\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w}) - \boldsymbol{\pi}_T^k\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w})\|_{L^{r'}(\partial T,\mathbb{R}^{d\times d})}^{r'}\right)^{\frac{1}{r'}} \left(\sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T} h_F^{1-r}\|\boldsymbol{v}_F - \boldsymbol{v}_T\|_{L^r(F,\mathbb{R}^d)}^r\right)^{\frac{1}{r}} \\ &\lesssim h^{(k+1)(\tilde r-1)}|\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w})|_{W^{(k+1)(\tilde r-1),r'}(\mathcal{T}_h,\mathbb{R}^{d\times d})}\|\underline{\boldsymbol{v}}_h\|_{r,h},\end{aligned} \tag{59}$$

where the conclusion follows using the $((k+1)(\tilde r-1),r')$-trace approximation properties (11b) of $\boldsymbol{\pi}_T^k$ along with $h_T \le h$ for the first factor and the definition (13) of the $\|\cdot\|_{r,h}$-norm for the second.

For the second term, using the Hölder inequality and again (41a), we get

$$|\mathcal{T}_2| \le \|\boldsymbol{\sigma}(\cdot,\mathbf{G}_{\mathrm{s},h}^k\underline{\hat{\boldsymbol{w}}}_h) - \boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w})\|_{L^{r'}(\Omega,\mathbb{R}^{d\times d})}\|\underline{\boldsymbol{v}}_h\|_{r,h}. \tag{60}$$

We estimate the first factor as follows:

$$\begin{aligned}&\|\boldsymbol{\sigma}(\cdot,\mathbf{G}_{\mathrm{s},h}^k\underline{\hat{\boldsymbol{w}}}_h) - \boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w})\|_{L^{r'}(\Omega,\mathbb{R}^{d\times d})} \\ &\quad\le \sigma_{\mathrm{hc}}\left\|\left(\sigma_{\mathrm{de}}^r + |\mathbf{G}_{\mathrm{s},h}^k\underline{\hat{\boldsymbol{w}}}_h|_{d\times d}^r + |\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w}|_{d\times d}^r\right)^{\frac{r-\tilde r}{r}}|\mathbf{G}_{\mathrm{s},h}^k\underline{\hat{\boldsymbol{w}}}_h - \boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w}|_{d\times d}^{\tilde r-1}\right\|_{L^{r'}(\Omega,\mathbb{R})} \\ &\quad\lesssim \sigma_{\mathrm{hc}}\left(\sigma_{\mathrm{de}}^r + \|\mathbf{G}_{\mathrm{s},h}^k\underline{\hat{\boldsymbol{w}}}_h\|_{L^r(\Omega,\mathbb{R}^{d\times d})}^r + \|\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w}\|_{L^r(\Omega,\mathbb{R}^{d\times d})}^r\right)^{\frac{r-\tilde r}{r}}\|\mathbf{G}_{\mathrm{s},h}^k\underline{\hat{\boldsymbol{w}}}_h - \boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w}\|_{L^r(\Omega,\mathbb{R}^{d\times d})}^{\tilde r-1} \\ &\quad\lesssim \sigma_{\mathrm{hc}}\left(\sigma_{\mathrm{de}}^r + \|\underline{\hat{\boldsymbol{w}}}_h\|_{r,h}^r + |\boldsymbol{w}|_{W^{1,r}(\Omega,\mathbb{R}^d)}^r\right)^{\frac{r-\tilde r}{r}}\|\boldsymbol{\pi}_h^k(\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w}) - \boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{w}\|_{L^r(\Omega,\mathbb{R}^{d\times d})}^{\tilde r-1} \\ &\quad\lesssim h^{(k+1)(\tilde r-1)}\sigma_{\mathrm{hc}}\left(\sigma_{\mathrm{de}}^r + |\boldsymbol{w}|_{W^{1,r}(\Omega,\mathbb{R}^d)}^r\right)^{\frac{r-\tilde r}{r}}|\boldsymbol{w}|_{W^{k+2,r}(\mathcal{T}_h,\mathbb{R}^d)}^{\tilde r-1},\end{aligned}$$

where we have used the Hölder continuity (3c) of $\boldsymbol{\sigma}$ in the first bound, the $(r';\frac{r}{r-\tilde r},\frac{r}{\tilde r-1})$-Hölder inequality (39) in the second, the boundedness of Ω along with (41a) and the commutation property (16) of $\mathbf{G}_{\mathrm{s},h}^k$ in the third, and we have concluded invoking the $(k+1,r,0)$-approximation property (11a) of $\boldsymbol{\pi}_T^k$. Plugging this estimate into (60), we get

$$|\mathcal{T}_2| \lesssim h^{(k+1)(\tilde r-1)}\sigma_{\mathrm{hc}}\left(\sigma_{\mathrm{de}}^r + |\boldsymbol{w}|_{W^{1,r}(\Omega,\mathbb{R}^d)}^r\right)^{\frac{r-\tilde r}{r}}|\boldsymbol{w}|_{W^{k+2,r}(\mathcal{T}_h,\mathbb{R}^d)}^{\tilde r-1}\|\underline{\boldsymbol{v}}_h\|_{r,h}. \tag{61}$$

Finally, using the fact that $\gamma \le \sigma_{\mathrm{hc}}$ together with the consistency (42) of s_h and the norm equivalence (41a), we obtain for the third term

$$|\mathcal{T}_3| \lesssim h^{(k+1)(\tilde r-1)}\sigma_{\mathrm{hc}}|\boldsymbol{w}|_{W^{1,r}(\Omega,\mathbb{R}^d)}^{r-\tilde r}|\boldsymbol{w}|_{W^{k+2,r}(\mathcal{T}_h,\mathbb{R}^d)}^{\tilde r-1}\|\underline{\boldsymbol{v}}_h\|_{r,h}. \tag{62}$$

Plug the bounds (59), (61), and (62) into (57) and pass to the supremum to conclude. □

6.3.2. *Consistency of the pressure-velocity coupling bilinear form*

Lemma 6.6 (Consistency of b_h)**.** *Let $q \in W^{1,r'}(\Omega,\mathbb{R}) \cap W^{(k+1)(\tilde r-1),r'}(\mathcal{T}_h,\mathbb{R})$. Let $\mathcal{E}_{\mathrm{b},h}(q;\cdot) : \underline{\boldsymbol{U}}_h^k \to \mathbb{R}$ be the pressure consistency error linear form such that, for all $\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_h^k$,*

$$\mathcal{E}_{\mathrm{b},h}(q;\underline{\boldsymbol{v}}_h) \coloneqq \int_\Omega \boldsymbol{\nabla}q\cdot\boldsymbol{v}_h - \mathrm{b}_h(\underline{\boldsymbol{v}}_h,\pi_h^k q). \tag{63}$$

