A Virtual Element method for non-Newtonian fluid flows *

Paola F. Antonietti[†1], Lourenço Beirão da Veiga[‡2], Michele Botti [§1], Giuseppe Vacca [¶3], and Marco Verani [‖1]

[1]MOX-Laboratory for Modeling and Scientific Computing, Dipartimento di Matematica, Politecnico di Milano, Piazza Leonardo da Vinci 32 - 20133 Milano, Italy
[3]Dipartimento di Matematica e Applicazioni, Università degli Studi di Milano-Bicocca, Via Roberto Cozzi 55 - 20125 Milano, Italy
[3]Dipartimento di Matematica, Università degli Studi di Bari, Via Edoardo Orabona 4 - 70125 Bari, Italy

May 25, 2024

Abstract

In this paper, we design and analyze a Virtual Element discretization for the steady motion of non-Newtonian, incompressible fluids. A specific stabilization, tailored to mimic the monotonicity and boundedness properties of the continuous operator, is introduced and theoretically investigated. The proposed method has several appealing features, including the exact enforcement of the divergence free condition and the possibility of making use of fully general polygonal meshes. A complete well-posedness and convergence analysis of the proposed method is presented under mild assumptions on the non-linear laws, encompassing common examples such as the Carreau–Yasuda model. Numerical experiments validating the theoretical bounds as well as demonstrating the practical capabilities of the proposed formulation are presented.

1 Introduction

In recent years, a novel approach known as Polytopal Finite Element methods has emerged. Polytopal Finite Element methods are Galerkin-type approximation schemes where the discretization space can support computational grids composed of arbitrarily polygonal or polyhedral (polytopal, for short) elements. Several methodologies have been developed in the last decade to extend the classical (i.e. based on tetrahedral, hexahedral and prismatic meshes) finite element paradigm to non-conventional elements, see, e.g., [3, 4, 12, 17, 23, 39, 44, 45, 47, 48, 49] and the references therein. Among them, the Virtual Element Method (VEM), first introduced in [23] for second-order elliptic problems and subsequently expanded to cover various other differential equations, has emerged as one of the most promising polytopal approaches.

Virtual Element Methods have been extensively developed in the last decade to approximate Newtonian fluid flows modeled by Stokes and Navier-Stokes equations. In [6] a novel stream formulation of the VEM for the solution of the Stokes problem, based on suitable stream function space characterizing the divergence free subspace of discrete velocities has been proposed and analyzed. Still for the Stokes problem, VEM have been further studied in [18, 40, 37, 15, 42, 61, 29, 16, 28], see also [52, 13] where arbitrary-order pressure-robust VEM methods have been studied. Virtual

*This is the Author Accepted Manuscript (AAM). The published version is available online since 5 June 2024 at https://doi.org/10.1016/j.cma.2024.117079 under subscription to Elsevier B.V.: Computer Methods in Applied Mechanics and Engineering 428 (2024) 117079.

[†]paola.antonietti@polimi.it
[‡]lourenco.beirao@unimib.it
[§]michele.botti@polimi.it
[¶]giuseppe.vacca@uniba.it
[‖]marco.verani@polimi.it

Element discretization of Navier-Stokes equations has been first studied in [20] and then further investigated in [53, 21, 1, 54], cf. also the recent works [2, 64] and [38] for quasi-Newtonian Stokes flows. VEM for the coupled Navier-Stokes and heat equations have been proposed and analyzed in [5, 25], whereas Least-squares type VEMs for the Stokes and Navier-Stokes problems have been recently addressed in [63, 70], respectively. We refer to the book [7] for a comprehensive overview of the recent developments in Virtual Element Methods.

On the other hand, complex fluids relevant to engineering and life-science applications often behave as non-Newtonian fluids, involving shear-rate dependent non-linear viscosity. Indeed, non-Newtonian fluids' rheological properties, which differ from the linear relationship between shear stress and shear rate observed in Newtonian fluids, make them applicable in a variety of applications, including for example (bio)polymer manufacturing processes, biomedical engineering (blood flow modeling, tissue engineering, and drug delivery), energy production systems through complex geological formations, and modeling of food processing to mention a few.

In this paper, we consider the steady motion of a non-Newtonian incompressible fluid, where the constitutive law obeys the so-called Carreau-Yasuda model, usually employed to describe pseudoplastic flow with asymptotic viscosities at zero and infinite shear rates. From the numerical viewpoint, the numerical discretization of non-Newtonian incompressible fluid flows has been extensively addressed, starting from the seminal work of [8] where a Finite Element approximation of a non-Newtonian flow model where the viscosity obeys the Carreau law or the power law is considered. Still in the framework of Finite Element methods, improved error bounds have been proved by [67, 10, 9]. More precisely, in the pioneering papers [10] and [9], (in some case optimal) velocity and pressure error bounds in suitable quasi-norms for a non-Newtonian flow model where the viscosity obeys the Carreau or the power law model are proved. The case of a non-degenerate power law is considered in [51]. Other notable recent contributions can be found, e.g., in [26, 57, 59, 58].

The approximation of non-Newtonian fluid flow problems requires numerical schemes able to represent the local features emerging from the nonlinear stress-strain relation and to handle complex unstructured and highly-adapted meshes. To this end, discretization methods that can possibly support general polyhedral have been recently explored. We mention in particular the discontinuous Galerkin and hybridizable discontinuous Galerkin methods of [65] and [55], respectively. Additionally, Hybrid High-Order methods have been applied to the simulation of non-Newtonian fluids governed by the Stokes equations in [32] and Navier–Stokes equations with nonlinear convection in [41]. Here, we consider the steady motion of a non-Newtonian incompressible fluid, where the constitutive law obeys the so-called Carreau-Yasuda model with the exponent r in (2) such that $r \in (1, 2]$ (shear thinning behavior). Our aim is to develop a Virtual Element approximation of such a problem and present a comprehensive theoretical analysis. More precisely, we consider as starting point the divergence-free Virtual spaces of [18, 20] and propose a new (divergence-free) Virtual Element formulation for the numerical discretization. Such approach exhibits two important advantages. The first one is the possibility of using general polygonal meshes, which allow, for instance, easier descriptions of complex domains and more efficient mesh adaptive strategies. Furthermore, the proposed approach yields an exactly divergence free discrete velocity solution and error estimates for the velocity field which do not depend on the pressure variable; this is a recognized advantage already for the standard linear Stokes problem, which becomes even more relevant in the current nonlinear setting (see for instance [60] and references therein). We develop a-priori error bounds for both the velocity and the pressure, focusing on the degenerate model with $\delta = 0$ in (2), which correspond to the power-law equation and which is recognized as the most complex case. The theoretical results could be extended to the case $\delta > 0$, cf. (2), combining the present results with, e.g., the approaches in [8, 10, 9]. Our analysis requires, in particular, the development of many technical results, such as the inf-sup stability of the discrete velocities-pressures coupling in non-hilbertian norms, and the proof of the continuity and coercivity properties of our novel stabilization form, specifically tailored for the current problem. To the best of our knowledge, this is the first work in the literature where $O(h^{kr/2})$ velocity error bounds are proved for a polytopal approximation method for the degenerate case and regular solution, with k denoting the polynomial order and $r \in (1, 2]$ the model exponent, cf. (2). Our theoretical estimates agree with the classical results presented, e.g., in [9] in the context of classical finite element discretizations. We demonstrate the practical capabilities of the proposed formulation in a wide set of numerical experiments. On the one hand, the numerical results confirm the theoretical estimates, and on the other hand, we show that, despite not being covered by our theoretical analysis, the proposed

formulation can be successfully employed also in the case $\delta \neq 0$ in (2).

The rest of the manuscript is organized as follows. The next section introduces the notation we are going to employ throughout the manuscript. The weak formulation of the model problem together with its well-posedness is discussed in Section 2. The proposed divergence-free Virtual Element discretization is described in Section 3 and the a-priori error analysis is carried out in Section 4. A wide set of numerical experiments validating the theoretical bounds as well as showing in practice the capabilities of the proposed formulation are presented in Section 5. Finally, Appendix A contains some technical results.

1.1 Notation

The vector spaces considered hereafter are over $\mathbb{R}$. We denote by $\mathbb{R}_+$ the set of non-negative real numbers. Given a vector space V with norm $\|\cdot\|_V$, the notation V' denotes its dual space and ${}_{V'}\langle\cdot,\cdot\rangle_V$ the duality between V and V'. The notation $\boldsymbol{v}\cdot\boldsymbol{w}$ and $\boldsymbol{v}\times\boldsymbol{w}$ designate the scalar and vector products of two vectors $\boldsymbol{v},\boldsymbol{w}\in\mathbb{R}^d$, and $|\boldsymbol{v}|$ denotes the Euclidean norm of $\boldsymbol{v}$ in $\mathbb{R}^d$. The inner product in $\mathbb{R}^{d\times d}$ is defined for $\boldsymbol{\tau},\boldsymbol{\eta}\in\mathbb{R}^{d\times d}$ by $\boldsymbol{\tau}:\boldsymbol{\eta}:=\sum_{i,j=1}^d \tau_{i,j}\eta_{i,j}$ and the induced norm is given by $|\boldsymbol{\tau}|:=\sqrt{\boldsymbol{\tau}:\boldsymbol{\tau}}$.

Let $\Omega\subset\mathbb{R}^d$ denote a bounded, connected, polyhedral open set with Lipschitz boundary $\partial\Omega$ and let $\boldsymbol{n}$ be the outward unit normal to $\partial\Omega$. To simplify the exposition, we restrict the presentation to the two-dimensional case, i.e. $d=2$, but the analysis path remains valid in the three-dimensional case $d=3$ as well, with obvious minor technical differences. We denote with $\boldsymbol{x}:=(x_1,x_2)$ the independent variable. We assume that the boundary is partitioned in two disjoint subsets $\partial\Omega:=\Gamma_D\cup\Gamma_N$, with $|\Gamma_D|>0$, such that a Dirichlet condition is given on Γ_D and a Neumann condition on Γ_N.

Throughout the article, spaces of functions, vector fields, and tensor fields, defined over any $X\subset\overline{\Omega}$ are denoted by italic capitals, boldface Roman capital, and special Roman capitals, respectively. The subscript s denotes a space of symmetric tensor fields. For example, $L^2(X)$, $\boldsymbol{L}^2(X)$, and $\mathbb{L}^2_s(X)$ denote the spaces of square-integrable functions, vector fields, and symmetric tensor fields, respectively. The notation $W^{m,r}(X)$, for $m\geq 0$ and $r\in[1,+\infty]$, with the convention that $W^{0,r}(X):=L^r(X)$ and $W^{m,2}(X):=H^m(X)$, designate the classical Sobolev spaces. The trace map is denoted by $\gamma:W^{1,r}(\Omega)\to W^{1-\frac{1}{r},r}(\partial\Omega)$. Finally, given $\Gamma\subset\partial\Omega$, we denote by $W^{1,r}_{0,\Gamma}(\Omega)$ the subspace of $W^{1,r}(\Omega)$ spanned by functions having zero-trace on Γ.

The symbol ∇ denotes the gradient for scalar functions, while $\boldsymbol{\nabla}$, $\boldsymbol{\epsilon}:=\frac{\boldsymbol{\nabla}+\boldsymbol{\nabla}^{\mathrm{T}}}{2}$, and $\nabla\cdot$ denote the gradient, the symmetric gradient operator, and the divergence operator, whereas $\boldsymbol{\nabla}\cdot$ denotes the vector-valued divergence operator for tensor fields.

2 Model problem

The incompressible flow of a non-Newtonian fluid occupying Ω and subjected to a volumetric force field $\boldsymbol{f}:\Omega\to\mathbb{R}^d$ and a normal stress $\boldsymbol{g}:\Gamma_N\to\mathbb{R}^d$ is described by the non-linear Stokes equations

$$\begin{aligned} -\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}(\cdot,\boldsymbol{\epsilon}(\boldsymbol{u}))+\boldsymbol{\nabla}p &= \boldsymbol{f} && \text{in } \Omega,\\ \nabla\cdot\boldsymbol{u} &= 0 && \text{in } \Omega,\\ \boldsymbol{\sigma}(\cdot,\boldsymbol{\epsilon}(\boldsymbol{u}))\boldsymbol{n}-p\boldsymbol{n} &= \boldsymbol{g} && \text{on } \Gamma_N,\\ \boldsymbol{\gamma}(\boldsymbol{u}) &= \boldsymbol{0} && \text{on } \Gamma_D, \end{aligned} \tag{1}$$

where $\boldsymbol{u}:\Omega\to\mathbb{R}^d$ and $p:\Omega\to\mathbb{R}$ denote the velocity field and the pressure field, respectively. The homogeneous Dirichlet boundary condition (1) can be generalized to non-homogeneous data by minor modifications.

In this work, we consider as a reference model for the non-linear shear stress-strain rate relation the Carreau–Yasuda model introduced in [72], i.e.

$$\boldsymbol{\sigma}(\boldsymbol{x},\boldsymbol{\epsilon}(\boldsymbol{v}))=\mu(\boldsymbol{x})(\delta^\alpha+|\boldsymbol{\epsilon}(\boldsymbol{v})|^\alpha)^{\frac{r-2}{\alpha}}\boldsymbol{\epsilon}(\boldsymbol{v}), \tag{2}$$

where $\mu:\Omega\to[\mu_-,\mu_+]$, with $0<\mu_-<\mu_+<+\infty$, $\alpha\in[1,\infty)$, and $\delta\geq 0$. The Carreau–Yasuda law is a generalization of the Carreau model which corresponds to the case $\alpha=2$. The case $\delta=0$ corresponds to the classical power-law model. Most real fluids that can be described by a constitutive

relation of type (2) exhibits shear thinning behavior corresponding to the case $r < 2$. For $r = 2$, problem (1) reduces to the standard Stokes system for Newtonian fluids. For the sake of conciseness, in what follows, we only consider the pseudoplastic case $r \leq 2$ that is the most common in practical applications and the one presenting more challenges in terms of numerical analysis. Nevertheless, the following arguments can be easily adapted to establish the results for the dilatant case $r > 2$.

According to [32, Appendix A], the stress-strain law (2) satisfy the following assumption:

Assumption 1. The shear stress-strain rate law $\boldsymbol{\sigma} : \Omega \times \mathbb{R}^{d\times d}_{\mathrm{s}} \to \mathbb{R}^{d\times d}_{\mathrm{s}}$ appearing in (1) is a Caratheodory function satisfying $\boldsymbol{\sigma}(\cdot, \mathbf{0}) = \mathbf{0}$ and for a fixed $r \in (1, 2]$ there exist real numbers $\delta \in [0, +\infty)$ and $\sigma_{\mathrm{c}}, \sigma_{\mathrm{m}} \in (0, +\infty)$ such that the following conditions hold:

$$|\boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{\tau}) - \boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{\eta})| \leq \sigma_{\mathrm{c}} \left(\delta^r + |\boldsymbol{\tau}|^r + |\boldsymbol{\eta}|^r\right)^{\frac{r-2}{r}} |\boldsymbol{\tau} - \boldsymbol{\eta}|, \qquad \text{(Hölder continuity)} \quad (3a)$$

$$(\boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{\tau}) - \boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{\eta})) : (\boldsymbol{\tau} - \boldsymbol{\eta}) \geq \sigma_{\mathrm{m}} \left(\delta^r + |\boldsymbol{\tau}|^r + |\boldsymbol{\eta}|^r\right)^{\frac{r-2}{r}} |\boldsymbol{\tau} - \boldsymbol{\eta}|^2, \qquad \text{(strong monotonicity)} \quad (3b)$$

for almost every $\boldsymbol{x} \in \Omega$ and all $\boldsymbol{\tau}, \boldsymbol{\eta} \in \mathbb{R}^{d\times d}_{\mathrm{s}}$.

We observe that the positive constants $\sigma_{\mathrm{c}}, \sigma_{\mathrm{m}}$ in (3) for the Carreau-Yasuda model (2) depends on μ, α and r and are such that

$$\mu_-(r-1)2^{\frac{(r-1)(r-2)}{r}} \leq \sigma_{\mathrm{m}} \leq \sigma_{\mathrm{c}} \leq \frac{\mu_+}{r-1} 2^{\frac{2r(2-r)+1}{r}}.$$

Finally, for further use, we recall the strong monotonicity bound

$$|\boldsymbol{x} - \boldsymbol{y}|^2 (\delta^r + |x|^r + |y|^r)^{\frac{r-2}{r}} \lesssim \left\{(\delta^r + |\boldsymbol{x}|^r)^{\frac{r-2}{r}} \boldsymbol{x} - (\delta^r + |\boldsymbol{y}|^r)^{\frac{r-2}{r}} \boldsymbol{y}\right\} \cdot (\boldsymbol{x} - \boldsymbol{y}) \quad (4)$$

for $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\delta \geq 0$. Here and in the following, to avoid the proliferation of constants, we adopt the notation $\mathsf{a} \lesssim \mathsf{b}$ to denote the inequality $\mathsf{a} \leq C\mathsf{b}$, for a positive constant C that might depend on $\sigma_{\mathrm{c}}, \sigma_{\mathrm{m}}$ (or related parameters) in Assumption 1 and on r, but is independent of the discretization parameter h. The obvious extensions $\mathsf{a} \gtrsim \mathsf{b}$ and $\mathsf{a} \simeq \mathsf{b}$ hold.