Then, we have that

$$\sup_{\underline{\boldsymbol{v}}_h\in\underline{\boldsymbol{U}}_{h,0}^k,\|\underline{\boldsymbol{v}}_h\|_{r,h}=1}\mathcal{E}_{\mathrm{b},h}(q;\underline{\boldsymbol{v}}_h) \lesssim h^{(k+1)(\tilde r-1)}|q|_{W^{(k+1)(\tilde r-1),r'}(\mathcal{T}_h,\mathbb{R})}. \tag{64}$$

Proof. Let $\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_{h,0}^k$. Integrating by parts element by element, we can reformulate the first term in the right-hand side of (63) as follows:

$$\int_\Omega \boldsymbol{\nabla} q \cdot \boldsymbol{v}_h = -\sum_{T\in\mathcal{T}_h}\left(\int_T q(\boldsymbol{\nabla}\cdot\boldsymbol{v}_T) + \sum_{F\in\mathcal{F}_T}\int_F q(\boldsymbol{v}_F-\boldsymbol{v}_T)\cdot\boldsymbol{n}_{TF}\right), \tag{65}$$

where the introduction of $\boldsymbol{v}_F$ in the boundary term is justified by the fact that the jumps of q vanish across interfaces by the assumed regularity and that $\boldsymbol{v}_F = \boldsymbol{0}$ on every boundary face $F \in \mathcal{F}_h^{\mathrm{b}}$. On the other hand, expanding, for each $T \in \mathcal{T}_h$, D_T^k according to its definition (27), we get

$$-\mathrm{b}_h(\underline{\boldsymbol{v}}_h, \pi_h^k q) = \sum_{T\in\mathcal{T}_h}\left(\int_T \pi_T^k q\ (\boldsymbol{\nabla}\cdot\boldsymbol{v}_T) + \sum_{F\in\mathcal{F}_T}\int_F \pi_T^k q\ (\boldsymbol{v}_F-\boldsymbol{v}_T)\cdot\boldsymbol{n}_{TF}\right). \tag{66}$$

Summing (65) and (66) and observing that the first terms in parentheses cancel out by the definition (10) of π_T^k since $\boldsymbol{\nabla}\cdot\boldsymbol{v}_T \in \mathbb{P}^{k-1}(T,\mathbb{R}) \subset \mathbb{P}^k(T,\mathbb{R})$ for all $T \in \mathcal{T}_h$, we can write

$$\begin{aligned}
\mathcal{E}_{\mathrm{b},h}(q;\underline{\boldsymbol{v}}_h) &= \sum_{T\in\mathcal{T}_h}\left(\cancel{\int_T (\pi_T^k q - q)(\boldsymbol{\nabla}\cdot\boldsymbol{v}_T)} + \sum_{F\in\mathcal{F}_T}\int_F (\pi_T^k q - q)(\boldsymbol{v}_F-\boldsymbol{v}_T)\cdot\boldsymbol{n}_{TF}\right)\\
&\le \left(\sum_{T\in\mathcal{T}_h} h_T \|\pi_T^k q - q\|_{L^{r'}(\partial T,\mathbb{R})}^{r'}\right)^{\frac{1}{r'}}\left(\sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T} h_F^{1-r}\|\boldsymbol{v}_F-\boldsymbol{v}_T\|_{L^r(F,\mathbb{R}^d)}^r\right)^{\frac{1}{r}}\\
&\lesssim h^{(k+1)(\tilde r-1)}|q|_{W^{(k+1)(\tilde r-1),r'}(\mathcal{T}_h,\mathbb{R})}\|\underline{\boldsymbol{v}}_h\|_{r,h},
\end{aligned}$$

where we have used the Hölder inequality along with $h_F \ge h_T$ whenever $F \in \mathcal{F}_T$ in the second line and the $((k+1)(\tilde r-1), r')$-trace approximation property (11b) of π_T^k together with the bound $h_F \le h$ and the definition (13) of the $\|\cdot\|_{r,h}$-norm to conclude. Passing to the supremum yields (64). □

6.3.3. *Proof of Theorem 3.7*

Proof of Theorem 3.7. Let $(\underline{\boldsymbol{e}}_h, \epsilon_h) \coloneqq (\underline{\boldsymbol{u}}_h - \underline{\hat{\boldsymbol{u}}}_h, p_h - \hat p_h) \in \underline{\boldsymbol{U}}_{h,0}^k \times P_h^k$ where $\underline{\hat{\boldsymbol{u}}}_h \coloneqq \underline{\boldsymbol{I}}_h^k \boldsymbol{u} \in \underline{\boldsymbol{U}}_{h,0}^k$ and $\hat p_h \coloneqq \pi_h^k p \in P_h^k$.

Step 1. *Consistency error.* Let $\mathcal{E}_h : \underline{\boldsymbol{U}}_{h,0}^k \to \mathbb{R}$ be the consistency error linear form such that, for all $\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_{h,0}^k$,

$$\mathcal{E}_h(\underline{\boldsymbol{v}}_h) \coloneqq \int_\Omega \boldsymbol{f}\cdot\boldsymbol{v}_h - \mathrm{a}_h(\underline{\hat{\boldsymbol{u}}}_h, \underline{\boldsymbol{v}}_h) - \mathrm{b}_h(\underline{\boldsymbol{v}}_h, \hat p_h). \tag{67}$$

Using in the above expression the fact that $\boldsymbol{f} = -\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}(\cdot,\boldsymbol{\nabla}_{\mathrm{s}}\boldsymbol{u}) + \boldsymbol{\nabla} p$ almost everywhere in Ω to write $\mathcal{E}_h(\underline{\boldsymbol{v}}_h) = \mathcal{E}_{\mathrm{a},h}(\boldsymbol{u};\underline{\boldsymbol{v}}_h) + \mathcal{E}_{\mathrm{b},h}(p;\underline{\boldsymbol{v}}_h)$, and invoking the consistency properties (56) of a_h and (64) of b_h, we obtain

$$\$ \coloneqq \sup_{\underline{\boldsymbol{v}}_h\in\underline{\boldsymbol{U}}_{h,0}^k, \|\underline{\boldsymbol{v}}_h\|_{r,h}=1} \mathcal{E}_h(\underline{\boldsymbol{v}}_h) \lesssim h^{(k+1)(\tilde r-1)}\mathcal{N}_{\boldsymbol{\sigma},\boldsymbol{u},p}. \tag{68}$$

Step 2. *Error estimate for the velocity.* Using the strong monotonicity (43b) of a_h, we get

$$\begin{aligned}
\|\underline{\boldsymbol{e}}_h\|_{r,h}^{r+2-\tilde r} &\lesssim \sigma_{\mathrm{sm}}^{-1}\left(\sigma_{\mathrm{de}}^r + \|\underline{\boldsymbol{u}}_h\|_{r,h}^r + \|\underline{\hat{\boldsymbol{u}}}_h\|_{r,h}^r\right)^{\frac{2-\tilde r}{r}}\left(\mathrm{a}_h(\underline{\boldsymbol{u}}_h,\underline{\boldsymbol{e}}_h) - \mathrm{a}_h(\underline{\hat{\boldsymbol{u}}}_h,\underline{\boldsymbol{e}}_h)\right)\\
&\lesssim \sigma_{\mathrm{sm}}^{-1}\mathcal{N}_{\boldsymbol{f}}^{2-\tilde r}\left(\mathrm{a}_h(\underline{\boldsymbol{u}}_h,\underline{\boldsymbol{e}}_h) - \mathrm{a}_h(\underline{\hat{\boldsymbol{u}}}_h,\underline{\boldsymbol{e}}_h)\right),
\end{aligned} \tag{69}$$

where we have used the a priori bound (30a) on the discrete solution along with the boundedness (49a) of the global interpolator and the a priori bound (8) on the continuous solution to conclude. Using then the discrete

mass equation (29b) along with (49b) (written for $\boldsymbol{v} = \boldsymbol{u}$) and the continuous mass equation (6b) to write $\mathrm{b}_h(\underline{\boldsymbol{I}}_h^k \boldsymbol{u}, q_h) = b(\boldsymbol{u}, q_h) = 0$, we get $\mathrm{b}_h(\underline{\boldsymbol{e}}_h, q_h) = 0$ for all $q_h \in P_h^k$. Hence, combining this result with (67) and the discrete momentum equation (29a) (with $\underline{\boldsymbol{v}}_h = \underline{\boldsymbol{e}}_h$), we obtain

$$\mathrm{a}_h(\underline{\boldsymbol{u}}_h, \underline{\boldsymbol{e}}_h) - \mathrm{a}_h(\widehat{\underline{\boldsymbol{u}}}_h, \underline{\boldsymbol{e}}_h) = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{e}_h - \mathrm{a}_h(\widehat{\underline{\boldsymbol{u}}}_h, \underline{\boldsymbol{e}}_h) - \cancel{\mathrm{b}_h(\underline{\boldsymbol{e}}_h, p_h)} = \mathcal{E}_h(\underline{\boldsymbol{e}}_h). \tag{70}$$

Plugging (70) into (69), we get

$$\|\underline{\boldsymbol{e}}_h\|_{r,h}^{r+2-\tilde{r}} \le \sigma_{\mathrm{sm}}^{-1} \mathcal{N}_{\boldsymbol{f}}^{2-\tilde{r}} \$ \|\underline{\boldsymbol{e}}_h\|_{r,h}.$$

Simplifying, using (68), and taking the $(r+1-\tilde{r})$th root of the resulting inequality yields (31a).

Step 3. *Error estimate for the pressure.* Using the Hölder continuity (43a) of a_h, we have, for all $\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_{h,0}^k$,

$$\begin{aligned} |\mathrm{a}_h(\widehat{\underline{\boldsymbol{u}}}_h, \underline{\boldsymbol{v}}_h) - \mathrm{a}_h(\underline{\boldsymbol{u}}_h, \underline{\boldsymbol{v}}_h)| &\lesssim \sigma_{\mathrm{hc}} \left(\sigma_{\mathrm{de}}^r + \|\widehat{\underline{\boldsymbol{u}}}_h\|_{r,h}^r + \|\underline{\boldsymbol{u}}_h\|_{r,h}^r\right)^{\frac{r-\tilde{r}}{r}} \|\underline{\boldsymbol{e}}_h\|_{r,h}^{\tilde{r}-1} \|\underline{\boldsymbol{v}}_h\|_{r,h} \\ &\lesssim \sigma_{\mathrm{hc}} \mathcal{N}_{\boldsymbol{f}}^{r-\tilde{r}} \|\underline{\boldsymbol{e}}_h\|_{r,h}^{\tilde{r}-1} \|\underline{\boldsymbol{v}}_h\|_{r,h}, \end{aligned} \tag{71}$$

where the first factor is estimated as in (69). Thus, using the inf-sup condition (48), we can write

$$\begin{aligned} \|\epsilon_h\|_{L^{r'}(\Omega,\mathbb{R})} &\lesssim \sup_{\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_{h,0}^k, \|\underline{\boldsymbol{v}}_h\|_{r,h}=1} \mathrm{b}_h(\underline{\boldsymbol{v}}_h, \epsilon_h) \\ &= \sup_{\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_{h,0}^k, \|\underline{\boldsymbol{v}}_h\|_{r,h}=1} \left(\mathcal{E}_h(\underline{\boldsymbol{v}}_h) + \mathrm{a}_h(\widehat{\underline{\boldsymbol{u}}}_h, \underline{\boldsymbol{v}}_h) - \mathrm{a}_h(\underline{\boldsymbol{u}}_h, \underline{\boldsymbol{v}}_h)\right) \\ &\lesssim \$ + \sigma_{\mathrm{hc}} \mathcal{N}_{\boldsymbol{f}}^{r-\tilde{r}} \|\underline{\boldsymbol{e}}_h\|_{r,h}^{\tilde{r}-1} \\ &\lesssim h^{(k+1)(\tilde{r}-1)} \mathcal{N}_{\boldsymbol{\sigma},\boldsymbol{u},p} + h^{(k+1)(\tilde{r}-1)^2} \sigma_{\mathrm{hc}} \mathcal{N}_{\boldsymbol{f}}^{|r-2|(\tilde{r}-1)} \left(\sigma_{\mathrm{sm}}^{-1} \mathcal{N}_{\boldsymbol{\sigma},\boldsymbol{u},p}\right)^{\frac{\tilde{r}-1}{r+1-\tilde{r}}}, \end{aligned} \tag{72}$$

where we have used the definition (67) of the consistency error together with equation (29a) to pass to the second line, (71) to pass to the third line (recall that \$ denotes here the supremum in the left-hand side of (68)), and the bounds (68) and (31a) (proved in Step 2) to conclude. □

Appendix A. Power-framed functions

In the following theorem, we introduce the notion of power-framed function and discuss sufficient conditions for this property to hold.