2.1 Weak formulation

Before deriving the variational formulation of problem (1) and discussing its well-posedness, we introduce additional notation and recall some basic results concerning Sobolev spaces. First of all, the expressions of the conjugate index and the Sobolev index are given by

$$r' := \begin{cases} \frac{r}{r-1} & \text{if } r > 1, \\ +\infty & \text{if } r = 1, \end{cases} \qquad r^* := \begin{cases} \frac{dr}{d-r} & \text{if } r < d, \\ +\infty & \text{if } r \geq d, \end{cases}$$

respectively. From the classical Sobolev embedding theorems [36, Section 9.3] it is inferred that $W^{1,r}(\Omega) \subset L^q(\Omega)$ for all $q \in [1, r^*]$ (excluding $\{+\infty\}$ if $r = d$) and the embedding is compact for $q < r^*$. We also recall Korn's first inequality (see, e.g., [43, Theorem 1.2] and [56, 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 $\boldsymbol{v} \in \boldsymbol{W}^{1,r}_{0,\Gamma_D}(\Omega)$,

$$\|\boldsymbol{v}\|_{\boldsymbol{W}^{1,r}(\Omega)} \leq C_{\mathrm{K}} \|\boldsymbol{\epsilon}(\boldsymbol{v})\|_{\mathbb{L}^r(\Omega)}. \quad (5)$$

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. Let $r \in (1, 2]$ be the Sobolev exponent dictated by the non-linear stress-strain law characterizing problem (1) and satisfying Assumption 1. We define the following velocity and pressure spaces incorporating the homogeneous boundary condition on Γ_D and the zero-average constraint in the case $\Gamma_D := \partial\Omega$, respectively:

$$\boldsymbol{U} := \boldsymbol{W}^{1,r}_{0,\Gamma_D}(\Omega) \qquad P = \begin{cases} L^{r'}(\Omega) & \text{if } |\Gamma_D| < |\partial\Omega| \\ L^{r'}_0(\Omega) := \left\{q \in L^{r'}(\Omega) \ : \ \int_\Omega q = 0\right\} & \text{if } \Gamma_D = \partial\Omega. \end{cases}$$

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

$$\begin{aligned} a(\boldsymbol{u}, \boldsymbol{v}) + b(\boldsymbol{v}, p) &= \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v} + \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{\gamma}(\boldsymbol{v}) && \forall \boldsymbol{v} \in \boldsymbol{U}, \\ -b(\boldsymbol{u}, q) &= 0 && \forall q \in P, \end{aligned} \quad (6)$$

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

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

Let us introduce the kernel of the bilinear form $b(\cdot, \cdot)$ that corresponds to the functions in $\boldsymbol{U}$ with vanishing divergence, i.e.

$$\boldsymbol{Z} := \{\boldsymbol{v} \in \boldsymbol{U} \quad \text{s.t.} \quad \nabla \cdot \boldsymbol{v} = 0\}.$$

Then, Problem (6) can be formulated in the equivalent kernel form: Find $\boldsymbol{u} \in \boldsymbol{Z}$ such that

$$a(\boldsymbol{u}, \boldsymbol{v}) = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v} + \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{\gamma}(\boldsymbol{v}) \qquad \forall \boldsymbol{v} \in \boldsymbol{Z}. \tag{8}$$

2.2 Well-posedness

In this section, after reporting the properties of the viscous function $a(\cdot, \cdot)$ and the coupling bilinear form $b(\cdot, \cdot)$ defined in (7), we prove the well-posedness of problem (6).

Lemma 1 (Continuity and strong monotonicity of a)**.** *For all $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \boldsymbol{U}$, setting $\boldsymbol{e} := \boldsymbol{u} - \boldsymbol{v}$, it holds*

$$|a(\boldsymbol{u}, \boldsymbol{v}) - a(\boldsymbol{w}, \boldsymbol{v})| \le \sigma_{\mathrm{c}} \|\boldsymbol{e}\|^{r-1}_{\boldsymbol{W}^{1,r}(\Omega)} \|\boldsymbol{v}\|_{\boldsymbol{W}^{1,r}(\Omega)} \tag{9a}$$

$$a(\boldsymbol{u}, \boldsymbol{e}) - a(\boldsymbol{w}, \boldsymbol{e}) \gtrsim \sigma_{\mathrm{m}} \left(\delta^r + \|\boldsymbol{u}\|^r_{\boldsymbol{W}^{1,r}(\Omega)} + \|\boldsymbol{w}\|^r_{\boldsymbol{W}^{1,r}(\Omega)}\right)^{\frac{r-2}{r}} \|\boldsymbol{e}\|^2_{\boldsymbol{W}^{1,r}(\Omega)}. \tag{9b}$$

Proof. Let $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \boldsymbol{U}$ and set $\boldsymbol{e} := \boldsymbol{u} - \boldsymbol{w}$.
(i) Hölder continuity. First, we make a preliminary observation. For every $\boldsymbol{\tau}, \boldsymbol{\eta} \in \mathbb{R}^{d \times d}_{\mathrm{s}}$ the triangle inequality implies that

$$2^{1-r} |\boldsymbol{\tau} - \boldsymbol{\eta}|^r \le |\boldsymbol{\tau}|^r + |\boldsymbol{\eta}|^r.$$

Therefore, since $\delta \ge 0$ and $r - 2 < 0$ we have

$$(\delta^r + |\boldsymbol{\tau}|^r + |\boldsymbol{\eta}|^r)^{\frac{r-2}{r}} \le \left(2^{1-r} |\boldsymbol{\tau} - \boldsymbol{\eta}|^r\right)^{\frac{r-2}{r}} \le 2^{\frac{(1-r)(r-2)}{r}} |\boldsymbol{\tau} - \boldsymbol{\eta}|^{r-2} \lesssim |\boldsymbol{\tau} - \boldsymbol{\eta}|^{r-2}. \tag{10}$$

Recalling the Hölder continuity property (3a) and using (10) followed by the Hölder inequality with exponents (r', r), we have (here $\boldsymbol{e} := \boldsymbol{u} - \boldsymbol{w}$)

$$\begin{aligned}
|a(\boldsymbol{u}, \boldsymbol{v}) - a(\boldsymbol{w}, \boldsymbol{v})| &\le \int_\Omega |[\boldsymbol{\sigma}(\cdot, \boldsymbol{\epsilon}(\boldsymbol{u})) - \boldsymbol{\sigma}(\cdot, \boldsymbol{\epsilon}(\boldsymbol{w}))] : \boldsymbol{\epsilon}(\boldsymbol{v})| \\
&\le \sigma_{\mathrm{c}} \int_\Omega (\delta^r + |\boldsymbol{\epsilon}(\boldsymbol{u})|^r + |\boldsymbol{\epsilon}(\boldsymbol{w})|^r)^{\frac{r-2}{r}} |\boldsymbol{\epsilon}(\boldsymbol{e})| |\boldsymbol{\epsilon}(\boldsymbol{v})| \\
&\lesssim \sigma_{\mathrm{c}} \int_\Omega |\boldsymbol{\epsilon}(\boldsymbol{e})|^{r-1} |\boldsymbol{\epsilon}(\boldsymbol{v})| \\
&\lesssim \sigma_{\mathrm{c}} \|\boldsymbol{e}\|^{r-1}_{\boldsymbol{W}^{1,r}(\Omega)} \|\boldsymbol{v}\|_{\boldsymbol{W}^{1,r}(\Omega)}.
\end{aligned} \tag{11}$$

(ii) Strong monotonicity. Using the Korn's inequality (5) together with the monotonicity property of $\boldsymbol{\sigma}$ in (3b) and the Hölder inequality with exponents $(\frac{2}{2-r}, \frac{2}{r})$, we obtain

$$\begin{aligned}
\sigma_{\mathrm{m}} \|\boldsymbol{e}\|^2_{\boldsymbol{W}^{1,r}(\Omega)} &\le C_{\mathrm{K}}^2 \sigma_{\mathrm{m}} \left(\int_\Omega |\boldsymbol{\epsilon}(\boldsymbol{u} - \boldsymbol{w})|^r\right)^{\frac{2}{r}} \lesssim \sigma_{\mathrm{m}} \left(\int_\Omega \left(|\boldsymbol{\epsilon}(\boldsymbol{u}) - \boldsymbol{\epsilon}(\boldsymbol{w})|^2\right)^{\frac{r}{2}}\right)^{\frac{2}{r}} \\
&\lesssim \left(\int_\Omega (\delta^r + |\boldsymbol{\epsilon}(\boldsymbol{u})|^r + |\boldsymbol{\epsilon}(\boldsymbol{w})|^r)^{\frac{2-r}{2}} \left((\boldsymbol{\sigma}(\cdot, \boldsymbol{\epsilon}(\boldsymbol{u})) - \boldsymbol{\sigma}(\cdot, \boldsymbol{\epsilon}(\boldsymbol{w}))) : \boldsymbol{\epsilon}(\boldsymbol{u} - \boldsymbol{w})\right)^{\frac{r}{2}}\right)^{\frac{2}{r}} \\
&\lesssim \left(\int_\Omega \left(\delta^r + |\boldsymbol{\epsilon}(\boldsymbol{u})|^r + |\boldsymbol{\epsilon}(\boldsymbol{w})|^r\right)\right)^{\frac{2-r}{r}} \left(\int_\Omega (\boldsymbol{\sigma}(\cdot, \boldsymbol{\epsilon}(\boldsymbol{u})) - \boldsymbol{\sigma}(\cdot, \boldsymbol{\epsilon}(\boldsymbol{w}))) : \boldsymbol{\epsilon}(\boldsymbol{u} - \boldsymbol{w})\right) \\
&\lesssim \left(|\Omega| \delta^r + \|\boldsymbol{u}\|^r_{\boldsymbol{W}^{1,r}(\Omega)} + \|\boldsymbol{w}\|^r_{\boldsymbol{W}^{1,r}(\Omega)}\right)^{\frac{2-r}{r}} (a(\boldsymbol{u}, \boldsymbol{e}) - a(\boldsymbol{w}, \boldsymbol{e})).
\end{aligned}$$

Rearranging the previous inequality, yields the conclusion. □

The following result is needed to infer the existence of a unique pressure $p \in P$ solving problem (6) from the well-posedness of problem (8). For its proof we refer to [31, Theorem 1].

Lemma 2 (Inf-sup condition)**.** *For any $r \in (1, \infty)$ there exists a positive constant $\beta(r)$ such that the bilinear form $b(\cdot, \cdot)$ defined in (7) satisfies*

$$\inf_{q \in P} \sup_{\boldsymbol{w} \in \boldsymbol{U} \setminus \{\boldsymbol{0}\}} \frac{b(\boldsymbol{w}, q)}{\|q\|_{L^{r'}(\Omega)} \|\boldsymbol{w}\|_{\boldsymbol{W}^{1,r}(\Omega)}} \geq \beta(r) > 0. \tag{12}$$

We are now ready to establish the well-posedness of problem (6).

Proposition 3 (Well-posedness)**.** *For any $r \in (1, 2)$, there exists a unique solution $(\boldsymbol{u}, p) \in \boldsymbol{U} \times P$ to problem problem (6) satisfying the a-priori estimates*

$$\|\boldsymbol{u}\|_{\boldsymbol{W}^{1,r}(\Omega)} \lesssim \delta^{2-r} \mathcal{N}(\boldsymbol{f}, \boldsymbol{g}) + \mathcal{N}(\boldsymbol{f}, \boldsymbol{g})^{\frac{1}{r-1}}, \tag{13}$$

$$\|p\|_{L^{r'}(\Omega)} \lesssim \mathcal{N}(\boldsymbol{f}, \boldsymbol{g}) + \delta^{(2-r)(r-1)} \mathcal{N}(\boldsymbol{f}, \boldsymbol{g})^{r-1}, \tag{14}$$

where

$$\mathcal{N}(\boldsymbol{f}, \boldsymbol{g}) := 2^{\frac{2-r}{r}} \sigma_{\mathrm{m}}^{-1} \left(\|\boldsymbol{f}\|_{\boldsymbol{L}^{r'}(\Omega)} + \|\boldsymbol{g}\|_{\boldsymbol{L}^{r'}(\Gamma_N)} \right). \tag{15}$$

Proof. We focus on the proof of the a-priori bounds (13) and (14). For the uniqueness and existence we refer to [11, 27]. Using (9b) with $\boldsymbol{w} = \boldsymbol{0}$ and taking $\boldsymbol{v} = \boldsymbol{u}$ in (8), it is inferred that

$$\sigma_{\mathrm{m}} \left(\delta^r + \|\boldsymbol{u}\|^r_{\boldsymbol{W}^{1,r}(\Omega)} \right)^{\frac{r-2}{r}} \|\boldsymbol{u}\|^2_{\boldsymbol{W}^{1,r}(\Omega)} \lesssim a(\boldsymbol{u}, \boldsymbol{u}) = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{u} + \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{\gamma}(\boldsymbol{u}).$$

Thus owing to the Hölder inequality and the continuity of the trace map, we obtain

$$\sigma_{\mathrm{m}} \left(\delta^r + \|\boldsymbol{u}\|^r_{\boldsymbol{W}^{1,r}(\Omega)} \right)^{\frac{r-2}{r}} \|\boldsymbol{u}\|^{\cancel{2}}_{\boldsymbol{W}^{1,r}(\Omega)} \lesssim \left(\|\boldsymbol{f}\|_{\boldsymbol{L}^{r'}(\Omega)} + \|\boldsymbol{g}\|_{\boldsymbol{L}^{r'}(\Gamma_N)} \right) \cancel{\|\boldsymbol{u}\|_{\boldsymbol{W}^{1,r}(\Omega)}}. \tag{16}$$

If $\|\boldsymbol{u}\|_{\boldsymbol{W}^{1,r}(\Omega)} \geq \delta$, from the previous bound it follows that

$$\|\boldsymbol{u}\|^{r-1}_{\boldsymbol{W}^{1,r}(\Omega)} \lesssim 2^{\frac{2-r}{r}} \sigma_{\mathrm{m}}^{-1} \left(\|\boldsymbol{f}\|_{L^{r'}(\Omega, \mathbb{R}^d)} + \|\boldsymbol{g}\|_{L^{r'}(\Gamma_N, \mathbb{R}^d)} \right) := \mathcal{N}(\boldsymbol{f}, \boldsymbol{g}). \tag{17}$$

Otherwise, owing to $\|\boldsymbol{u}\|^r_{\boldsymbol{W}^{1,r}(\Omega)} \geq \delta^r$ and (16) we obtain

$$\|\boldsymbol{u}\|_{\boldsymbol{W}^{1,r}(\Omega)} \lesssim \delta^{2-r} \mathcal{N}(\boldsymbol{f}, \boldsymbol{g}). \tag{18}$$

Therefore, from (17) and (18) we infer that

$$\|\boldsymbol{u}\|_{\boldsymbol{W}^{1,r}(\Omega)} \lesssim \delta^{2-r} \mathcal{N}(\boldsymbol{f}, \boldsymbol{g}) + \mathcal{N}(\boldsymbol{f}, \boldsymbol{g})^{\frac{1}{r-1}}.$$

We now move to the estimate of the pressure. Owing to the inf-sup condition (12) and equation (6), it holds

$$\beta(r) \|p\|_{L^{r'}(\Omega)} \leq \sup_{\boldsymbol{v} \in \boldsymbol{U} \setminus \{\boldsymbol{0}\}} \frac{b(\boldsymbol{v}, p)}{\|\boldsymbol{v}\|_{\boldsymbol{W}^{1,r}(\Omega)}} = \sup_{\boldsymbol{v} \in \boldsymbol{U} \setminus \{\boldsymbol{0}\}} \frac{\int_\Omega \boldsymbol{f} \cdot \boldsymbol{v} + \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{\gamma}(\boldsymbol{v}) - a(\boldsymbol{u}, \boldsymbol{v})}{\|\boldsymbol{v}\|_{\boldsymbol{W}^{1,r}(\Omega)}}.$$

Applying the Hölder inequality, the continuity of a in (9a) with $\boldsymbol{w} = \boldsymbol{0}$, and the a-priori estimate of the velocity (13), we obtain

$$\begin{aligned} \|p\|_{L^{r'}(\Omega)} &\lesssim \|\boldsymbol{f}\|_{\boldsymbol{L}^{r'}(\Omega)} + \|\boldsymbol{g}\|_{\boldsymbol{L}^{r'}(\Gamma_N)} + \sigma_{\mathrm{c}} \|\boldsymbol{u}\|^{r-1}_{\boldsymbol{W}^{1,r}(\Omega)} \\ &\lesssim \sigma_{\mathrm{c}} \mathcal{N}(\boldsymbol{f}, \boldsymbol{g}) + \sigma_{\mathrm{c}} \delta^{(2-r)(r-1)} \mathcal{N}(\boldsymbol{f}, \boldsymbol{g})^{r-1}. \end{aligned}$$

□

3 Virtual Elements discretization

In the present section, we present the divergence-free Virtual Elements for non-Newtonian fluid flows. In Subsection 3.1, we introduce some basic tools and notations useful in the construction and the theoretical analysis of Virtual Element Methods. In Subsections 3.2, we outline an overview of the inf-sup stable divergence–free velocities-pressures pair of spaces. In Subsection 3.3, we define the discrete computable form, in particular we design a VEM stabilizing form that is suited in the proposed non-linear setting. Finally, in Subsection 3.4, we introduce the virtual elements discretization of Problems (6) and (8) and we establish the well-posedness of the discrete problem.