Theorem A.1 (Power-framed function)**.** *Let U be a measurable subset of $\mathbb{R}^n$ with $n \ge 1$, $(W, (\cdot,\cdot)_W)$ an inner product space, and $\boldsymbol{\sigma} : U \times W \to W$. Assume that there exists a Carathéodory function $\varsigma : U \times [0,\infty) \to \mathbb{R}$ such that, for all $\boldsymbol{\tau} \in W$ and almost every $\boldsymbol{x} \in U$,*

$$\boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{\tau}) = \varsigma(\boldsymbol{x}, \|\boldsymbol{\tau}\|_W)\boldsymbol{\tau}, \tag{73a}$$

where $\|\cdot\|_W$ is the norm induced by $(\cdot,\cdot)_W$. Additionally assume that, for almost every $\boldsymbol{x} \in U$, $\varsigma(\boldsymbol{x},\cdot)$ is differentiable on $(0,\infty)$ and there exist $\varsigma_{\mathrm{de}} \in [0,\infty)$ and $\varsigma_{\mathrm{sm}}, \varsigma_{\mathrm{hc}} \in (0,\infty)$ independent of $\boldsymbol{x}$ such that, for all $\alpha \in (0,\infty)$,

$$\varsigma_{\mathrm{sm}} (\varsigma_{\mathrm{de}}^r + \alpha^r)^{\frac{r-2}{r}} \le \frac{\partial(\alpha\varsigma(\boldsymbol{x},\alpha))}{\partial\alpha} \le \varsigma_{\mathrm{hc}} (\varsigma_{\mathrm{de}}^r + \alpha^r)^{\frac{r-2}{r}}. \tag{73b}$$

*Then, $\boldsymbol{\sigma}$ is an r-*power-framed function*, i.e., for all $(\boldsymbol{\tau}, \boldsymbol{\eta}) \in W^2$ with $\boldsymbol{\tau} \ne \boldsymbol{\eta}$ and almost every $\boldsymbol{x} \in U$, the function $\boldsymbol{\sigma}$ verifies the Hölder continuity property*

$$\|\boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{\tau}) - \boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{\eta})\|_W \le \sigma_{\mathrm{hc}} \left(\sigma_{\mathrm{de}}^r + \|\boldsymbol{\tau}\|_W^r + \|\boldsymbol{\eta}\|_W^r\right)^{\frac{r-2}{r}} \|\boldsymbol{\tau} - \boldsymbol{\eta}\|_W, \tag{74a}$$

and the strong monotonicity property

$$(\boldsymbol{\sigma}(\boldsymbol{x},\boldsymbol{\tau})-\boldsymbol{\sigma}(\boldsymbol{x},\boldsymbol{\eta}),\boldsymbol{\tau}-\boldsymbol{\eta})_W \geq \sigma_{\mathrm{sm}}\left(\sigma_{\mathrm{de}}^r+\|\boldsymbol{\tau}\|_W^r+\|\boldsymbol{\eta}\|_W^r\right)^{\frac{r-2}{r}}\|\boldsymbol{\tau}-\boldsymbol{\eta}\|_W^2, \tag{74b}$$

with $\sigma_{\mathrm{de}} \coloneqq \varsigma_{\mathrm{de}}$, $\sigma_{\mathrm{hc}} \coloneqq 2^{2-\tilde{r}+r^{-1}\lceil 2-\tilde{r}\rceil}(\tilde{r}-1)^{-1}\varsigma_{\mathrm{hc}}$, *and* $\sigma_{\mathrm{sm}} \coloneqq 2^{\tilde{r}-r-\lceil r^{-1}(r-\tilde{r})\rceil}(r+1-\tilde{r})^{-1}\varsigma_{\mathrm{sm}}$, *where* $\tilde{r}$ *is given by* (2) *and* $\lceil\cdot\rceil$ *is the ceiling function.*

Remark A.2 (Notation)**.** The boldface notation for the elements of W is reminiscent of the fact that Theorem A.1 is used with $W=\mathbb{R}^{d\times d}_{\mathrm{s}}$ in Corollary A.3 to characterize the Carreau-Yasuda law as an r-power-framed function and in Lemma 3.3 with $W=\mathbb{R}^d$ to study the local stabilization function s_T.

Proof of Theorem A.1. Let $\boldsymbol{x}\in U$ be such that (73) holds, and $\boldsymbol{\tau},\boldsymbol{\eta}\in W$. By symmetry of inequalities (74) and the fact that $\boldsymbol{\sigma}$ is continuous, we can assume, without loss of generality, that $\|\boldsymbol{\tau}\|_W>\|\boldsymbol{\eta}\|_W>0$.

(i) *Strong monotonicity.* Let $\beta\in(0,\infty)$ and let $g:[\beta,\infty)\to\mathbb{R}$ be such that, for all $\alpha\in[\beta,\infty)$,

$$g(\alpha) \coloneqq \alpha\varsigma(\boldsymbol{x},\alpha)-\beta\varsigma(\boldsymbol{x},\beta)-C_{\mathrm{sm}}(\varsigma_{\mathrm{de}}^r+\alpha^r+\beta^r)^{\frac{r-2}{r}}(\alpha-\beta), \quad \text{with} \quad C_{\mathrm{sm}} \coloneqq \tfrac{2^{\tilde{r}-r}}{r+1-\tilde{r}}\varsigma_{\mathrm{sm}}.$$

Differentiating g and using the first inequality in (73b), we obtain, for all $\alpha\in[\beta,\infty)$,

$$\begin{aligned}\frac{\partial}{\partial\alpha}g(\alpha) &\geq \varsigma_{\mathrm{sm}}(\varsigma_{\mathrm{de}}^r+\alpha^r)^{\frac{r-2}{r}}-C_{\mathrm{sm}}\left((r-2)(\varsigma_{\mathrm{de}}^r+\alpha^r+\beta^r)^{-\frac{2}{r}}(\alpha-\beta)\alpha^{r-1}+(\varsigma_{\mathrm{de}}^r+\alpha^r+\beta^r)^{\frac{r-2}{r}}\right)\\ &\geq \varsigma_{\mathrm{sm}}(\varsigma_{\mathrm{de}}^r+\alpha^r)^{\frac{r-2}{r}}-(r+1-\tilde{r})C_{\mathrm{sm}}(\varsigma_{\mathrm{de}}^r+\alpha^r+\beta^r)^{\frac{r-2}{r}}\\ &\geq \varsigma_{\mathrm{sm}}2^{\tilde{r}-r}(\varsigma_{\mathrm{de}}^r+\alpha^r+\beta^r)^{\frac{r-2}{r}}-(r+1-\tilde{r})C_{\mathrm{sm}}(\varsigma_{\mathrm{de}}^r+\alpha^r+\beta^r)^{\frac{r-2}{r}}=0,\end{aligned}$$

where, to pass to the second line, we have removed negative contributions if $r<2$ and used the fact that $(\alpha-\beta)\alpha^{r-1}\leq\varsigma_{\mathrm{de}}^r+\alpha^r+\beta^r$ if $r\geq 2$, to pass to the third line we have used the fact that $t\mapsto t^{r-2}$ is non-increasing if $r<2$, and the fact that $\beta\leq\alpha$ otherwise, while the conclusion follows from the definition of C_{sm}. This shows that g is non-decreasing. Hence, for all $\alpha\in[\beta,\infty)$, $g(\alpha)\geq g(\beta)=0$, i.e.

$$\alpha\varsigma(\boldsymbol{x},\alpha)-\beta\varsigma(\boldsymbol{x},\beta)\geq C_{\mathrm{sm}}(\varsigma_{\mathrm{de}}^r+\alpha^r+\beta^r)^{\frac{r-2}{r}}(\alpha-\beta). \tag{75}$$

Moreover, for all $\alpha,\beta\in(0,\infty)$, using (75) (with $\beta=0$) along with the fact that $t\mapsto t^{r-2}$ is decreasing if $r<2$ and inequality (36) if $r\geq 2$, we infer that

$$\varsigma(\boldsymbol{x},\alpha)+\varsigma(\boldsymbol{x},\beta)\geq C_{\mathrm{sm}}\left((\varsigma_{\mathrm{de}}^r+\alpha^r)^{\frac{r-2}{r}}+(\varsigma_{\mathrm{de}}^r+\beta^r)^{\frac{r-2}{r}}\right)\geq C_{\mathrm{sm}}2^{1-\lceil\frac{r-\tilde{r}}{r}\rceil}(\varsigma_{\mathrm{de}}^r+\alpha^r+\beta^r)^{\frac{r-2}{r}}. \tag{76}$$

We conclude that $\boldsymbol{\sigma}$ verifies (74b) by using (75) and (76) with $\alpha=\|\boldsymbol{\tau}\|_W$ and $\beta=\|\boldsymbol{\eta}\|_W$ as follows:

$$\begin{aligned}&(\boldsymbol{\sigma}(\boldsymbol{x},\boldsymbol{\tau})-\boldsymbol{\sigma}(\boldsymbol{x},\boldsymbol{\eta}),\boldsymbol{\tau}-\boldsymbol{\eta})_W\\ &\quad=(\boldsymbol{\tau}\varsigma(\boldsymbol{x},\|\boldsymbol{\tau}\|_W)-\boldsymbol{\eta}\varsigma(\boldsymbol{x},\|\boldsymbol{\eta}\|_W),\boldsymbol{\tau}-\boldsymbol{\eta})_W\\ &\quad=\|\boldsymbol{\tau}\|_W^2\varsigma(\boldsymbol{x},\|\boldsymbol{\tau}\|_W)+\|\boldsymbol{\eta}\|_W^2\varsigma(\boldsymbol{x},\|\boldsymbol{\eta}\|_W)-(\boldsymbol{\tau},\boldsymbol{\eta})_W\left[\varsigma(\boldsymbol{x},\|\boldsymbol{\tau}\|_W)+\varsigma(\boldsymbol{x},\|\boldsymbol{\eta}\|_W)\right]\\ &\quad=\left[\|\boldsymbol{\tau}\|_W\varsigma(\boldsymbol{x},\|\boldsymbol{\tau}\|_W)-\|\boldsymbol{\eta}\|_W\varsigma(\boldsymbol{x},\|\boldsymbol{\eta}\|_W)\right](\|\boldsymbol{\tau}\|_W-\|\boldsymbol{\eta}\|_W)\\ &\qquad+\left[\varsigma(\boldsymbol{x},\|\boldsymbol{\tau}\|_W)+\varsigma(\boldsymbol{x},\|\boldsymbol{\eta}\|_W)\right](\|\boldsymbol{\tau}\|_W\|\boldsymbol{\eta}\|_W-(\boldsymbol{\tau},\boldsymbol{\eta})_W)\\ &\quad\geq C_{\mathrm{sm}}2^{-\lceil\frac{r-\tilde{r}}{r}\rceil}(\varsigma_{\mathrm{de}}^r+\|\boldsymbol{\tau}\|_W^r+\|\boldsymbol{\eta}\|_W^r)^{\frac{r-2}{r}}\left[(\|\boldsymbol{\tau}\|_W-\|\boldsymbol{\eta}\|_W)^2+2(\|\boldsymbol{\tau}\|_W\|\boldsymbol{\eta}\|_W-(\boldsymbol{\tau},\boldsymbol{\eta})_W)\right]\\ &\quad=C_{\mathrm{sm}}2^{-\lceil\frac{r-\tilde{r}}{r}\rceil}(\varsigma_{\mathrm{de}}^r+\|\boldsymbol{\tau}\|_W^r+\|\boldsymbol{\eta}\|_W^r)^{\frac{r-2}{r}}\|\boldsymbol{\tau}-\boldsymbol{\eta}\|_W^2.\end{aligned}$$

(ii) *Hölder continuity.* Now, setting $C_{\mathrm{hc}} \coloneqq \frac{\varsigma_{\mathrm{hc}}}{\tilde{r}-1}$ and reasoning in a similar way as for the proof of (75) to leverage the second inequality in (73b), we have, for all $\alpha \in [\beta, \infty)$,

$$\alpha\varsigma(\boldsymbol{x}, \alpha) - \beta\varsigma(\boldsymbol{x}, \beta) \leq C_{\mathrm{hc}} \left(\varsigma_{\mathrm{de}}^r + \alpha^r + \beta^r\right)^{\frac{r-2}{r}} (\alpha - \beta). \tag{77}$$