3.1 Preliminaries

Let $\{\Omega_h\}_h$ be a sequence of decompositions of the domain $\Omega \subset \mathbb{R}^2$ into general polytopal elements E where $h := \sup_{E\in\Omega_h} h_E$. We suppose that $\{\Omega_h\}_h$ fulfils the following assumption.

Assumption 2. (Mesh assumptions). There exists a positive constant ρ such that for any $E \in \{\Omega_h\}_h$

- E is star-shaped with respect to a ball B_E of radius $\geq \rho\, h_E$;
- any edge e of E has length $\geq \rho\, h_E$.

We remark that the hypotheses above, though not too restrictive in many practical cases, could possibly be further relaxed, combining the present analysis with the studies in [24, 33, 35, 22]. Using standard VEM notations, for $n \in \mathbb{N}$ and $m \in \mathbb{R}_+$ and $p = 1, \ldots, +\infty$ let us introduce the spaces:

- $\mathbb{P}_n(\omega)$: the set of polynomials on $\omega \subset \Omega$ of degree $\leq n$ (with $\mathbb{P}_{-1}(\omega) = \{0\}$),
- $\mathbb{P}_n(\Omega_h) := \{q \in L^2(\Omega) \quad \text{s.t} \quad q|_E \in \mathbb{P}_n(E) \quad \text{for all } E \in \Omega_h\}$,
- $W^{m,p}(\Omega_h) := \{v \in L^p(\Omega) \quad \text{s.t} \quad v|_E \in W^{m,p}(E) \quad \text{for all } E \in \Omega_h\}$,

equipped with the broken norm and seminorm

$$\|v\|^p_{W^{m,p}(\Omega_h)} := \sum_{E\in\Omega_h} \|v\|^p_{W^{m,p}(E)}\,, \qquad |v|^p_{W^{m,p}(\Omega_h)} := \sum_{E\in\Omega_h} |v|^p_{W^{m,p}(E)}\,, \qquad \text{if } 1 \leq p < \infty. \tag{19}$$

Let $E \in \Omega_h$, we denote with h_E the diameter, with $|E|$ the area, with $\boldsymbol{x}_E := (x_{E,1}, x_{E,2})$ the centroid. A natural basis associated with the space $\mathbb{P}_n(E)$ is the set of normalized monomials

$$\mathbb{M}_n(E) := \{m_{\boldsymbol{\alpha}}, \ \text{ with } \ |\boldsymbol{\alpha}| \leq n\}$$

where, for any multi-index $\boldsymbol{\alpha} := (\alpha_1, \alpha_2) \in \mathbb{N}^2$

$$m_{\boldsymbol{\alpha}} := \prod_{i=1}^{2} \left(\frac{x_i - x_{E,i}}{h_E}\right)^{\alpha_i} \qquad \text{and} \qquad |\boldsymbol{\alpha}| := \sum_{i=1}^{2} \alpha_i\,.$$

For any e edge of Ω_h, the normalized monomial set $\mathbb{M}_n(e)$ is defined analogously as the span of all one-dimensional normalized monomials of degree up to n. Moreover for any $m \leq n$ we denote with

$$\widehat{\mathbb{P}}_{n\setminus m}(E) := \text{span}\,\{m_{\boldsymbol{\alpha}}, \ \text{ with } \ m+1 \leq |\boldsymbol{\alpha}| \leq n\}\ .$$

For any E, let us define the following polynomial projections:

- the L^2-projection $\Pi^{0,E}_n : L^2(E) \to \mathbb{P}_n(E)$, given by

$$\int_E q_n(v - \Pi^{0,E}_n v)\,\mathrm{d}E := 0 \qquad \text{for all } v \in L^2(E) \text{ and } q_n \in \mathbb{P}_n(E), \tag{20}$$

  with obvious extension for vector functions $\Pi^{0,E}_n : \boldsymbol{L}^2(E) \to [\mathbb{P}_n(E)]^d$ and tensor functions $\boldsymbol{\Pi}^{0,E}_n : \mathbb{L}^2(E) \to [\mathbb{P}_n(E)]^{d\times d}$;
- the H^1-seminorm projection $\Pi^{\nabla,E}_n : H^1(E) \to \mathbb{P}_n(E)$, defined by

$$\begin{cases} \displaystyle\int_E \nabla q_n \cdot \nabla(v - \Pi^{\nabla,E}_n v)\,\mathrm{d}E = 0 \quad \text{for all } v \in H^1(E) \text{ and } q_n \in \mathbb{P}_n(E), \\ \displaystyle\int_{\partial E} (v - \Pi^{\nabla,E}_n v)\,\mathrm{d}s := 0\,, \end{cases}$$

  with extension for vector fields $\Pi^{\nabla,E}_n : \boldsymbol{H}^1(E) \to [\mathbb{P}_n(E)]^d$.

Further we define the operator $\Pi^0_n \colon L^2(\Omega) \to \mathbb{P}_n(\Omega_h)$ such that $\Pi^0_n|_E := \Pi^{0,E}_n$ for any $E \in \Omega_h$. We finally recall the following well know useful results:

- Trace inequality with scaling [34]: For any $E \in \Omega_h$ and for any function $v \in W^{1,r}(E)$ it holds
$$\|v\|^r_{L^r(\partial E)} \lesssim h_E^{-1}\|v\|^r_{L^r(E)} + h_E^{r-1}\|\nabla v\|^r_{\boldsymbol{L}^r(E)}\,. \tag{21}$$
- Polynomial inverse estimate [34, Theorem 4.5.11]: Let $1 \le q, p \le \infty$, then for any $E \in \Omega_h$
$$\|p_n\|_{L^q(E)} \lesssim h_E^{2/q-2/p}\|p_n\|_{L^p(E)} \quad \text{for any } p_n \in \mathbb{P}_n(E). \tag{22}$$

3.2 Virtual Elements velocity and pressure spaces

Let $k \ge 2$ be the polynomial order of the method. We consider on each polygonal element $E \in \Omega_h$ the "enhanced" virtual space [19, 69]:

$$\begin{aligned}\boldsymbol{U}_h(E) := \Big\{ \boldsymbol{v}_h \in [C^0(\overline{E})]^2 \ \text{s.t.}\ &(i)\ \boldsymbol{\Delta}\boldsymbol{v}_h + \nabla s \in \boldsymbol{x}^\perp \mathbb{P}_{k-1}(E), \quad \text{for some } s \in L^2_0(E),\\ &(ii)\ \nabla\cdot \boldsymbol{v}_h \in \mathbb{P}_{k-1}(E)\,,\\ &(iii)\ \boldsymbol{v}_{h|e} \in [\mathbb{P}_k(e)]^2 \ \ \forall e \in \partial E,\\ &(iv)\ (\boldsymbol{v}_h - \Pi^{\nabla,E}_k \boldsymbol{v}_h, \boldsymbol{x}^\perp \widehat{p}_{k-1})_E = 0 \ \ \forall \widehat{p}_{k-1} \in \widehat{\mathbb{P}}_{k-1\setminus k-3}(E) \Big\},\end{aligned} \tag{23}$$

where $\boldsymbol{x}^\perp := (x_2, -x_1)$. Next, we summarize the main properties of the space $\boldsymbol{U}_h(E)$ (we refer to [20] for a deeper analysis).

(P1) Polynomial inclusion: $[\mathbb{P}_k(E)]^2 \subseteq \boldsymbol{U}_h(E)$;

(P2) Degrees of freedom: the following linear operators $\mathbf{D}_{\boldsymbol{U}}$ constitute a set of DoFs for $\boldsymbol{U}_h(E)$:

$\mathbf{D}_{\boldsymbol{U}}$1 the values of $\boldsymbol{v}_h$ at the vertexes of the polygon E,

$\mathbf{D}_{\boldsymbol{U}}$2 the values of $\boldsymbol{v}_h$ at $k-1$ distinct points of every edge $e \in \partial E$,

$\mathbf{D}_{\boldsymbol{U}}$3 the moments of $\boldsymbol{v}_h$
$$\frac{1}{|E|}\int_E \boldsymbol{v}_h \cdot \boldsymbol{m}^\perp m_{\boldsymbol{\alpha}}\,\mathrm{d}E \qquad \text{for any } m_{\boldsymbol{\alpha}} \in \mathbb{M}_{k-3}(E),$$
where $\boldsymbol{m}^\perp := \frac{1}{h_E}(x_2 - x_{2,E}, -x_1 + x_{1,E})$,

$\mathbf{D}_{\boldsymbol{U}}$4 the moments of $\nabla\cdot\boldsymbol{v}_h$
$$\frac{h_E}{|E|}\int_E (\nabla\cdot\boldsymbol{v}_h)\, m_{\boldsymbol{\alpha}}\,\mathrm{d}E \qquad \text{for any } m_{\boldsymbol{\alpha}} \in \mathbb{M}_{k-1}(E) \text{ with } |\boldsymbol{\alpha}| > 0;$$

(P3) Polynomial projections: the DoFs $\mathbf{D}_{\boldsymbol{U}}$ allow us to compute the following linear operators:
$$\Pi^{0,E}_k \colon \boldsymbol{U}_h(E) \to [\mathbb{P}_k(E)]^2, \qquad \boldsymbol{\Pi}^{0,E}_{k-1} \colon \boldsymbol{\nabla}\boldsymbol{U}_h(E) \to [\mathbb{P}_{k-1}(E)]^{2\times 2}\,.$$

The global velocity space $\boldsymbol{U}_h$ is defined by gluing the local spaces with the obvious associated sets of global DoFs:
$$\boldsymbol{U}_h := \{\boldsymbol{v}_h \in \boldsymbol{U} \quad \text{s.t.} \quad \boldsymbol{v}_h|_E \in \boldsymbol{U}_h(E) \quad \text{for all } E \in \Omega_h\}\,. \tag{24}$$
The discrete pressure space P_h is given by the piecewise polynomial functions of degree $k-1$, i.e.
$$P_h := \{q_h \in P \quad \text{s.t.} \quad q_h|E \in \mathbb{P}_{k-1}(E) \quad \text{for all } E \in \Omega_h\}\,. \tag{25}$$
The couple of spaces $(\boldsymbol{U}_h, P_h)$ is well known to be inf-sup stable in the classical Hilbertian setting. The inf-sup stability for $r \neq 2$ is proven below in Section 3.3.2. Let us introduce the discrete kernel
$$\boldsymbol{Z}_h := \{\boldsymbol{v}_h \in \boldsymbol{U}_h \quad \text{s.t.} \quad b(\boldsymbol{v}_h, q_h) = 0 \quad \text{for all } q_h \in P_h\} \tag{26}$$
then recalling (ii) in (23) and (25), the following kernel inclusion holds
$$\boldsymbol{Z}_h \subseteq \boldsymbol{Z}\,,$$
i.e. the functions in the discrete kernel are exactly divergence-free.

Uniquely for the purpose of defining our interpolant in the space $\boldsymbol{U}_h$, we consider also the alternative set of edge degrees of freedom (which can substitute $\mathbf{D}_{\boldsymbol{U}}$**2**)

$\mathbf{D_U 2'}$ the moments of $\boldsymbol{v}_h$

$$\frac{1}{|e|}\int_e \boldsymbol{v}_h \cdot \boldsymbol{t}_e m_\alpha \,\mathrm{d}s\,, \qquad \frac{1}{|e|}\int_e \boldsymbol{v}_h \cdot \boldsymbol{n}_e m_\alpha \,\mathrm{d}s \qquad \text{for any } m_\alpha \in \mathbb{M}_{k-2}(e)\,, \tag{27}$$

where $\boldsymbol{t}_e$ and $\boldsymbol{n}_e$ denote the tangent and the normal vectors to the edge e respectively.

Given any $\boldsymbol{v} \in \boldsymbol{W}^{s,p}(E)$, with $p \in (1,\infty)$ and $s \in \mathbb{R}_+$, $s > 2/p$, we define its approximant $\boldsymbol{v}_I \in \boldsymbol{U}_h$ as the unique function in $\boldsymbol{U}_h$ that interpolates $\boldsymbol{v}$ with respect to the DoF set $\mathbf{D_U 1}$, $\mathbf{D_U 2'}$, $\mathbf{D_U 3}$, $\mathbf{D_U 4}$. It is easy to check that, whenever $\mathrm{div}\boldsymbol{v} = 0$, then $\boldsymbol{v}_I \in \boldsymbol{Z}_h$. Furthermore, the following approximation property is a trivial generalization of the results in [66].

Lemma 4. *Let $E \in \Omega_h$, $n \in \mathbb{N}$, $p \in [1,\infty]$, $s \in \mathbb{R}_+$ and $\boldsymbol{v} \in \boldsymbol{W}^{s,p}(E)$. For $s > 2/p$, let $\boldsymbol{v}_I \in \boldsymbol{U}_h$ be the interpolant of $\boldsymbol{v}$ defined above. It holds*

$$\begin{aligned} &|\boldsymbol{v} - \Pi_n^{0,E}\boldsymbol{v}|_{\boldsymbol{W}^{m,p}(E)} \lesssim h_E^{s-m}|\boldsymbol{v}|_{\boldsymbol{W}^{s,p}(E)} && \text{for } 0 \le m \le s \le n+1,\\ &|\boldsymbol{v} - \boldsymbol{v}_I|_{\boldsymbol{W}^{m,p}(E)} \lesssim h_E^{s-m}|\boldsymbol{v}|_{\boldsymbol{W}^{s,p}(E)} && \text{for } 2/p < s \le k+1,\ m \in \{0,1\}. \end{aligned}$$

The first bound above extends identically to the scalar and tensor-valued case.

3.3 Virtual Element form: stabilization and inf-sup condition

The next step in the construction of the method is the definition of a discrete version of the form $a(\cdot,\cdot)$ in (7) and the approximation of the right-hand side of (6). In the present analysis, the main issue is the design of a VEM stabilizing form that is suited for the non-linearity under consideration. Following the usual procedure in the VEM setting, we need to construct discrete forms that are computable employing the DoF values only. In the light of property **(P3)** we define the computable discrete form

$$a_h^E(\boldsymbol{v}_h,\boldsymbol{w}_h) := \int_E \boldsymbol{\sigma}(\cdot,\boldsymbol{\Pi}_{k-1}^{0,E}\boldsymbol{\epsilon}(\boldsymbol{v}_h)) : \boldsymbol{\Pi}_{k-1}^{0,E}\boldsymbol{\epsilon}(\boldsymbol{w}_h) + S^E((I-\Pi_k^{0,E})\boldsymbol{v}_h,\,(I-\Pi_k^{0,E})\boldsymbol{w}_h)\,,$$

where $S^E(\cdot,\cdot) : \boldsymbol{U}_h(E)\times\boldsymbol{U}_h(E) \to \mathbb{R}$ is the VEM stabilizing term. Many examples for the linear case can be found in the VEM literature [23, 14, 22]. In the present paper, we consider in the non-linear setting the so-called `dofi-dofi` stabilization defined as follows. Let N_E be the dimension of $\boldsymbol{U}_h(E)$ and let, for $i = 1,\dots,N_E$, $\chi_i : \boldsymbol{U}_h(E) \to \mathbb{R}$ be the function that associates to each $\boldsymbol{v}_h \in \boldsymbol{U}_h(E)$ the value of the i-th local degree of freedom in **(P2)**, and let $\boldsymbol{\chi} \in \mathbb{R}^{N_E}$ be the corresponding vector. Then, we propose the following two choices for S^E (coherently with the choice in (2))

$$\begin{aligned} S_1^E(\boldsymbol{v}_h,\boldsymbol{w}_h) &:= \overline{\mu}_E\left(\delta^\alpha + h_E^{-\alpha}|\boldsymbol{\chi}(\boldsymbol{v}_h)|^\alpha\right)^{\frac{r-2}{\alpha}}\boldsymbol{\chi}(\boldsymbol{v}_h)\cdot\boldsymbol{\chi}(\boldsymbol{w}_h)\,,\\ S_2^E(\boldsymbol{v}_h,\boldsymbol{w}_h) &:= \overline{\mu}_E\sum_{i=1}^{N_E}\left(\delta^\alpha + h_E^{-\alpha}|\chi_i(\boldsymbol{v}_h)|^\alpha\right)^{\frac{r-2}{\alpha}}\chi_i(\boldsymbol{v}_h)\chi_i(\boldsymbol{w}_h)\,, \end{aligned} \tag{28}$$

where $\overline{\mu}_E = \Pi_0^{0,E}\mu$. In Section 3.3.1 we will show that the above choice indeed satisfies a suitable stability property. Since the two forms above are equivalent, see Lemma 6, the following results apply to each of them. Whenever a technical difference arises, for ease of exposition we will restrict the discussion to the S_1^E choice.

Let $S(\cdot,\cdot) : \boldsymbol{U}_h \times \boldsymbol{U}_h \to \mathbb{R}$ be defined by

$$S(\boldsymbol{v}_h,\boldsymbol{w}_h) := \sum_{E\in\Omega_h} S^E(\boldsymbol{v}_h,\boldsymbol{w}_h) \qquad \text{for all } \boldsymbol{v}_h,\boldsymbol{w}_h \in \boldsymbol{U}_h. \tag{29}$$

Then the global forms $a : \boldsymbol{U}_h \times \boldsymbol{U}_h \to \mathbb{R}$ is defined by summing the local contributions, i.e.

$$a_h(\boldsymbol{v}_h,\boldsymbol{w}_h) := \int_\Omega \boldsymbol{\sigma}(\cdot,\boldsymbol{\Pi}_{k-1}^0\boldsymbol{\epsilon}(\boldsymbol{v}_h)) : \boldsymbol{\Pi}_{k-1}^0\boldsymbol{\epsilon}(\boldsymbol{w}_h) + S((I-\Pi_k^0)\boldsymbol{v}_h,(I-\Pi_k^0)\boldsymbol{w}_h) \qquad \text{for all } \boldsymbol{v}_h,\boldsymbol{w}_h \in \boldsymbol{U}_h. \tag{30}$$

We finally define the discrete external force

$$\boldsymbol{f}_h := \Pi_k^0\boldsymbol{f}\,, \tag{31}$$

and observe that the ensuing right-hand side $(\boldsymbol{f}_h,\boldsymbol{v}_h)$ is computable by property **(P3)**.

Remark 5. (Choiche of δ) To avoid a more cumbersome analysis, in the following, we tailor our theoretical developments for the more challenging and interesting power-law model, i.e. (2) with $\delta = 0$. Our derivations could be extended to the case $\delta > 0$ combining the present results with, e.g., the approaches in [8, 10, 9].

3.3.1 Properties of the stabilization form

The following lemma will be useful in the sequel.

Lemma 6. *Let the mesh regularity assumptions stated in Assumption 2 hold. For any $E \in \Omega_h$ and for all $\boldsymbol{v}_h \in \boldsymbol{U}_h(E)$ we have*

$$S_1^E(\boldsymbol{v}_h, \boldsymbol{v}_h) \simeq S_2^E(\boldsymbol{v}_h, \boldsymbol{v}_h) \simeq h_E^{2-r} |\boldsymbol{\chi}(\boldsymbol{v}_h)|^r \simeq h_E^{2-r} \max_{1 \le i \le N_E} |\chi_i(\boldsymbol{v}_h)|^r .$$

Proof. By the equivalence of vector norms, since N_E is uniformly bounded for any $E \in \Omega_h$, we have

$$S_1^E(\boldsymbol{v}_h, \boldsymbol{v}_h) := \overline{\mu}_E h_E^{2-r} |\boldsymbol{\chi}(\boldsymbol{v}_h)|^r \simeq h_E^{2-r} \max_{1 \le i \le N_E} |\chi_i(\boldsymbol{v}_h)|^r .$$

Analogously,

$$S_2^E(\boldsymbol{v}_h, \boldsymbol{v}_h) := \overline{\mu}_E h_E^{2-r} \sum_{i=1}^{N_E} |\chi_i(\boldsymbol{v}_h)|^r \simeq h_E^{2-r} \max_{1 \le i \le N_E} |\chi_i(\boldsymbol{v}_h)|^r \simeq h_E^{2-r} |\boldsymbol{\chi}(\boldsymbol{v}_h)|^r .$$

□

We now show the main coercivity and continuity bounds in $\boldsymbol{W}^{1,r}(E)$ for the stabilization forms; later we will collect such results into a single Corollary taking into account also the presence of the symmetric gradient.

Lemma 7. *Let the mesh regularity assumptions stated in Assumption 2 hold. For any $E \in \Omega_h$ we have*

$$|\boldsymbol{v}_h|^r_{\boldsymbol{W}^{1,r}(E)} \lesssim S^E(\boldsymbol{v}_h, \boldsymbol{v}_h) \qquad \text{for all } \boldsymbol{v}_h \in \boldsymbol{U}_h(E).$$

Proof. Let $E \in \Omega_h$ and $\boldsymbol{v}_h \in \boldsymbol{U}_h(E)$. Then first by a Hölder inequality with exponents $(\frac{2}{r}, \frac{2}{2-r})$ and recalling $|E| \simeq h_E^2$, then applying [66, Theorem 2], we obtain

$$|\boldsymbol{v}_h|^r_{\boldsymbol{W}^{1,r}(E)} \lesssim h_E^{2-r} |\boldsymbol{v}_h|^r_{\boldsymbol{W}^{1,2}(E)} \lesssim h_E^{2-r} |\boldsymbol{\chi}(\boldsymbol{v}_h)|^r .$$

The result follows combining the bound above with Lemma 6. □

Lemma 8. *Let the mesh regularity assumptions stated in Assumption 2 hold. For any $E \in \Omega_h$ we have*

$$S^E(\boldsymbol{v}_h, \boldsymbol{v}_h) \lesssim |\boldsymbol{v}_h|^r_{\boldsymbol{W}^{1,r}(E)} \qquad \text{for all } \boldsymbol{v}_h \in \boldsymbol{U}_h(E) \text{ s.t. } \Pi_k^{0,E} \boldsymbol{v}_h = 0.$$

Proof. Let $E \in \Omega_h$ and $\boldsymbol{v}_h \in \boldsymbol{U}_h(E)$. Recalling Lemma 6, it is sufficient to show that

$$\max_{1 \le i \le N_E} h_E^{2-r} |\chi_i(\boldsymbol{v}_h)|^r \lesssim |\boldsymbol{v}_h|^r_{\boldsymbol{W}^{1,r}(E)} \qquad \text{for } 1 \le i \le N_E,$$

uniformly in E and $\boldsymbol{v}_h$. We need to handle boundary and bulk types of degrees of freedom separately.

Boundary DoFs. Let χ_i be a degree of freedom of type $\mathbf{D}_U\mathbf{1}$ or $\mathbf{D}_U\mathbf{2}$, that is a pointwise evaluation at a generic point v on ∂E. Then, by recalling that $\boldsymbol{v}_h$ is piecewise polynomial on ∂E and applying an inverse estimate, we get

$$|\chi_i(\boldsymbol{v}_h)|^r = |\boldsymbol{v}_h(v)|^r \le \|\boldsymbol{v}_h\|^r_{\boldsymbol{L}^\infty(\partial E)} \le h_E^{-1} \|\boldsymbol{v}_h\|^r_{\boldsymbol{L}^r(\partial E)} .$$

Since $\Pi_k^{0,E} \boldsymbol{v}_h = 0$, we have $\int_E \boldsymbol{v}_h = 0$. Therefore, we obtain the bound by applying the trace inequality (21) followed by Poincaré's inequality

$$h_E^{2-r} |\chi_i(\boldsymbol{v}_h)|^r \le h_E^{1-r} \|\boldsymbol{v}_h\|^r_{\boldsymbol{L}^r(\partial E)} \le h_E^{1-r} \left(h_E^{-1} \|\boldsymbol{v}_h\|^r_{\boldsymbol{L}^r(E)} + h_E^{r-1} |\boldsymbol{v}_h|^r_{\boldsymbol{W}^{1,r}(E)} \right) \lesssim |\boldsymbol{v}_h|^r_{\boldsymbol{W}^{1,r}(E)} .$$

Bulk DoFs. Let now χ_i be a degree of freedom of type $\mathbf{D_U 3}$. We focus on this case only since the proof for $\mathbf{D_U 4}$ follows with very similar steps. Then, for $m_\alpha \in \mathbb{M}_{k-3}(E)$,

$$|\chi_i(\boldsymbol{v}_h)| = \left|\frac{1}{|E|}\int_E \boldsymbol{v}_h \cdot \mathbf{m}^\perp m_\alpha \mathrm{d}E\right|.$$

We apply the Hölder inequality, recall that $|E| \simeq h_E^2$, and finally the inverse estimate (22) on the polynomial m_α, yielding

$$\begin{aligned} |\chi_i(\boldsymbol{v}_h)| &\lesssim h_E^{-2}\|\boldsymbol{v}_h\|_{\boldsymbol{L}^r(E)}\|\boldsymbol{m}^\perp\|_{\boldsymbol{L}^\infty(E)}\|m_\alpha\|_{L^{r'}(E)} \lesssim h_E^{-2}\|\boldsymbol{v}_h\|_{\boldsymbol{L}^r(E)} h_E^{2/r'}\|m_\alpha\|_{L^\infty(E)} \\ &\lesssim h_E^{-2} h_E^{2(1-1/r)}\|\boldsymbol{v}_h\|_{\boldsymbol{L}^r(E)} = h_E^{-2/r}\|\boldsymbol{v}_h\|_{\boldsymbol{L}^r(E)}, \end{aligned} \tag{32}$$

where we also used that $\|\mathbf{m}^\perp\|_{\boldsymbol{L}^\infty(E)}$ and $\|m_\alpha\|_{\boldsymbol{L}^\infty(E)} \lesssim 1$. As mentioned above, we recall that $\int_E \boldsymbol{v}_h = 0$. First using the above bound, then applying a (scaled) Poincaré inequality, finally gives

$$h_E^{2-r}|\chi_i(\boldsymbol{v}_h)|^r \lesssim h_E^{-r}\|\boldsymbol{v}_h\|^r_{\boldsymbol{L}^r(E)} \lesssim |\boldsymbol{v}_h|^r_{\boldsymbol{W}^{1,r}(E)}.$$

□

The next instrumental result follows from Lemma 6, Lemma 7, and Lemma 8.

Corollary 9. *Let the mesh regularity assumptions stated in Assumption 2 hold. Then for any $E \in \Omega_h$ we have the following inverse estimate*

$$|\boldsymbol{v}_h|_{\boldsymbol{W}^{1,2}(E)} \lesssim h_E^{(r-2)/r}|\boldsymbol{v}_h|_{\boldsymbol{W}^{1,r}(E)} \qquad \text{for all } \boldsymbol{v}_h \in \boldsymbol{U}_h(E). \tag{33}$$

Proof. We start by combining Lemma 7 (for $r = 2$) and Lemma 6 (still for $r = 2$), obtaining

$$|\boldsymbol{v}_h|^2_{\boldsymbol{W}^{1,2}(E)} \lesssim \max_{1\le i\le N_E}|\chi_i(\boldsymbol{v}_h)|^2.$$

We now take the square root and, using that N_E is uniformly bounded, manipulate as follows

$$|\boldsymbol{v}_h|_{\boldsymbol{W}^{1,2}(E)} \lesssim \max_{1\le i\le N_E}|\chi_i(\boldsymbol{v}_h)| \lesssim h_E^{(r-2)/r}\left(h_E^{2-r}\max_{1\le i\le N_E}|\chi_i(\boldsymbol{v}_h)|^r\right)^{1/r}.$$

We now apply Lemma 6 and Lemma 8, yielding

$$|\boldsymbol{v}_h|_{\boldsymbol{W}^{1,2}(E)} \lesssim h_E^{(r-2)/r} S_E(\boldsymbol{v}_h,\boldsymbol{v}_h)^{1/r} \lesssim h_E^{(r-2)/r}|\boldsymbol{v}_h|_{\boldsymbol{W}^{1,r}(E)},$$

where we observe that in Lemma 8 it is sufficient that $\int_E \boldsymbol{v}_h = 0$, something which is not restrictive to assume here since this result only involves semi-norms. □

Corollary 10. *Let the mesh regularity assumptions stated in Assumption 2 hold. For any $E \in \Omega_h$ we have*

$$\|\boldsymbol{\epsilon}(\boldsymbol{v}_h)\|^r_{\mathbb{L}^r(E)} \lesssim S^E(\boldsymbol{v}_h,\boldsymbol{v}_h) \lesssim \|\boldsymbol{\epsilon}(\boldsymbol{v}_h)\|^r_{\mathbb{L}^r(E)} \qquad \text{for all } \boldsymbol{v}_h \in \boldsymbol{U}_h(E) \text{ s.t. } \Pi_k^{0,E}\boldsymbol{v}_h = 0.$$

Proof. The first bound follows immediately from Lemma 7. The second bound is a consequence of Lemma 8 combined with Korn's second inequality, cf. [71, Theorem 2]. Indeed, the assumption $\Pi_k^{0,E}\boldsymbol{v}_h = 0$ for $k \ge 1$ implies that $\boldsymbol{v}_h \in \boldsymbol{U}_h(E)$ is rigid body motion-free and, as a result of [62, Theorem 3.3 and Remark 3.4], we have

$$S^E(\boldsymbol{v}_h,\boldsymbol{v}_h) \lesssim |\boldsymbol{v}_h|^r_{W^{1,r}(E)} \lesssim \|\boldsymbol{\epsilon}(\boldsymbol{v}_h)\|^r_{L^r(E)},$$

with hidden constants depending on r but independent of E due to Assumption 2. □

We close this section with the following results regarding the stabilization form.

Lemma 11 (Strong monotonicity of $S^E(\cdot,\cdot)$)**.** *Let the mesh regularity assumptions stated in Assumption 2 hold. Let $\boldsymbol{u}_h$, $\boldsymbol{w}_h \in \boldsymbol{U}_h(E)$ and set $\boldsymbol{e}_h := \boldsymbol{u}_h - \boldsymbol{w}_h$. Then there holds*

$$S^E(\boldsymbol{u}_h,\boldsymbol{e}_h) - S^E(\boldsymbol{w}_h,\boldsymbol{e}_h) \gtrsim S^E(\boldsymbol{e}_h,\boldsymbol{e}_h)^{\frac{2}{r}} \left(h_E^{2-r}(|\chi(\boldsymbol{u}_h)|^r + |\chi(\boldsymbol{w}_h)|^r)\right)^{\frac{r-2}{r}}. \tag{34}$$

Moreover if $\boldsymbol{u}_h$ and $\boldsymbol{w}_h$ are s.t. $\Pi_k^{0,E}\boldsymbol{u}_h = \Pi_k^{0,E}\boldsymbol{w}_h = \boldsymbol{0}$, then

$$S^E(\boldsymbol{u}_h,\boldsymbol{e}_h) - S^E(\boldsymbol{w}_h,\boldsymbol{e}_h) \gtrsim \|\boldsymbol{\epsilon}(\boldsymbol{e}_h)\|_{\mathbb{L}^r(E)}^2 \left(\|\boldsymbol{\epsilon}(\boldsymbol{u}_h)\|_{\mathbb{L}^r(E)}^r + \|\boldsymbol{\epsilon}(\boldsymbol{w}_h)\|_{\mathbb{L}^r(E)}^r\right)^{\frac{r-2}{r}}. \tag{35}$$

Proof. Employing (4) with $\boldsymbol{x} = \chi(\boldsymbol{u}_h)$, $\boldsymbol{y} = \chi(\boldsymbol{w}_h)$ and $\delta = 0$ and recalling the definition of $S_1^E(\cdot,\cdot)$ in (28) we infer

$$\begin{aligned} |\chi(\boldsymbol{e}_h)|^2(|\chi(\boldsymbol{u}_h)|^r + |\chi(\boldsymbol{w}_h)|^r)^{\frac{r-2}{r}} &\lesssim \left((|\chi(\boldsymbol{u}_h)|^{r-2}\chi(\boldsymbol{u}_h) - |\chi(\boldsymbol{w}_h)|^{r-2}\chi(\boldsymbol{w}_h)\right)\cdot\chi(\boldsymbol{e}_h) \\ &\lesssim h_E^{r-2}\left(S_1^E(\boldsymbol{u}_h,\boldsymbol{e}_h) - S_1^E(\boldsymbol{w}_h,\boldsymbol{e}_h)\right). \end{aligned} \tag{36}$$

Owing to Lemma 6 it holds $S_1^E(\boldsymbol{e}_h,\boldsymbol{e}_h)^{2/r} \simeq h_E^{\frac{2(2-r)}{r}}|\chi(\boldsymbol{e}_h)|^2$. Therefore from (36) we derive

$$h_E^{\frac{2(r-2)}{r}} S_1^E(\boldsymbol{e}_h,\boldsymbol{e}_h)^{\frac{2}{r}}(|\chi(\boldsymbol{u}_h)|^r + |\chi(\boldsymbol{w}_h)|^r)^{\frac{r-2}{r}} \lesssim h_E^{r-2}\left(S_1^E(\boldsymbol{u}_h,\boldsymbol{e}_h) - S_1^E(\boldsymbol{w}_h,\boldsymbol{e}_h)\right).$$

Bound (34) easily follows from the bound above. Bound (35) follows combining in (34) Lemma 6 and Corollary 10.