First, let $r \geq 2$. Using (77) (with $\beta = 0$) and the fact that $t \mapsto t^{r-2}$ is non-decreasing, we have, for all $\alpha, \beta \in (0, \infty)$,

$$\varsigma(\boldsymbol{x}, \alpha)\varsigma(\boldsymbol{x}, \beta) \leq C_{\mathrm{hc}}^2 \left(\varsigma_{\mathrm{de}}^r + \alpha^r\right)^{\frac{r-2}{r}} \left(\varsigma_{\mathrm{de}}^r + \beta^r\right)^{\frac{r-2}{r}} \leq \left[C_{\mathrm{hc}} \left(\varsigma_{\mathrm{de}}^r + \alpha^r + \beta^r\right)^{\frac{r-2}{r}}\right]^2. \tag{78}$$

Thus, using inequalities (77) and (78) with $\alpha = \|\boldsymbol{\tau}\|_W$ and $\beta = \|\boldsymbol{\eta}\|_W$, we infer

$$\begin{aligned}
&\|\boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{\tau}) - \boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{\eta})\|_W^2 \\
&\quad = (\boldsymbol{\tau}\varsigma(\boldsymbol{x}, \|\boldsymbol{\tau}\|_W) - \boldsymbol{\eta}\varsigma(\boldsymbol{x}, \|\boldsymbol{\eta}\|_W), \boldsymbol{\tau}\varsigma(\boldsymbol{x}, \|\boldsymbol{\tau}\|_W) - \boldsymbol{\eta}\varsigma(\boldsymbol{x}, \|\boldsymbol{\eta}\|_W))_W \\
&\quad = \left[\|\boldsymbol{\tau}\|_W\varsigma(\boldsymbol{x}, \|\boldsymbol{\tau}\|_W) - \|\boldsymbol{\eta}\|_W\varsigma(\boldsymbol{x}, \|\boldsymbol{\eta}\|_W)\right]^2 \\
&\qquad + 2\varsigma(\boldsymbol{x}, \|\boldsymbol{\tau}\|_W)\varsigma(\boldsymbol{x}, \|\boldsymbol{\eta}\|_W) \left[\|\boldsymbol{\tau}\|_W\|\boldsymbol{\eta}\|_W - (\boldsymbol{\tau}, \boldsymbol{\eta})_W\right] \\
&\quad \leq \left[C_{\mathrm{hc}} \left(\varsigma_{\mathrm{de}}^r + \|\boldsymbol{\tau}\|_W^r + \|\boldsymbol{\eta}\|_W^r\right)^{\frac{r-2}{r}}\right]^2 \left[(\|\boldsymbol{\tau}\|_W - \|\boldsymbol{\eta}\|_W)^2 + 2(\|\boldsymbol{\tau}\|_W\|\boldsymbol{\eta}\|_W - (\boldsymbol{\tau}, \boldsymbol{\eta})_W)\right] \\
&\quad = \left[C_{\mathrm{hc}} \left(\varsigma_{\mathrm{de}}^r + \|\boldsymbol{\tau}\|_W^r + \|\boldsymbol{\eta}\|_W^r\right)^{\frac{r-2}{r}} \|\boldsymbol{\tau} - \boldsymbol{\eta}\|_W\right]^2,
\end{aligned} \tag{79}$$

hence $\boldsymbol{\sigma}$ verifies (74a) for $r \geq 2$. Assume now $r < 2$. Using a triangle inequality followed by (77) and the left inequality in (36), it is inferred that

$$\begin{aligned}
\|\boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{\tau}) - \boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{\eta})\|_W &\leq \varsigma(\boldsymbol{x}, \|\boldsymbol{\tau}\|_W)\|\boldsymbol{\tau}\|_W + \varsigma(\boldsymbol{x}, \|\boldsymbol{\eta}\|_W)\|\boldsymbol{\eta}\|_W \\
&\leq C_{\mathrm{hc}} \left(\left(\varsigma_{\mathrm{de}}^r + \|\boldsymbol{\tau}\|_W^r\right)^{\frac{r-1}{r}} + \left(\varsigma_{\mathrm{de}}^r + \|\boldsymbol{\eta}\|_W^r\right)^{\frac{r-1}{r}}\right) \\
&\leq 2^{\frac{1}{r}} C_{\mathrm{hc}} (2\varsigma_{\mathrm{de}}^r + \|\boldsymbol{\tau}\|_W^r + \|\boldsymbol{\eta}\|_W^r)^{\frac{r-1}{r}} \\
&= 2^{\frac{1}{r}} C_{\mathrm{hc}} (2\varsigma_{\mathrm{de}}^r + \|\boldsymbol{\tau}\|_W^r + \|\boldsymbol{\eta}\|_W^r)^{\frac{r-2}{r}} (2\varsigma_{\mathrm{de}}^r + \|\boldsymbol{\tau}\|_W^r + \|\boldsymbol{\eta}\|_W^r)^{\frac{1}{r}}, \\
&\leq 2^{\frac{1}{r}} C_{\mathrm{hc}} (\varsigma_{\mathrm{de}}^r + \|\boldsymbol{\tau}\|_W^r + \|\boldsymbol{\eta}\|_W^r)^{\frac{r-2}{r}} (2\varsigma_{\mathrm{de}} + \|\boldsymbol{\tau}\|_W + \|\boldsymbol{\eta}\|_W),
\end{aligned}$$

where the last line follows from the fact that $t \mapsto t^{r-2}$ is decreasing and again (36). If $2\varsigma_{\mathrm{de}} + \|\boldsymbol{\tau}\|_W + \|\boldsymbol{\eta}\|_W \leq 2^{2-r}\|\boldsymbol{\tau} - \boldsymbol{\eta}\|_W$, from the previous bound we directly get the conclusion, i.e. (74a) with $\sigma_{\mathrm{hc}} = 2^{2-r+\frac{1}{r}} C_{\mathrm{hc}}$. Otherwise, using (36) and a triangle inequality yields