The bounds for the stabilization $S_2^E(\cdot,\cdot)$ in (28) can be derived using analogous arguments. □

Corollary 12 (Strong monotonicity of $S(\cdot,\cdot)$)**.** *Let the mesh regularity assumptions stated in Assumption 2 hold. Let $\boldsymbol{u}_h$, $\boldsymbol{w}_h \in \boldsymbol{U}_h$ and set $\boldsymbol{e}_h := \boldsymbol{u}_h - \boldsymbol{w}_h$. Then there holds*

$$S(\boldsymbol{u}_h,\boldsymbol{e}_h) - S(\boldsymbol{w}_h,\boldsymbol{e}_h) \gtrsim S(\boldsymbol{e}_h,\boldsymbol{e}_h)^{\frac{2}{r}} \left(\sum_{E\in\Omega_h} h_E^{2-r}(|\chi(\boldsymbol{u}_h)|^r + |\chi(\boldsymbol{w}_h)|^r)\right)^{\frac{r-2}{r}}. \tag{37}$$

Moreover if $\boldsymbol{u}_h$ and $\boldsymbol{w}_h$ are s.t. $\Pi_k^0\boldsymbol{u}_h = \Pi_k^0\boldsymbol{w}_h = \boldsymbol{0}$ then

$$S(\boldsymbol{u}_h,\boldsymbol{e}_h) - S(\boldsymbol{w}_h,\boldsymbol{e}_h) \gtrsim \|\boldsymbol{\epsilon}(\boldsymbol{e}_h)\|_{\mathbb{L}^r(\Omega_h)}^2 \left(\|\boldsymbol{\epsilon}(\boldsymbol{u}_h)\|_{\mathbb{L}^r(\Omega_h)}^r + \|\boldsymbol{\epsilon}(\boldsymbol{w}_h)\|_{\mathbb{L}^r(\Omega_h)}^r\right)^{\frac{r-2}{r}}. \tag{38}$$

Proof. Applying Lemma 11 and employing the Hölder inequality with exponents $(\frac{2}{r}, \frac{2}{2-r})$, direct computations yield

$$\begin{aligned} S(\boldsymbol{e}_h,\boldsymbol{e}_h) = \sum_{E\in\Omega_h} S^E(\boldsymbol{e}_h,\boldsymbol{e}_h) &\lesssim \sum_{E\in\Omega_h} \left(S^E(\boldsymbol{u}_h,\boldsymbol{e}_h) - S^E(\boldsymbol{w}_h,\boldsymbol{e}_h)\right)^{\frac{r}{2}} \left(h_E^{2-r}(|\chi(\boldsymbol{u}_h)|^r + |\chi(\boldsymbol{w}_h)|^r)\right)^{\frac{2-r}{2}} \\ &\lesssim \left(S(\boldsymbol{u}_h,\boldsymbol{e}_h) - S(\boldsymbol{w}_h,\boldsymbol{e}_h)\right)^{\frac{r}{2}} \left(\sum_{E\in\Omega_h} h_E^{2-r}(|\chi(\boldsymbol{u}_h)|^r + |\chi(\boldsymbol{w}_h)|^r)\right)^{\frac{2-r}{2}} \end{aligned}$$

Raising both sides of the bound above to $2/r$ we obtain (37). Recalling definition (19) and employing again the Hölder inequality with exponents $(\frac{2}{r}, \frac{2}{2-r})$, bound (38) can be derived from (35) as follows

$$\begin{aligned} \|\boldsymbol{\epsilon}(\boldsymbol{e}_h)\|_{\mathbb{L}^r(\Omega_h)}^r &\lesssim \sum_{E\in\Omega_h} \left(S^E(\boldsymbol{u}_h,\boldsymbol{e}_h) - S^E(\boldsymbol{w}_h,\boldsymbol{e}_h)\right)^{\frac{r}{2}} \left(\|\boldsymbol{\epsilon}(\boldsymbol{u}_h)\|_{\mathbb{L}^r(E)}^r + \|\boldsymbol{\epsilon}(\boldsymbol{w}_h)\|_{\mathbb{L}^r(E)}^r\right)^{\frac{2-r}{2}} \\ &\lesssim \left(S(\boldsymbol{u}_h,\boldsymbol{e}_h) - S(\boldsymbol{w}_h,\boldsymbol{e}_h)\right)^{\frac{r}{2}} \left(\|\boldsymbol{\epsilon}(\boldsymbol{u}_h)\|_{\mathbb{L}^r(\Omega_h)}^r + \|\boldsymbol{\epsilon}(\boldsymbol{w}_h)\|_{\mathbb{L}^r(\Omega_h)}^r\right)^{\frac{2-r}{2}}. \end{aligned}$$

Raising both sides of the bound above to $2/r$ we obtain (38). □

Lemma 13 (Hölder continuity of $S^E(\cdot,\cdot)$)**.** *Let the mesh regularity assumptions stated in Assumption 2 hold. Let $\boldsymbol{u}_h$, $\boldsymbol{w}_h \in \boldsymbol{U}_h(E)$ and set $\boldsymbol{e}_h := \boldsymbol{u}_h - \boldsymbol{w}_h$. Then there holds*

$$|S^E(\boldsymbol{u}_h,\boldsymbol{v}_h) - S^E(\boldsymbol{w}_h,\boldsymbol{v}_h)| \lesssim h_E^{2-r}|\chi(\boldsymbol{e}_h)|^{r-1}|\chi(\boldsymbol{v}_h)| \quad \textit{for all } \boldsymbol{v}_h \in \boldsymbol{U}_h(E). \tag{39}$$

Moreover if $\boldsymbol{u}_h$, $\boldsymbol{w}_h$ and $\boldsymbol{v}_h$ are s.t. $\Pi_k^{0,E}\boldsymbol{u}_h = \Pi_k^{0,E}\boldsymbol{w}_h = \Pi_k^{0,E}\boldsymbol{v}_h = \boldsymbol{0}$, then

$$|S^E(\boldsymbol{u}_h,\boldsymbol{v}_h) - S^E(\boldsymbol{w}_h,\boldsymbol{v}_h)| \lesssim \|\boldsymbol{e}_h\|_{\boldsymbol{W}^{1,r}(E)}^{r-1}\|\boldsymbol{v}_h\|_{\boldsymbol{W}^{1,r}(E)}. \tag{40}$$

Proof. Recalling the definition of $S_1^E(\cdot,\cdot)$ in (28) with $\delta = 0$, employing (3a) and (10), bound (39) can be derived as follows

$$\begin{aligned} S_1^E(\boldsymbol{u}_h, \boldsymbol{v}_h) - S_1^E(\boldsymbol{w}_h, \boldsymbol{v}_h) &= h_E^{2-r} \left(|\boldsymbol{\chi}(\boldsymbol{u}_h)|^{r-2}\boldsymbol{\chi}(\boldsymbol{u}_h) - |\boldsymbol{\chi}(\boldsymbol{w}_h)|^{r-2}\boldsymbol{\chi}(\boldsymbol{w}_h)) \right) \cdot \boldsymbol{\chi}(\boldsymbol{v}_h) \\ &\lesssim h_E^{2-r} \left(|\boldsymbol{\chi}(\boldsymbol{u}_h)|^r + |\boldsymbol{\chi}(\boldsymbol{w}_h)|^r \right)^{\frac{r-2}{r}} |\boldsymbol{\chi}(\boldsymbol{e}_h)|\, |\boldsymbol{\chi}(\boldsymbol{v}_h)| \\ &\lesssim h_E^{2-r} |\boldsymbol{\chi}(\boldsymbol{e}_h)|^{r-2} |\boldsymbol{\chi}(\boldsymbol{e}_h)|\, |\boldsymbol{\chi}(\boldsymbol{v}_h)| = h_E^{2-r} |\boldsymbol{\chi}(\boldsymbol{e}_h)|^{r-1}\, |\boldsymbol{\chi}(\boldsymbol{v}_h)|\,. \end{aligned}$$

Bound (40) is a direct consequence of (39), Lemma 6 and Lemma 8. The result for $S_2^E(\cdot,\cdot)$ follows by the same argument applying the Hölder inequality. □

Lemma 14 (Hölder continuity of $S(\cdot,\cdot)$)**.** *Let the mesh regularity assumptions stated in Assumption 2 hold. Let $\boldsymbol{u}_h$, $\boldsymbol{w}_h \in \boldsymbol{U}_h$ and set $\boldsymbol{e}_h := \boldsymbol{u}_h - \boldsymbol{w}_h$. Then there holds*

$$|S(\boldsymbol{u}_h, \boldsymbol{v}_h) - S(\boldsymbol{w}_h, \boldsymbol{v}_h)| \lesssim \sum_{E\in\Omega_h} h_E^{2-r} |\boldsymbol{\chi}(\boldsymbol{e}_h)|^{r-1} |\boldsymbol{\chi}(\boldsymbol{v}_h)| \quad \text{for all } \boldsymbol{v}_h \in \boldsymbol{U}_h(E). \tag{41}$$

Moreover if $\boldsymbol{u}_h$ and $\boldsymbol{w}_h$ are s.t. $\Pi_k^0 \boldsymbol{u}_h = \Pi_k^0 \boldsymbol{w}_h = \boldsymbol{0}$, then

$$|S(\boldsymbol{u}_h, \boldsymbol{v}_h) - S(\boldsymbol{w}_h, \boldsymbol{v}_h)| \lesssim \|\boldsymbol{e}_h\|^{r-1}_{\boldsymbol{W}^{1,r}(\Omega_h)} \|\boldsymbol{v}_h\|_{\boldsymbol{W}^{1,r}(\Omega_h)}\,. \tag{42}$$

Proof. Bound (41) easily follows from (39). Employing the the Hölder inequality with exponents (r', r), bound (42) can be derived from (40) as follows

$$\begin{aligned} |S(\boldsymbol{u}_h, \boldsymbol{v}_h) - S(\boldsymbol{w}_h, \boldsymbol{v}_h)| &\le \sum_{E\in\Omega_h} |S^E(\boldsymbol{u}_h, \boldsymbol{v}_h) - S^E(\boldsymbol{w}_h, \boldsymbol{v}_h)| \lesssim \sum_{E\in\Omega_h} \|\boldsymbol{e}_h\|^{r-1}_{\boldsymbol{W}^{1,r}(E)} \|\boldsymbol{v}_h\|_{\boldsymbol{W}^{1,r}(E)} \\ &\lesssim \|\boldsymbol{e}_h\|^{r-1}_{\boldsymbol{W}^{1,r}(\Omega_h)} \|\boldsymbol{v}_h\|_{\boldsymbol{W}^{1,r}(\Omega_h)}\,. \end{aligned}$$

□

3.3.2 Discrete inf-sup condition

We here prove a discrete inf-sup condition analogous to the continuous one (12). Note that the argument, especially in the first part, has important differences from that developed in [19] for $r = 2$, because here we cannot exploit a "minimum energy" argument on the elements. We use a more direct approach based on the previous lemmas regarding the discrete stability form.

Lemma 15 (Discrete inf-sup)**.** *Let the mesh regularity assumptions stated in Assumption 2 hold. Then, for any $r \in (1,\infty)$ it exists a constant $\overline{\beta}(r)$, such that*

$$\inf_{q_h\in P_h} \sup_{\boldsymbol{w}_h\in\boldsymbol{U}_h} \frac{b(\boldsymbol{w}_h, q_h)}{\|q_h\|_{L^{r'}(\Omega)} \|\boldsymbol{w}_h\|_{\boldsymbol{W}^{1,r}(\Omega)}} \ge \overline{\beta}(r) > 0.$$

Proof. Due to (12), it is sufficient to show the existence of a Fortin operator, see for example [30]. The proof is divided into two parts.

Part 1. We start by introducing a suitable lowest-order Clément type interpolant in $\boldsymbol{U}_h$ which, given any $\mathbf{w} \in \boldsymbol{W}^{1,r}_{0,\Gamma_D}(\Omega)$, we will denote by $\mathbf{w}_c$. Given any vertex v of the mesh, we denote by ω_v the set given by the union of all elements in Ω_h sharing v as a vertex. Given $E \in \Omega_h$, we denote by ω_E the union of all the ω_v for v vertexes of E. We will prove that, for any $E \in \Omega_h$ and $\mathbf{w} \in \boldsymbol{W}^{1,r}_{0,\Gamma_D}(\Omega)$, the quasi-interpolant satisfies

$$\|\mathbf{w} - \mathbf{w}_c\|_{\boldsymbol{L}^r(E)} + h_E |\mathbf{w} - \mathbf{w}_c|_{\boldsymbol{W}^{1,r}(E)} \lesssim h_E |\mathbf{w}|_{\boldsymbol{W}^{1,r}(\omega_E)}. \tag{43}$$

We start by defining $\mathbf{w}_c \in \boldsymbol{U}_h$ through its DoF values as follows. We initially impose

$$\mathbf{w}_c(v) = \frac{1}{|\omega_v|} \int_{\omega_v} \mathbf{w}\, \mathrm{d}\omega_v \qquad \text{for any } v \text{ vertex of } \Omega_h \text{ not in } \Gamma_D.$$

The remaining skeletal DoFs, which are the point-wise evaluations on edges, are simply set by linearly interpolating the corresponding vertex values of each edge. We further set the DoFs $\mathbf{D}_{\boldsymbol{U}}\mathbf{3}$ as follows

$$\frac{1}{|E|} \int_E \boldsymbol{w}_c \cdot \boldsymbol{m}^\perp m_{\boldsymbol{\alpha}}\, \mathrm{d}E := \frac{1}{|E|} \int_E \boldsymbol{w} \cdot \boldsymbol{m}^\perp m_{\boldsymbol{\alpha}}\, \mathrm{d}E \qquad \text{for any } m_{\boldsymbol{\alpha}} \in \mathbb{M}_{k-3}(E),$$

and we enforce all degrees of freedom $\mathbf{D_U 4}$ of $\boldsymbol{w}_c$ equal to zero.