$$\begin{aligned}
(\varsigma_{\mathrm{de}}^r + \|\boldsymbol{\tau}\|_W^r)^{\frac{1}{r}} (\varsigma_{\mathrm{de}}^r + \|\boldsymbol{\eta}\|_W^r)^{\frac{1}{r}} &\geq 2^{-\frac{2}{r'}} (\varsigma_{\mathrm{de}} + \|\boldsymbol{\tau}\|_W)(\varsigma_{\mathrm{de}} + \|\boldsymbol{\eta}\|_W) \\
&= 2^{-2(\frac{1}{r'}+1)} \left[(2\varsigma_{\mathrm{de}} + \|\boldsymbol{\tau}\|_W + \|\boldsymbol{\eta}\|_W)^2 - (\|\boldsymbol{\tau}\|_W - \|\boldsymbol{\eta}\|_W)^2\right] \\
&\geq 2^{-2(\frac{1}{r'}+1)} \left[(2\varsigma_{\mathrm{de}} + \|\boldsymbol{\tau}\|_W + \|\boldsymbol{\eta}\|_W)^2 - \|\boldsymbol{\tau} - \boldsymbol{\eta}\|_W^2\right] \\
&\geq 2^{-2(\frac{1}{r'}+1)} (1 - 4^{r-2}) (2\varsigma_{\mathrm{de}} + \|\boldsymbol{\tau}\|_W + \|\boldsymbol{\eta}\|_W)^2 \\
&\geq 2^{\frac{2}{(r-2)r} - 2} (\varsigma_{\mathrm{de}}^r + \|\boldsymbol{\tau}\|_W^r + \|\boldsymbol{\eta}\|_W^r)^{\frac{2}{r}},
\end{aligned} \tag{80}$$

where we concluded with (36) together with the fact that $2^{-2(\frac{1}{r'}+1)} \left(1 - 4^{r-2}\right) \geq 2^{\frac{2}{(r-2)r} - 2}$. Finally, raising both sides of (80) to the power $r-2$, we get a relation analogous to (78). Hence, proceeding as in (79), we infer (74a). □

Corollary A.3 (Carreau–Yasuda)**.** *The strain rate-shear stress law of the (μ, δ, a, r)-Carreau–Yasuda fluid defined in Example 1.5 is an r-power-framed function.*

Proof. Let $\boldsymbol{x} \in \Omega$ and $g : (0, \infty) \to \mathbb{R}$ be such that, for all $\alpha \in (0, \infty)$,

$$g(\alpha) \coloneqq \frac{\partial}{\partial \alpha} \left[\alpha \mu(\boldsymbol{x}) \left(\delta^{a(\boldsymbol{x})} + \alpha^{a(\boldsymbol{x})} \right)^{\frac{r-2}{a(\boldsymbol{x})}} \right] = \mu(\boldsymbol{x}) \left(\delta^{a(\boldsymbol{x})} + \alpha^{a(\boldsymbol{x})} \right)^{\frac{r-2}{a(\boldsymbol{x})} - 1} \left(\delta^{a(\boldsymbol{x})} + (r-1)\alpha^{a(\boldsymbol{x})} \right).$$

We have for all $\alpha \in (0, \infty)$,

$$\mu_-(\tilde{r} - 1) \left(\delta^{a(\boldsymbol{x})} + \alpha^{a(\boldsymbol{x})} \right)^{\frac{r-2}{a(\boldsymbol{x})}} \leq g(\alpha) \leq \mu_+(r + 1 - \tilde{r}) \left(\delta^{a(\boldsymbol{x})} + \alpha^{a(\boldsymbol{x})} \right)^{\frac{r-2}{a(\boldsymbol{x})}},$$

and we conclude using (36) together with Theorem A.1. □

The work of M. Botti was funded by the European Commission through the H2020-MSCA-IF-EF project PDGeoFF (Grant no. 896616). This support is gratefully acknowledged.

References

[1] J. Aghili, S. Boyaval, and D. A. Di Pietro. Hybridization of mixed high-order methods on general meshes and application to the Stokes equations. *Comput. Meth. Appl. Math.*, 15(2):111–134, 2015.

[2] J. W. Barrett and W. B. Liu. Quasi-norm error bounds for the finite element approximation of a non-Newtonian flow. *Numer. Math.*, 68(4):437–456, 1994.

[3] H. Beirão da Veiga. On the global regularity of shear thinning flows in smooth domains. *J. Math. Anal. Appl.*, 349(2):335–360, 2009.

[4] L. Beirão da Veiga, C. Lovadina, and G. Vacca. Divergence free Virtual Elements for the Stokes problem on polygonal meshes. *ESAIM: Math. Model. Numer. Anal. (M2AN)*, 51(2):509–535, 2017.

[5] L. Beirão da Veiga, C. Lovadina, and G. Vacca. Virtual Elements for the Navier–Stokes Problem on Polygonal Meshes. *SIAM J. Numer. Anal.*, 56(3):1210–1242, 2018.

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

[7] L. C. Berselli and M. Růžička. Global regularity for systems with p-structure depending on the symmetric gradient. *Adv. Nonlinear Anal.*, 9(1):176–192, 2020.

[8] R. B. Bird, R. C. Armstrong, and O. Hassager. *Dynamics of Polymeric Liquids*, volume 1. John Wiley, New York, 2nd ed. edition, 1987.

[9] D. Boffi, F. Brezzi, and M. Fortin. *Mixed finite element methods and applications*, volume 44 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2013.

[10] M. Bogovskiĭ. Solution of the first boundary value problem for an equation of continuity of an incompressible medium. *Dokl. Akad. Nauk SSSR*, 248(5):1037–1040, 1979.

[11] L. Botti and D. A. Di Pietro. p-multilevel preconditioners for hho discretizations of the Stokes equations with static condensation. *Commun. Appl. Math. Comput.*, 2021. Accepted for publication.

[12] M. Botti, D. A. Di Pietro, and A. Guglielmana. A low-order nonconforming method for linear elasticity on general meshes. *Comput. Methods Appl. Mech. Engrg.*, 354:96–118, 2019.

[13] M. Botti, D. A. Di Pietro, and P. Sochala. A hybrid high-order method for nonlinear elasticity. *SIAM J. Numer. Anal.*, 55(6):2687–2717, 2017.

[14] M. Botti, D. A. Di Pietro, and P. Sochala. A nonconforming high-order method for nonlinear poroelasticity. In *Finite volumes for complex applications VIII—hyperbolic, elliptic and parabolic problems*, volume 200 of *Springer Proc. Math. Stat.*, pages 537–545. Springer, Cham, 2017.