It is clear that the above operator preserves constants, in the sense that if $E \in \Omega_h$ is an element without vertexes on Γ_D, then if $\mathbf{w}|_{\omega_E} = \mathbf{p}_0 \in [\mathbb{P}_0(\omega_E)]^2$ it will hold $\mathbf{w}_{c|E} = \mathbf{p}_0$. Furthermore, from the DoFs definition and employing (32), it is easy to check that for any $E \in \Omega_h$

$$|\chi(\mathbf{w}_c)| \lesssim \max_{\nu \in \partial E \text{ vertex}} |\mathbf{w}_c(\nu)| + h_E^{-2/r} \|\mathbf{w}_c\|_{\boldsymbol{L}^r(E)} \,. \tag{44}$$

We now start from a well known (scaled) Poincarè type inequality (the uniformity of the involved constant on E following from Assumption 2) and afterwards apply Lemma 7 and a trivial calculation

$$\|\mathbf{w}_c\|_{\boldsymbol{L}^r(E)} \lesssim h_E |\mathbf{w}_c|_{\boldsymbol{W}^{1,r}(E)} + h_E^{2/r-1} \left| \int_{\partial E} \mathbf{w}_c \right| \lesssim h_E S^E(\mathbf{w}_c, \mathbf{w}_c)^{1/r} + h_E^{2/r} \max_{\nu \in \partial E \text{ vertex}} |\mathbf{w}_c(\nu)|.$$

From the above bound, since the maximum of all the vertex value norms is clearly bounded by $|\chi(\mathbf{w}_c)|$, using Lemma 6 we obtain

$$\|\mathbf{w}_c\|_{\boldsymbol{L}^r(E)} \lesssim h_E S^E(\mathbf{w}_c, \mathbf{w}_c)^{1/r} + h_E^{2/r} h_E^{(r-2)/r} S^E(\mathbf{w}_c, \mathbf{w}_c)^{1/r} \lesssim h_E S^E(\mathbf{w}_c, \mathbf{w}_c)^{1/r}. \tag{45}$$

To avoid repetition of similar ideas, we will prove only the bound for the first term in the left-hand side of (43). Furthermore, we will consider only the case $\overline{E} \cap \Gamma_D = \emptyset$, the other one following very similarly but using a classical Poincarè inequality instead of the constant preserving property. By a triangle inequality and recalling that the quasi-interpolant preserves constants (in the sense detailed above), for any $\mathbf{p}_0 \in [\mathbb{P}_0(\Omega)]^2$ we have

$$\|\mathbf{w} - \mathbf{w}_c\|_{\boldsymbol{L}^r(E)} \leq \|\mathbf{w} - \mathbf{p}_0\|_{\boldsymbol{L}^r(E)} + \|(\mathbf{w} - \mathbf{p}_0)_c\|_{\boldsymbol{L}^r(E)}. \tag{46}$$

For the second term, we first apply (45), then Lemma 6 and finally (44):

$$\begin{aligned}\|(\mathbf{w} - \mathbf{p}_0)_c\|_{\boldsymbol{L}^r(E)} &\lesssim h_E S^E((\mathbf{w} - \mathbf{p}_0)_c, (\mathbf{w} - \mathbf{p}_0)_c)^{1/r} \lesssim h_E^{1+(2-r)/r} \max_{1 \leq i \leq N_E} |\chi_i((\mathbf{w} - \mathbf{p}_0)_c)| \\ &\lesssim h_E^{2/r} \max_{\nu \in \partial E \text{ vertex}} |(\mathbf{w} - \mathbf{p}_0)_c(\nu)| + \|\boldsymbol{w} - \mathbf{p}_0\|_{\boldsymbol{L}^r(E)} \,.\end{aligned}$$

By definition of the quasi-interpolant and due to the shape regularity of the mesh (yielding $|\omega_\nu| \simeq h_E^2$ for any ν vertex of E) we obtain from above

$$\|(\mathbf{w} - \mathbf{p}_0)_c\|_{\boldsymbol{L}^r(E)} \lesssim \max_{\nu \in \partial E \text{ vertex}} |\omega_\nu|^{1/r} \left| \frac{1}{|\omega_\nu|} \int_{\omega_\nu} (\mathbf{w} - \mathbf{p}_0) \right| + \|\boldsymbol{w} - \mathbf{p}_0\|_{\boldsymbol{L}^r(E)} \,.$$

Using an Holder inequality (r, r') we obtain

$$\|(\mathbf{w} - \mathbf{p}_0)_c\|_{\boldsymbol{L}^r(E)} \lesssim \max_{\nu \in \partial E \text{ vertex}} \|\mathbf{w} - \mathbf{p}_0\|_{\boldsymbol{L}^r(\omega_\nu)} + \|\boldsymbol{w} - \mathbf{p}_0\|_{\boldsymbol{L}^r(E)} \leq \|\mathbf{w} - \mathbf{p}_0\|_{\boldsymbol{L}^r(\omega_E)}.$$

Combining the above bound with (46) and standard polynomial approximation estimates we finally obtain the desired bound

$$\|\mathbf{w} - \mathbf{w}_c\|_{\boldsymbol{L}^r(E)} \lesssim h_E \, |\mathbf{w}|_{\boldsymbol{W}^{1,r}(\omega_E)}.$$

Part 2. We define a preliminary Fortin operator $\widehat{\Pi} : \boldsymbol{W}^{1,r}_{0,\Gamma_D}(\Omega) \to \boldsymbol{U}_h$ by setting its DoF values as follows. For any $\mathbf{w} \in \boldsymbol{W}^{1,r}_{0,\Gamma_D}(\Omega)$, we set to zero all DOFs of type $\mathbf{D_U 1}$, $\mathbf{D_U 3}$ and of type $\mathbf{D_U 2'}$ (cf. (27)) apart from

$$\frac{1}{|e|} \int_e \widehat{\Pi}\mathbf{w} \cdot \mathbf{n}_e = \frac{1}{|e|} \int_e \mathbf{w} \cdot \mathbf{n}_e \,,$$

on each edge e (not in Γ_D). We set the DoFs $\mathbf{D_U 4}$ as

$$\frac{h_E}{|E|} \int_E (\nabla \cdot \widehat{\Pi}\mathbf{w}) \, m_{\boldsymbol{\alpha}} \, \mathrm{d}E := \frac{h_E}{|E|} \int_E (\nabla \cdot \mathbf{w}) \, m_{\boldsymbol{\alpha}} \, \mathrm{d}E \quad \text{for all } m_{\boldsymbol{\alpha}} \in \mathbb{M}_{k-1}(E) \text{ with } |\boldsymbol{\alpha}| > 0.$$

It is trivial to check that the above operator satisfies the first Fortin condition by construction, that is $b(\widehat{\Pi}\mathbf{w}, q_h) = b(\mathbf{w}, q_h)$ for all $q_h \in P_h$. Furthermore, it also easily follows from the above definition

$$|\chi(\widehat{\Pi}\mathbf{w})| \lesssim \frac{h_E}{|E|} \max_{m_{\boldsymbol{\alpha}} \in \mathbb{M}_{k-1}(E)} \left| \int_E (\nabla \cdot \mathbf{w}) \, m_{\boldsymbol{\alpha}} \, \mathrm{d}E \right| + \max_{e \in \partial E} \frac{1}{|e|} \left| \int_e \mathbf{w} \cdot \mathbf{n}_e \, \mathrm{d}s \right|$$

which combined with a scaled trace inequality and a Holder inequality (r, r') yields

$$|\boldsymbol{\chi}(\widehat{\Pi}\mathbf{w})| \lesssim h_E^{-2/r}\Big(\|\mathbf{w}\|_{\boldsymbol{L}^r(E)} + h_E|\mathbf{w}|_{\boldsymbol{W}^{1,r}(E)}\Big). \tag{47}$$

We can now define the Fortin operator $\Pi^{\mathcal{F}} : \boldsymbol{W}^{1,r}_{0,\Gamma_D}(\Omega) \to \boldsymbol{U}_h$. We set

$$\Pi^{\mathcal{F}}\mathbf{w} = \mathbf{w}_c + \widehat{\Pi}(\mathbf{w} - \mathbf{w}_c) \qquad \forall\, \mathbf{w} \in \boldsymbol{W}^{1,r}_{0,\Gamma_D}(\Omega).$$

It is immediate to check that also $\Pi^{\mathcal{F}}$ satisfies the first Fortin condition, that is $b(\Pi^{\mathcal{F}}\mathbf{w}, q_h) = b(\mathbf{w}, q_h)$ for all $q_h \in P_h$. In order to conclude, we are left to check the second condition for a Fortin operator, which is its continuity in the $\boldsymbol{W}^{1,r}(\Omega)$ norm (see, eg. [30]).

By the triangle inequality and recalling (43), for all $E \in \Omega_h$

$$|\Pi^{\mathcal{F}}\mathbf{w}|_{\boldsymbol{W}^{1,r}(E)} \lesssim |\mathbf{w}|_{\boldsymbol{W}^{1,r}(\omega_E)} + |\widehat{\Pi}(\mathbf{w} - \mathbf{w}_c)|_{\boldsymbol{W}^{1,r}(E)}\,. \tag{48}$$

For the second term on the right hand side, we first apply again Lemma 6 and 7, then recall (47) and finally make use of (43), yielding

$$\begin{aligned}|\widehat{\Pi}(\mathbf{w} - \mathbf{w}_c)|_{\boldsymbol{W}^{1,r}(E)} &\lesssim h_E^{(2-r)/r}|\boldsymbol{\chi}(\widehat{\Pi}(\mathbf{w} - \mathbf{w}_c))| \\ &\lesssim h_E^{-1}\|\mathbf{w} - \mathbf{w}_c\|_{\boldsymbol{L}^r(E)} + |\mathbf{w} - \mathbf{w}_c|_{\boldsymbol{W}^{1,r}(E)} \lesssim |\mathbf{w}|_{\boldsymbol{W}^{1,r}(\omega_E)}.\end{aligned}$$

The continuity in $\boldsymbol{W}^{1,r}(\Omega)$ of $\Pi^{\mathcal{F}}$ follows trivially from (48) and the bound here above, summing on all elements $E \in \Omega_h$ and noting that, due to Assumption 2, $\{\omega_E\}_{E\in\Omega_h}$ overlaps every element a uniformly bounded number of times.

□

3.4 Virtual Element problem

Having in mind the spaces (24) and (25), the discrete form (30), the form (7), the discrete loading term (31), the virtual element discretization of Problem (6) is given by: Find $(\boldsymbol{u}_h, p_h) \in \boldsymbol{U}_h \times P_h$ such that

$$\begin{aligned} a_h(\boldsymbol{u}_h, \boldsymbol{v}_h) + b(\boldsymbol{v}_h, p_h) &= \int_\Omega \boldsymbol{f}_h \cdot \boldsymbol{v}_h + \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{\gamma}(\boldsymbol{v}_h) \qquad && \forall \boldsymbol{v}_h \in \boldsymbol{U}_h, \\ b(\boldsymbol{u}_h, q_h) &= 0 && \forall q_h \in P_h\,. \end{aligned} \tag{49}$$

Recalling the definition of the discrete kernel $\boldsymbol{Z}_h$ in (26), the previous problem can be also written in the kernel formulation: Find $\boldsymbol{u}_h \in \boldsymbol{Z_h}$ such that

$$a_h(\boldsymbol{u}_h, \boldsymbol{v}_h) = \int_\Omega \boldsymbol{f}_h \cdot \boldsymbol{v}_h + \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{\gamma}(\boldsymbol{v}_h) \qquad \forall \boldsymbol{v}_h \in \boldsymbol{Z_h}\,. \tag{50}$$

In order to prove the well-posedness of the discrete problem (49), we establish the strong monotonicity and the Hölder continuity of the non-linear function $a_h(\cdot,\cdot)$.

Lemma 16 (Strong monotonicity of $a_h(\cdot,\cdot)$)**.** *Let $\boldsymbol{u}_h$, $\boldsymbol{w}_h \in \boldsymbol{U}_h$ and set $\boldsymbol{e}_h := \boldsymbol{u}_h - \boldsymbol{w}_h \in \boldsymbol{U}_h$. Then there holds*

$$a_h(\boldsymbol{u}_h, \boldsymbol{e}_h) - a_h(\boldsymbol{w}_h, \boldsymbol{e}_h) \gtrsim \Big(\|\boldsymbol{u}_h\|^r_{\boldsymbol{W}^{1,r}(\Omega)} + \|\boldsymbol{w}_h\|^r_{\boldsymbol{W}^{1,r}(\Omega)}\Big)^{\frac{r-2}{r}} \|\boldsymbol{e}_h\|^2_{\boldsymbol{W}^{1,r}(\Omega)}\,.$$

Proof. Recalling the definition of $a_h(\cdot,\cdot)$ in (30) and definition (29) we have

$$a_h(\boldsymbol{u}_h, \boldsymbol{e}_h) - a_h(\boldsymbol{w}_h, \boldsymbol{e}_h) := T_1 + T_2 \tag{51}$$

where

$$\begin{aligned} T_1 &:= \int_\Omega \boldsymbol{\sigma}(\cdot, \boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h)) : \boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{e}_h) - \int_\Omega \boldsymbol{\sigma}(\cdot, \boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{w}_h)) : \boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{e}_h)\,, \\ T_2 &:= S((I - \Pi^0_k)\boldsymbol{u}_h, (I - \Pi^0_k)\boldsymbol{e}_h) - S((I - \Pi^0_k)\boldsymbol{w}_h, (I - \Pi^0_k)\boldsymbol{e}_h)\,. \end{aligned}$$

We estimate separately each term above. The same identical proof used to derive (9b) (with $\delta = 0$) and the continuity of the L^2-projection with respect to the L^r-norm yield

$$\begin{aligned} T_1 &\gtrsim \sigma_{\mathrm{m}} \left(\|\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h)\|^r_{\mathbb{L}^r(\Omega)} + \|\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{w}_h)\|^r_{\mathbb{L}^r(\Omega)} \right)^{\frac{r-2}{r}} \|\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{e}_h)\|^2_{\mathbb{L}^r(\Omega)} \\ &\gtrsim \sigma_{\mathrm{m}} \left(\|\boldsymbol{\epsilon}(\boldsymbol{u}_h)\|^r_{\mathbb{L}^r(\Omega)} + \|\boldsymbol{\epsilon}(\boldsymbol{w}_h)\|^r_{\mathbb{L}^r(\Omega)} \right)^{\frac{r-2}{r}} \|\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{e}_h)\|^2_{\mathbb{L}^r(\Omega)} \\ &\gtrsim \sigma_{\mathrm{m}} \left(\|\boldsymbol{u}_h\|^r_{\boldsymbol{W}^{1,r}(\Omega)} + \|\boldsymbol{w}_h\|^r_{\boldsymbol{W}^{1,r}(\Omega)} \right)^{\frac{r-2}{r}} \|\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{e}_h)\|^2_{\mathbb{L}^r(\Omega)} . \end{aligned} \tag{52}$$

Employing bound (38) and the properties of the L^2-projection we infer

$$\begin{aligned} T_2 &\gtrsim \left(\|\boldsymbol{\epsilon}((I-\Pi^0_k)\boldsymbol{u}_h)\|^r_{\mathbb{L}^r(\Omega_h)} + \|\boldsymbol{\epsilon}((I-\Pi^0_k)\boldsymbol{w}_h)\|^r_{\boldsymbol{L}^r(\Omega_h)} \right)^{\frac{r-2}{r}} \|\boldsymbol{\epsilon}((I-\Pi^0_k)\boldsymbol{e}_h)\|^2_{\mathbb{L}^r(\Omega_h)} \\ &\geq \left(|(I-\Pi^0_k)\boldsymbol{u}_h|^r_{\boldsymbol{W}^{1,r}(\Omega_h)} + |(I-\Pi^0_k)\boldsymbol{w}_h|^r_{\boldsymbol{W}^{1,r}(\Omega_h)} \right)^{\frac{r-2}{r}} \|(I-\boldsymbol{\Pi}^0_{k-1})\boldsymbol{\epsilon}(\boldsymbol{e}_h)\|^2_{\mathbb{L}^r(\Omega)} \\ &\gtrsim \left(\|\boldsymbol{u}_h\|^r_{\boldsymbol{W}^{1,r}(\Omega_h)} + \|\boldsymbol{w}_h\|^r_{\boldsymbol{W}^{1,r}(\Omega_h)} \right)^{\frac{r-2}{r}} \|(I-\boldsymbol{\Pi}^0_{k-1})\boldsymbol{\epsilon}(\boldsymbol{e}_h)\|^2_{\mathbb{L}^r(\Omega)} . \end{aligned} \tag{53}$$

The proof follows collecting (52) and (53) in (51) and observing that by Korn inequality (5) we have

$$\|\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{e}_h)\|^2_{\mathbb{L}^r(\Omega)} + \|(I-\boldsymbol{\Pi}^0_{k-1})\boldsymbol{\epsilon}(\boldsymbol{e}_h)\|^2_{\mathbb{L}^r(\Omega)} \gtrsim \|\boldsymbol{\epsilon}(\boldsymbol{e}_h)\|^2_{\mathbb{L}^r(\Omega)} \gtrsim \|\boldsymbol{e}_h\|^2_{\boldsymbol{W}^{1,r}(\Omega)} .$$

□

Lemma 17 (Hölder continuity of $a_h(\cdot,\cdot)$)**.** *Let $\boldsymbol{u}_h$, $\boldsymbol{w}_h \in \boldsymbol{U}_h$ and set $\boldsymbol{e}_h := \boldsymbol{u}_h - \boldsymbol{w}_h \in \boldsymbol{U}_h$. Then for any $\boldsymbol{v}_h \in \boldsymbol{U}_h$ there holds*

$$|a_h(\boldsymbol{u}_h,\boldsymbol{v}_h) - a_h(\boldsymbol{w}_h,\boldsymbol{v}_h)| \lesssim \|\boldsymbol{e}_h\|^{r-1}_{\boldsymbol{W}^{1,r}(\Omega)} \|\boldsymbol{v}_h\|_{\boldsymbol{W}^{1,r}(\Omega)} .$$

Proof. Recalling the definition of $a_h(\cdot,\cdot)$ in (30) and definition (29) we have

$$|a_h(\boldsymbol{u}_h,\boldsymbol{v}_h) - a_h(\boldsymbol{w}_h,\boldsymbol{v}_h)| \leq |T_1| + |T_2| \tag{54}$$

where

$$\begin{aligned} T_1 &:= \int_\Omega \boldsymbol{\sigma}(\cdot, \boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h)) : \boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{v}_h) - \int_\Omega \boldsymbol{\sigma}(\cdot, \boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{w}_h)) : \boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{v}_h) , \\ T_2 &:= S((I-\Pi^0_k)\boldsymbol{u}_h, (I-\Pi^0_k)\boldsymbol{v}_h) - S((I-\Pi^0_k)\boldsymbol{w}_h, (I-\Pi^0_k)\boldsymbol{v}_h) . \end{aligned}$$

Using the same computations as in (11) and the continuity of the L^2-projection w.r.t. the L^2-norm we have

$$|T_1| \lesssim \|\boldsymbol{e}_h\|^{r-1}_{\boldsymbol{W}^{1,r}(\Omega)} \|\boldsymbol{v}_h\|_{\boldsymbol{W}^{1,r}(\Omega)} . \tag{55}$$

From (42) and Lemma 4 we infer

$$|T_2| \lesssim \|(I-\Pi^0_k)\boldsymbol{e}_h\|^{r-1}_{\boldsymbol{W}^{1,r}(\Omega)} \|(I-\Pi^0_k)\boldsymbol{v}_h\|_{\boldsymbol{W}^{1,r}(\Omega)} \lesssim \|\boldsymbol{e}_h\|^{r-1}_{\boldsymbol{W}^{1,r}(\Omega)} \|\boldsymbol{v}_h\|_{\boldsymbol{W}^{1,r}(\Omega)} . \tag{56}$$

The proof follows inserting (55) and (56) in (54). □

Theorem 18 (Well-posedness of (50))**.** *For any $r \in (1,2)$, there exists a unique solution $\boldsymbol{u}_h \in \boldsymbol{Z}_h$ to the discrete problem (50) satisfying the a-priori estimate (cf. (15))*

$$\|\boldsymbol{u}_h\|_{\boldsymbol{W}^{1,r}(\Omega)} \lesssim \mathcal{N}(\boldsymbol{f},\boldsymbol{g})^{\frac{1}{r-1}} . \tag{57}$$

Proof. (i) Existence. Let the mesh Ω_h be fixed. We equip the space $\boldsymbol{Z_h}$ with the $\boldsymbol{H}^1$-inner product, denoted by $(\cdot,\cdot)_{\boldsymbol{H}^1}$, and corresponding induced norm $\|\cdot\|_{\boldsymbol{H}^1(\Omega)}$. Owing to the Poincaré inequality, the equivalence of vector norms in finite-dimensional spaces, and the inverse estimate (33), we infer

$$\|\boldsymbol{v}_h\|_{\boldsymbol{H}^1(\Omega)} \lesssim \left(\sum_{E\in\Omega_h} |\boldsymbol{v}_h|^2_{\boldsymbol{H}^1(E)} \right)^{\frac{1}{2}} \lesssim \left(\sum_{E\in\Omega_h} h_E^{r-2} |\boldsymbol{v}_h|^r_{\boldsymbol{W}^{1,r}(E)} \right)^{\frac{1}{r}} \lesssim \left(\min_{E\in\Omega_h} h_E \right)^{\frac{r-2}{r}} \|\boldsymbol{v}_h\|_{\boldsymbol{W}^{1,r}(\Omega)} . \tag{58}$$

Let now $\boldsymbol{\Phi}_h : \boldsymbol{Z_h} \to \boldsymbol{Z_h}$ be defined such that

$$(\boldsymbol{\Phi}_h(\boldsymbol{v}_h), \boldsymbol{w}_h)_{\boldsymbol{H}^1(\Omega)} := a_h(\boldsymbol{v}_h, \boldsymbol{w}_h), \qquad \forall\, \boldsymbol{v}_h, \boldsymbol{w}_h \in \boldsymbol{Z_h}.$$

Thus, the strong monotonicity of $a_h(\cdot,\cdot)$ established in Lemma 16 together with the bound in (58) leads to, for any $\boldsymbol{v}_h \in \boldsymbol{Z_h}$,

$$\begin{aligned}\lim_{\|\boldsymbol{v}_h\|_{\boldsymbol{H}^1(\Omega)} \to \infty} \frac{(\boldsymbol{\Phi}_h(\boldsymbol{v}_h), \boldsymbol{v}_h)_{\boldsymbol{H}^1(\Omega)}}{\|\boldsymbol{v}_h\|_{\boldsymbol{H}^1(\Omega)}} &\gtrsim \lim_{\|\boldsymbol{v}_h\|_{\boldsymbol{H}^1(\Omega)} \to \infty} \frac{\|\boldsymbol{v}_h\|^r_{\boldsymbol{W}^{1,r}(\Omega)}}{\|\boldsymbol{v}_h\|_{\boldsymbol{H}^1(\Omega)}} \\ &\gtrsim \lim_{\|\boldsymbol{v}_h\|_{\boldsymbol{H}^1(\Omega)} \to \infty} \left(\min_{E \in \Omega_h} h_E\right)^{2-r} \|\boldsymbol{v}_h\|^{r-1}_{\boldsymbol{H}^1(\Omega)} \to \infty.\end{aligned}$$

By applying [46, Theorem 3.3], the previous result shows that the operator $\boldsymbol{\Phi}_h$ is surjective. Let $z_h \in \boldsymbol{Z_h}$ be such that

$$(z_h, \boldsymbol{w}_h)_{\boldsymbol{H}^1(\Omega)} = \int_\Omega \boldsymbol{f}_h \cdot \boldsymbol{w}_h + \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{\gamma}(\boldsymbol{w}_h).$$

Hence, as a result of the surjectivity of $\boldsymbol{\Phi}_h$ and the definition of z_h, there exists $\boldsymbol{u}_h \in \boldsymbol{Z_h}$ such that $\boldsymbol{\Phi}_h(\boldsymbol{u}_h) = z_h$, namely $\boldsymbol{u}_h$ is a solution to the discrete problem (50).

(ii) Uniqueness. Let $\boldsymbol{u}_{h,1}, \boldsymbol{u}_{h,2} \in \boldsymbol{Z_h}$ solve (50). Subtracting (50) for $\boldsymbol{u}_{h,2}$ from (50) for $\boldsymbol{u}_{h,1}$ and then taking $\boldsymbol{v}_h = \boldsymbol{u}_{h,1} - \boldsymbol{u}_{h,2}$ as test function, we obtain

$$a_h(\boldsymbol{u}_{h,1}, \boldsymbol{u}_{h,1} - \boldsymbol{u}_{h,2}) - a_h(\boldsymbol{u}_{h,2}, \boldsymbol{u}_{h,1} - \boldsymbol{u}_{h,2}) = 0.$$

Thus, applying the strong monotonicity of $a_h(\cdot,\cdot)$ in Lemma 16 with $\boldsymbol{e}_h = \boldsymbol{u}_{h,1} - \boldsymbol{u}_{h,2}$ and owing to Korn's first inequality (5), we get $\|\boldsymbol{u}_{h,1} - \boldsymbol{u}_{h,2}\|_{\boldsymbol{W}^{1,r}(\Omega)} = 0$ and, as a result, $\boldsymbol{u}_{h,1} = \boldsymbol{u}_{h,2}$.

(iii) A-priori estimate. We follow the same lines of Proposition 3 (with $\delta = 0$), based on the strong monotonicity of $a_h(\cdot,\cdot)$ in Lemma 16. □

Thanks to the discrete inf-sup condition established in Lemma 15 and the equivalence of the discrete problems (49) and (50), the following result hold:

Corollary 19 (Well-posedness of (49))**.** *For any $r \in (1,2)$, there exists a unique solution $(\boldsymbol{u}_h, p_h) \in \boldsymbol{U_h} \times P_h$ to the discrete problem (49). Additionally, p_h satisfies the a-priori bound (cf. (15))*

$$\|p_h\|_{L^{r'}(\Omega)} \lesssim \mathcal{N}(\boldsymbol{f}, \boldsymbol{g})\,. \tag{59}$$

Proof. First, we observe that, by definition, the unique solution $\boldsymbol{u}_h$ to (50) satisfies the second equation in (49). Then, the existence and uniqueness of the discrete pressure solution is obtained as in the linear case $r = 2$, cf. [30, Theorem 4.2.1]. The a-priori estimate (59) follows, proceeding as in Proposition 3, from the discrete inf-sup stability in Lemma 15 and the Hölder continuity of $a_h(\cdot,\cdot)$ in Lemma 17. □

4 A-priori error analysis

4.1 Additional properties of the stress-strain law

We recall some important results regarding the stress-strain relation that are instrumental for the a-priori analysis of the scheme. We mainly follow [27, Section 3] and [57, Section 2]. First, we associate to the Sobolev exponent $r \in (1,2]$ the convex functions $\varphi, \varphi^* \in C^1(\mathbb{R}^+_0, \mathbb{R}^+_0)$ defined by

$$\varphi(t) := \frac{t^r}{r} \qquad \text{and} \qquad \varphi^*(t) := \frac{t^{r'}}{r'}.$$

We also introduce, for $a \geq 0$, the shifted functions $\varphi_a(t) := \int_0^t (a+s)^{r-2} s \,\mathrm{d}s$, which satisfy

$$(r-1)(a+t)^{r-2} \leq \varphi_a''(t) \leq (a+t)^{r-2},$$

from which the convexity of φ_a can be inferred and, as a result, $\varphi_a(t+s) \lesssim \varphi_a(t) + \varphi_a(s)$ uniformly in $t, s, a \in \mathbb{R}^+_0$. The following Lemma provides important properties of the shifted functions φ_a. We refer the reader to [50, Lemmata 28–32] for the detailed proof.

Lemma 20 (Young type inequalities)**.** *For all $\varepsilon > 0$ there exists $C(\varepsilon) > 0$ only depending on r and δ such that for all $s, t, a, \delta \geq 0$ and all $\boldsymbol{\tau}, \boldsymbol{\eta} \in \mathbb{R}^{d\times d}$ there holds*

$$s\varphi_a'(t) + t\varphi_a'(s) \leq \varepsilon\varphi_a(s) + c(\varepsilon)\varphi_a(t), \tag{60}$$

$$\varphi_{\delta+|\boldsymbol{\tau}|}(t) \leq \varepsilon\varphi_{\delta+|\boldsymbol{\eta}|}(|\boldsymbol{\tau}-\boldsymbol{\eta}|) + c(\varepsilon)\varphi_{\delta+|\boldsymbol{\eta}|}(t). \tag{61}$$

The next result showing the equivalence of several quantities is strictly related to the continuity and monotonicity assumptions given in Assumption 1. The proofs of the next lemma can be found in [57, Section 2.3]. The lemma here below applies to any (scalar or tensor valued) function $\boldsymbol{\sigma}$ which satisfies Assumption 1. In the following, with a slight abuse of notation, we will apply such lemma both to the constitutive law $\boldsymbol{\sigma}$ but also to the auxiliary scalar function $\sigma(\tau) := |\tau|^{r-2}\tau$.

Lemma 21. *Let $\boldsymbol{\sigma}$ satisfy (3) for $r \in (1, 2]$ and $\delta \geq 0$. Then, uniformly for all $\boldsymbol{\tau}, \boldsymbol{\eta} \in \mathbb{R}^{d\times d}_{\mathrm{s}}$ and all $\boldsymbol{v}, \boldsymbol{w} \in \boldsymbol{U}$ there hold*

$$|\boldsymbol{\sigma}(\cdot,\boldsymbol{\tau}) - \boldsymbol{\sigma}(\cdot,\boldsymbol{\eta})| \simeq (\delta + |\boldsymbol{\tau}| + |\boldsymbol{\eta}|)^{r-2}\,|\boldsymbol{\tau}-\boldsymbol{\eta}| \simeq \varphi'_{\delta+|\boldsymbol{\tau}|}(|\boldsymbol{\tau}-\boldsymbol{\eta}|), \tag{62a}$$

$$(\boldsymbol{\sigma}(\cdot,\boldsymbol{\tau}) - \boldsymbol{\sigma}(\cdot,\boldsymbol{\eta})) : (\boldsymbol{\tau}-\boldsymbol{\eta}) \simeq (\delta + |\boldsymbol{\tau}| + |\boldsymbol{\eta}|)^{r-2}\,|\boldsymbol{\tau}-\boldsymbol{\eta}|^2 \simeq \varphi_{\delta+|\boldsymbol{\tau}|}(|\boldsymbol{\tau}-\boldsymbol{\eta}|), \tag{62b}$$

where the hidden constants only depend on $\sigma_{\mathrm{c}}, \sigma_{\mathrm{m}}$ in Assumption 1 and on r.

4.2 Convergence

In the present section, we derive the rate of convergence for the proposed VEM scheme (49) in the case $\delta = 0$, see Remark 5.

Proposition 22. *Let $\boldsymbol{u}$ be the solution of problem (8) and let $\boldsymbol{u}_h$ be the solution of problem (50). Assume that $\boldsymbol{u} \in \boldsymbol{W}^{k_1+1,r}(\Omega_h)$, $\boldsymbol{\sigma}(\cdot,\boldsymbol{\epsilon}(\boldsymbol{u})) \in \mathbb{W}^{k_2,r'}(\Omega_h)$, $\boldsymbol{f} \in \boldsymbol{W}^{k_3+1,r'}(\Omega_h)$ for some k_1, k_2, $k_3 \leq k$. Let the mesh regularity assumptions stated in Assumption 2 hold. Then, we have*

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{W}^{1,r}(\Omega_h)} \lesssim h^{k_1 r/2} R_1^{r/2} + h^{k_1} R_1 + h^{k_2} R_2 + h^{k_3+2} R_3\,, \tag{63}$$

where the hidden constant depends on $\mathcal{N}(\boldsymbol{f},\boldsymbol{g})$ (cf. (15)) and the regularity terms are

$$R_1 := |\boldsymbol{u}|_{\boldsymbol{W}^{k_1+1,r}(\Omega_h)}\,, \qquad R_2 := |\boldsymbol{\sigma}(\cdot,\boldsymbol{\epsilon}(\boldsymbol{u}))|_{\mathbb{W}^{k_2,r'}(\Omega_h)}\,, \qquad R_3 := |\boldsymbol{f}|_{\boldsymbol{W}^{k_3+1,r'}(\Omega_h)}\,. \tag{64}$$

Proof. We set $\boldsymbol{\xi}_h := \boldsymbol{u}_h - \boldsymbol{u}_I$. For the sake of brevity in the following we employ the notation $\widetilde{\boldsymbol{v}} := (I - \Pi^0_k)\boldsymbol{v}$ for any $\boldsymbol{v} \in \boldsymbol{L}^2(\Omega)$ (that is, a tilde symbol over an $\boldsymbol{L}^2$ function denotes the application of the $(I - \Pi^0_k)$ operator). Manipulating (50) and (8) and recalling (30) we have

$$\begin{aligned} &a_h(\boldsymbol{u}_h,\boldsymbol{\xi}_h) - a_h(\boldsymbol{u}_I,\boldsymbol{\xi}_h) = a(\boldsymbol{u},\boldsymbol{\xi}_h) - a_h(\boldsymbol{u}_I,\boldsymbol{\xi}_h) + (\boldsymbol{f}_h - \boldsymbol{f},\boldsymbol{\xi}_h)\\ &= \left(\int_\Omega \boldsymbol{\sigma}(\cdot,\boldsymbol{\epsilon}(\boldsymbol{u})) : \boldsymbol{\epsilon}(\boldsymbol{\xi}_h) - \int_\Omega \boldsymbol{\sigma}(\cdot,\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_I)) : \boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{\xi}_h)\right) - S(\widetilde{\boldsymbol{u}}_I,\widetilde{\boldsymbol{\xi}}_h) + (\boldsymbol{f}_h - \boldsymbol{f},\boldsymbol{\xi}_h)\\ &=: T_1 + T_2 + T_3\,. \end{aligned} \tag{65}$$

We next estimate each term on the right-hand side above. In the following C will denote a generic positive constant independent of h that may change at each occurrence, whereas the positive parameter θ adopted in (67) and (74) will be specified later.