[15] D. Castanon Quiroz and D. A. Di Pietro. A Hybrid High-Order method for the incompressible Navier–Stokes problem robust for large irrotational body forces. *Comput. Math. Appl.*, 2020. Published online.

[16] K. Deimling. *Nonlinear functional analysis*. Springer-Verlag, Berlin, 1985.

[17] D. A. Di Pietro and J. Droniou. A Hybrid High-Order method for Leray-Lions elliptic equations on general meshes. *Math. Comp.*, 86(307):2159–2191, 2017.

[18] D. A. Di Pietro and J. Droniou. $W^{s,p}$-approximation properties of elliptic projectors on polynomial spaces, with application to the error analysis of a hybrid high-order discretisation of Leray-Lions problems. *Math. Models Methods Appl. Sci.*, 27(5):879–908, 2017.

[19] D. A. Di Pietro and J. Droniou. *The Hybrid High-Order method for polytopal meshes*. Number 19 in Modeling, Simulation and Application. Springer International Publishing, 2020.
[20] D. A. Di Pietro, J. Droniou, and A. Harnist. Improved error estimates for Hybrid High-Order discretizations of Leray–Lions problems. *Calcolo*, 58(19), 2021.
[21] D. A. Di Pietro, J. Droniou, and G. Manzini. Discontinuous Skeletal Gradient Discretisation methods on polytopal meshes. *J. Comput. Phys.*, 355:397–425, 2018.
[22] D. A. Di Pietro, A. Ern, A. Linke, and F. Schieweck. A discontinuous skeletal method for the viscosity-dependent Stokes problem. *Comput. Methods Appl. Mech. Engrg.*, 306:175–195, 2016.
[23] D. A. Di Pietro and S. Krell. A Hybrid High-Order method for the steady incompressible Navier–Stokes problem. *J. Sci. Comput.*, 74(3):1677–1705, 2018.
[24] L. Diening and F. Ettwein. Fractional estimates for non-differentiable elliptic systems with general growth. *Forum Math.*, 20(3):523–556, 2008.
[25] L. Diening, C. Kreuzer, and E. Süli. Finite element approximation of steady flows of incompressible fluids with implicit power-law-like rheology. *SIAM J. Numer. Anal.*, 51(2):984–1015, 2013.
[26] J. Droniou, R. Eymard, T. Gallouët, C. Guichard, and R. Herbin. *The gradient discretisation method*, volume 82 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer, Cham, 2018.
[27] R. Duran, M. A. Muschietti, E. Russ, and P. Tchamitchian. Divergence operator and Poincaré inequalities on arbitrary bounded domains. *Complex Var. Elliptic Equ.*, 55(8-10):795–816, 2010.
[28] G. P. Galdi, R. Rannacher, A. M. Robertson, and S. Turek. *Hemodynamical flows*, volume 37 of *Oberwolfach Seminars*, chapter Mathematical Problems in Classical and Non-Newtonian Fluid Mechanics. Birkhuser, 2008.
[29] G. Geymonat and P. M. Suquet. Functional spaces for Norton-Hoff materials. *Math. Methods Appl. Sci.*, 8(2):206–222, 1986.
[30] R. Glowinski and J. Rappaz. Approximation of a nonlinear elliptic problem arising in a non-Newtonian fluid flow model in glaciology. *M2AN Math. Model. Numer. Anal.*, 37(1):175–186, 2003.
[31] A. Hirn. Approximation of the p-Stokes equations with equal-order finite elements. *J. Math. Fluid Mech.*, 15(1):65–88, 2013.
[32] T. Isaac, G. Stadler, and O. Ghattas. Solution of nonlinear Stokes equations discretized by high-order finite elements on nonconforming and anisotropic meshes, with application to ice sheet dynamics. *SIAM J. Sci. Comput.*, 37(6):B804–B833, 2015.
[33] S. Ko and E. Süli. Finite element approximation of steady flows of generalized newtonian fluids with concentration-dependent power-law index. *Math. Comp.*, 88(6):1061–1090, 2018.
[34] Ko, S., Pustejovská, P., and Süli, E. Finite element approximation of an incompressible chemically reacting non-newtonian fluid. *M2AN Math. Model. Numer. Anal.*, 52(2):509–541, 2018.
[35] C. Kreuzer and E. Süli. Adaptive finite element approximation of steady flows of incompressible fluids with implicit power-law-like rheology. *M2AN Math. Model. Numer. Anal.*, 50(5):1333–1369, 2016.
[36] O. A. Ladyzhenskaya. *The Mathematical Theory of Viscous Incompressible Flow*. Gordon Breach, New York, 2nd ed. edition, 1969.
[37] W. M. Lai, S. C. Kuei, and V. C. Mow. Rheological equations for synovial fluids. *J. Biomech. Eng.*, 100(4):169–186, 11 1978.
[38] J.-L. Lions and E. Magenes. *Non-homogeneous boundary value problems and applications. Vol. I*. Springer-Verlag, New York-Heidelberg, 1972. Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181.
[39] J. Málek and K. R. Rajagopal. Mathematical issues concerning the Navier–Stokes equations and some of their generalizations. In *Evolutionary Equations*, volume 2 of *Handbook of Differential Equations*, pages 371–459. Elsevier/North-Holland, Amsterdam, 2005.
[40] 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. *Math. Models Methods Appl. Sci.*, 5(6):789–812, 1995.
[41] M. Růžička and L. Diening. Non-newtonian fluids and function spaces. In *Nonlinear Analysis, Function Spaces and Applications*, pages 95–143. Institute of Mathematics of the Academy of Sciences of the Czech Republic, 2007.
[42] D. Sandri. Sur l'approximation numérique des écoulements quasi-newtoniens dont la viscosité suit la loi puissance ou la loi de carreau. *M2AN Math. Model. Numer. Anal.*, 27(2):131–155, 1993.
[43] D. Sandri. Numerical analysis of a four-field model for the approximation of a fluid obeying the power law or carreau's law. *Japan Journal of Industrial and Applied Mathematics*, 31(3):633–663, 2014.
[44] G. Schubert, D. L. Turcotte, and P. Olson. *Mantle Convection in the Earth and Planets*. Cambridge University Press, 2001.
[45] H. D. Ursell. Inequalities between sums of powers. *Proc. London Math. Soc. (3)*, 9:432–450, 1959.
[46] K. Yasuda, R.C. Armstrong, and R.E. Cohen. Shear flow properties of concentrated solutions of linear and star branched polystyrenes. *Rheologica Acta*, 20(2):163–178, 3 1981.