- Estimate of T_1. Employing the definition of L^2-projection (20) we have

$$\begin{aligned} T_1 &= \int_\Omega \left(\boldsymbol{\sigma}(\cdot,\boldsymbol{\epsilon}(\boldsymbol{u})) - \boldsymbol{\Pi}^0_{k-1}\boldsymbol{\sigma}(\cdot,\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_I))\right) : \boldsymbol{\epsilon}(\boldsymbol{\xi}_h)\\ &= \int_\Omega \left((I - \boldsymbol{\Pi}^0_{k-1})\boldsymbol{\sigma}(\cdot,\boldsymbol{\epsilon}(\boldsymbol{u}))\right) : \boldsymbol{\epsilon}(\boldsymbol{\xi}_h) + \int_\Omega \left(\boldsymbol{\sigma}(\cdot,\boldsymbol{\epsilon}(\boldsymbol{u})) - \boldsymbol{\sigma}(\cdot,\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_I))\right) : \boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{\xi}_h)\\ &=: T_1^A + T_1^B\,. \end{aligned} \tag{66}$$

The term T_1^A can be bounded as follows

$$\begin{aligned} T_1^A &= \sum_{E\in\Omega_h}\int_E \left((I-\boldsymbol{\Pi}_{k-1}^{0,E})\boldsymbol{\sigma}(\cdot,\boldsymbol{\epsilon}(\boldsymbol{u}))\right) : \boldsymbol{\epsilon}(\boldsymbol{\xi}_h) \\ &\leq C\sum_{E\in\Omega_h} h_E^{k_2}|\boldsymbol{\sigma}(\cdot,\boldsymbol{\epsilon}(\boldsymbol{u}))|_{\mathbb{W}^{k_2,r'}(E)}|\boldsymbol{\xi}_h|_{\boldsymbol{W}^{1,r}(E)} \qquad ((r',r)\text{-Hölder ineq. \& Lemma 4}) \\ &\leq \frac{C}{\theta}h^{2k_2}|\boldsymbol{\sigma}(\cdot,\boldsymbol{\epsilon}(\boldsymbol{u}))|^2_{\mathbb{W}^{k_2,r'}(\Omega_h)} + \frac{\theta}{2}\|\boldsymbol{\xi}_h\|^2_{\boldsymbol{W}^{1,r}(\Omega)} \qquad ((r',r)\text{-Hölder ineq. \& Young ineq.}) \end{aligned} \tag{67}$$

Employing (62a) with $\delta = 0$ we obtain

$$\begin{aligned} T_1^B &= \sum_{E\in\Omega_h}\int_E |\boldsymbol{\sigma}(\cdot,\boldsymbol{\epsilon}(\boldsymbol{u})) - \boldsymbol{\sigma}(\cdot,\boldsymbol{\Pi}_{k-1}^{0,E}\boldsymbol{\epsilon}(\boldsymbol{u}_I))|\,|\boldsymbol{\Pi}_{k-1}^{0,E}\boldsymbol{\epsilon}(\boldsymbol{\xi}_h)| \\ &\leq \sum_{E\in\Omega_h}\int_E C\varphi'_{|\boldsymbol{\Pi}_{k-1}^{0,E}\boldsymbol{\epsilon}(\boldsymbol{u}_I)|}(|\boldsymbol{\Pi}_{k-1}^{0,E}\boldsymbol{\epsilon}(\boldsymbol{u}_I) - \boldsymbol{\epsilon}(\boldsymbol{u})|)\,|\boldsymbol{\Pi}_{k-1}^{0,E}\boldsymbol{\epsilon}(\boldsymbol{\xi}_h)|\,. \end{aligned}$$

Employing (60) we get that for every $\varepsilon > 0$ there exists a positive constant $C(\varepsilon)$ such that

$$T_1^B \leq \varepsilon\sum_{E\in\Omega_h}\int_E \varphi_{|\boldsymbol{\Pi}_{k-1}^{0,E}\boldsymbol{\epsilon}(\boldsymbol{u}_I)|}(|\boldsymbol{\Pi}_{k-1}^{0,E}\boldsymbol{\epsilon}(\boldsymbol{\xi}_h)|) + C(\varepsilon)\sum_{E\in\Omega_h}\int_E \varphi_{|\boldsymbol{\Pi}_{k-1}^{0,E}\boldsymbol{\epsilon}(\boldsymbol{u}_I)|}(|\boldsymbol{\Pi}_{k-1}^{0,E}\boldsymbol{\epsilon}(\boldsymbol{u}_I) - \boldsymbol{\epsilon}(\boldsymbol{u})|).$$

Using (62b) we obtain (with γ denoting the associated uniform hidden constant)

$$\begin{aligned} T_1^B &\leq \gamma\varepsilon(\boldsymbol{\sigma}(\cdot,\boldsymbol{\Pi}_{k-1}^0\boldsymbol{\epsilon}(\boldsymbol{u}_h)) - \boldsymbol{\sigma}(\cdot,\boldsymbol{\Pi}_{k-1}^0\boldsymbol{\epsilon}(\boldsymbol{u}_I)), \boldsymbol{\Pi}_{k-1}^0\boldsymbol{\epsilon}(\boldsymbol{\xi}_h)) \\ &\quad + C(\varepsilon)\sum_{E\in\Omega_h}\int_E (|\boldsymbol{\Pi}_{k-1}^{0,E}\boldsymbol{\epsilon}(\boldsymbol{u}_I)| + |\boldsymbol{\epsilon}(\boldsymbol{u})|)^{r-2}|\boldsymbol{\epsilon}(\boldsymbol{u}) - \boldsymbol{\Pi}_{k-1}^{0,E}\boldsymbol{\epsilon}(\boldsymbol{u}_I)|^2\,. \end{aligned}$$

Notice that the constant $C(\varepsilon)$ may depend on $\boldsymbol{\sigma}$, the degree k, the domain Ω but, given our mesh assumptions, it is independent of the particular mesh or mesh element within the family $\{\Omega_h\}_h$. Recalling that $r - 2 < 0$, and employing the fact that $|\boldsymbol{\epsilon}(\boldsymbol{u}) - \boldsymbol{\Pi}_{k-1}^{0,E}\boldsymbol{\epsilon}(\boldsymbol{u}_I)| \leq |\boldsymbol{\epsilon}(\boldsymbol{u})| + |\boldsymbol{\Pi}_{k-1}^{0,E}\boldsymbol{\epsilon}(\boldsymbol{u}_I)|$, from the last equation we get

$$T_1^B \leq \gamma\varepsilon(\boldsymbol{\sigma}(\cdot,\boldsymbol{\Pi}_{k-1}^0\boldsymbol{\epsilon}(\boldsymbol{u}_h)) - \boldsymbol{\sigma}(\cdot,\boldsymbol{\Pi}_{k-1}^0\boldsymbol{\epsilon}(\boldsymbol{u}_I)), \boldsymbol{\Pi}_{k-1}^0\boldsymbol{\epsilon}(\boldsymbol{\xi}_h)) + C(\varepsilon)\sum_{E\in\Omega_h}\|\boldsymbol{\epsilon}(\boldsymbol{u}) - \boldsymbol{\Pi}_{k-1}^{0,E}\boldsymbol{\epsilon}(\boldsymbol{u}_I)\|^r_{\mathbb{L}^r(E)}\,,$$

which, making use of Lemma 4, and taking $\varepsilon = \frac{1}{2\gamma}$ becomes

$$T_1^B \leq \frac{1}{2}(\boldsymbol{\sigma}(\cdot,\boldsymbol{\Pi}_{k-1}^0\boldsymbol{\epsilon}(\boldsymbol{u}_h)) - \boldsymbol{\sigma}(\cdot,\boldsymbol{\Pi}_{k-1}^0\boldsymbol{\epsilon}(\boldsymbol{u}_I)), \boldsymbol{\Pi}_{k-1}^0\boldsymbol{\epsilon}(\boldsymbol{\xi}_h)) + Ch^{k_1 r}|\boldsymbol{u}|^r_{\boldsymbol{W}^{k_1+1,r}(\Omega_h)}\,. \tag{68}$$

Combining (67) and (68) in (66) we infer

$$\begin{aligned} T_1 \leq\ & \frac{1}{2}(\boldsymbol{\sigma}(\cdot,\boldsymbol{\Pi}_{k-1}^0\boldsymbol{\epsilon}(\boldsymbol{u}_h)) - \boldsymbol{\sigma}(\cdot,\boldsymbol{\Pi}_{k-1}^0\boldsymbol{\epsilon}(\boldsymbol{u}_I)), \boldsymbol{\Pi}_{k-1}^0\boldsymbol{\epsilon}(\boldsymbol{\xi}_h)) + \\ & + Ch^{k_1 r}|\boldsymbol{u}|^r_{\boldsymbol{W}^{k_1+1,r}(\Omega_h)} + \frac{C}{\theta}h^{2k_2}|\boldsymbol{\sigma}(\cdot,\boldsymbol{\epsilon}(\boldsymbol{u}))|^2_{\mathbb{W}^{k_2,r'}(\Omega_h)} + \frac{\theta}{2}\|\boldsymbol{\xi}_h\|^2_{\boldsymbol{W}^{1,r}(\Omega)}\,. \end{aligned} \tag{69}$$

- Estimate of T_2. Recalling definitions (29) and (28) with $\delta = 0$ we have

$$T_2 = -\sum_{E\in\Omega_h} S^E(\widetilde{\boldsymbol{u}}_I, \widetilde{\boldsymbol{\xi}}_h) \leq \sum_{E\in\Omega_h} h_E^{2-r}|\boldsymbol{\chi}(\widetilde{\boldsymbol{u}}_I)|^{r-2}|\boldsymbol{\chi}(\widetilde{\boldsymbol{u}}_I)|\,|\boldsymbol{\chi}(\widetilde{\boldsymbol{\xi}}_h)|\,.$$

Employing (62a) in Lemma 21 to the scalar function $\sigma(\tau) = |\tau|^{r-2}\tau$ (hence $\delta = 0$ and $\eta = 0$) we have

$$T_2 \leq \sum_{E\in\Omega_h} C\varphi'_{|\boldsymbol{\chi}(\widetilde{\boldsymbol{u}}_I)|}(|\boldsymbol{\chi}(\widetilde{\boldsymbol{u}}_I)|)\,|\boldsymbol{\chi}(\widetilde{\boldsymbol{\xi}}_h)|\,.$$

Employing (60) we get that for every $\varepsilon > 0$ there exists a positive constant $C(\varepsilon)$ such that

$$T_2 \leq \varepsilon\sum_{E\in\Omega_h} h_E^{2-r}\varphi_{|\boldsymbol{\chi}(\widetilde{\boldsymbol{u}}_I)|}(|\boldsymbol{\chi}(\widetilde{\boldsymbol{\xi}}_h)|) + C(\varepsilon)\sum_{E\in\Omega_h} h_E^{2-r}\varphi_{|\boldsymbol{\chi}(\widetilde{\boldsymbol{u}}_I)|}(|\boldsymbol{\chi}(\widetilde{\boldsymbol{u}}_I)|)\,.$$

We now use (62b) and, denoting with γ the hidden constant, we infer

$$T_2 \leq \varepsilon\gamma \sum_{E\in\Omega_h} h_E^{2-r}(|\chi(\widetilde{\boldsymbol{u}}_I)| + |\chi(\widetilde{\boldsymbol{u}}_h)|)^{r-2}|\chi(\widetilde{\boldsymbol{\xi}}_h)|^2 + C(\varepsilon)\sum_{E\in\Omega_h} h_E^{2-r}|\chi(\widetilde{\boldsymbol{u}}_I)|^r =: T_2^A + T_B^2\,. \tag{70}$$

Employing (36) in Lemma 11 (with $\boldsymbol{u}_h = \widetilde{\boldsymbol{u}}_h$ and $\boldsymbol{w}_h = \widetilde{\boldsymbol{u}}_I$) and recalling definition (29) we obtain

$$T_2^A \leq \varepsilon C\gamma \sum_{E\in\Omega_h} (S^E(\widetilde{\boldsymbol{u}}_h, \widetilde{\boldsymbol{\xi}}_h) - S^E(\widetilde{\boldsymbol{u}}_I, \widetilde{\boldsymbol{\xi}}_h)) = \varepsilon C\gamma (S(\widetilde{\boldsymbol{u}}_h, \widetilde{\boldsymbol{\xi}}_h) - S(\widetilde{\boldsymbol{u}}_I, \widetilde{\boldsymbol{\xi}}_h))\,. \tag{71}$$

Employing definition (28), Lemma 8 (with $\boldsymbol{v}_h = \widetilde{\boldsymbol{u}}_I$), a triangular inequality and Lemma 4 we infer

$$\begin{aligned} T_2^B &\leq C(\varepsilon)\sum_{E\in\Omega_h} S^E(\widetilde{\boldsymbol{u}}_I, \widetilde{\boldsymbol{u}}_I) \leq C(\varepsilon)\sum_{E\in\Omega_h} |\widetilde{\boldsymbol{u}}_I|^r_{\boldsymbol{W}^{1,r}(E)} \\ &\leq C(\varepsilon)\sum_{E\in\Omega_h} \left(|\widetilde{\boldsymbol{u}} - \widetilde{\boldsymbol{u}}_I|^r_{\boldsymbol{W}^{1,r}(E)} + |\widetilde{\boldsymbol{u}}|^r_{\boldsymbol{W}^{1,r}(E)}\right) \leq C(\varepsilon) h^{k_1 r}|\boldsymbol{u}|^r_{\boldsymbol{W}^{k_1+1,r}(\Omega_h)}\,. \end{aligned} \tag{72}$$

Combining (71) and (72) in (70) and taking $\epsilon = \frac{1}{2C\gamma}$ we infer

$$T_2 \leq \frac{1}{2}(S(\widetilde{\boldsymbol{u}}_h, \widetilde{\boldsymbol{\xi}}_h) - S(\widetilde{\boldsymbol{u}}_I, \widetilde{\boldsymbol{\xi}}_h)) + Ch^{k_1 r}|\boldsymbol{u}|^r_{\boldsymbol{W}^{k_1+1,r}(\Omega_h)}\,. \tag{73}$$

- Estimate of T_3. We infer

$$\begin{aligned} T_3 &= \sum_{E\in\Omega_h}\int_E (\Pi_k^{0,E}\boldsymbol{f} - \boldsymbol{f})\cdot(\boldsymbol{\xi}_h - \Pi_0^{0,E}(\boldsymbol{\xi}_h)) && \text{(def. (31) \& def. (20))} \\ &\leq C\sum_{E\in\Omega_h} h_E^{k_3+2}|\boldsymbol{f}|_{\boldsymbol{W}^{k_3+1,r'}(E)}|\boldsymbol{\xi}_h|_{\boldsymbol{W}^{1,r}(E)} && \text{((r',r)-Hölder ineq. \& Lemma 4)} \\ &\leq \frac{C}{\theta}h^{2k_3+4}|\boldsymbol{f}|^2_{\boldsymbol{W}^{k_3+1,r'}(\Omega_h)} + \frac{\theta}{2}\|\boldsymbol{\xi}_h\|^2_{\boldsymbol{W}^{1,r}(\Omega)} && \text{((r',r)-Hölder ineq. \& Young ineq.)} \end{aligned} \tag{74}$$

Plugging the estimates in (69), (73) and (74) in (65) and recalling definition (30) we obtain

$$\begin{aligned} &\frac{1}{2}(a_h(\boldsymbol{u}_h, \boldsymbol{\xi}_h) - a_h(\boldsymbol{u}_I, \boldsymbol{\xi}_h)) \leq \\ &Ch^{k_1 r}|\boldsymbol{u}|^r_{\boldsymbol{W}^{k_1+1,r}(\Omega_h)} + \frac{C}{\theta}h^{2k_2}|\boldsymbol{\sigma}(\cdot,\boldsymbol{\epsilon}(\boldsymbol{u}))|^2_{\mathbb{W}^{k_2,r'}(\Omega_h)} + \frac{C}{\theta}h^{2k_3+4}|\boldsymbol{f}|^2_{\boldsymbol{W}^{k_3+1,r'}(\Omega_h)} + \theta\|\boldsymbol{\xi}_h\|^2_{\boldsymbol{W}^{1,r}(\Omega)}\,. \end{aligned} \tag{75}$$

We now derive a lower bound for the left-hand side of the equation above. Note that, since from bounds (13) and (57), estimate (63) clearly holds true if $\mathcal{N}(\boldsymbol{f},\boldsymbol{g}) = 0$, it is not restrictive to assume $\mathcal{N}(\boldsymbol{f},\boldsymbol{g}) > 0$. Employing Lemma 16, a triangular inequality, the stability bounds (13) and (57) and the interpolation bound in Lemma 4 we get

$$\begin{aligned} \frac{1}{2}(a_h(\boldsymbol{u}_h, \boldsymbol{\xi}_h) - a_h(\boldsymbol{u}_I, \boldsymbol{\xi}_h)) &\geq C\left(\|\boldsymbol{u}_h\|^r_{\boldsymbol{W}^{1,r}(\Omega)} + \|\boldsymbol{u}_I\|^r_{\boldsymbol{W}^{1,r}(\Omega)}\right)^{\frac{r-2}{r}}\|\boldsymbol{\xi}_h\|^2_{\boldsymbol{W}^{1,r}(\Omega)} \\ &\geq C\left(\|\boldsymbol{u}_h\|^r_{\boldsymbol{W}^{1,r}(\Omega)} + \|\boldsymbol{u}\|^r_{\boldsymbol{W}^{1,r}(\Omega)} + \|\boldsymbol{u} - \boldsymbol{u}_I\|^r_{\boldsymbol{W}^{1,r}(\Omega)}\right)^{\frac{r-2}{r}}\|\boldsymbol{\xi}_h\|^2_{\boldsymbol{W}^{1,r}(\Omega)} \\ &\geq C\left(\mathcal{N}(\boldsymbol{f},\boldsymbol{g})^{\frac{r}{r-1}} + h^{k_1 r}|\boldsymbol{u}|^r_{\boldsymbol{W}^{k_1+1,r}(\Omega)}\right)^{\frac{r-2}{r}}\|\boldsymbol{\xi}_h\|^2_{\boldsymbol{W}^{1,r}(\Omega)} \\ &\geq C\mathcal{N}(\boldsymbol{f},\boldsymbol{g})^{\frac{r-2}{r-1}}\|\boldsymbol{\xi}_h\|^2_{\boldsymbol{W}^{1,r}(\Omega)} \end{aligned} \tag{76}$$

where in the last inequality we assume that h is sufficiently small. Combining (75) and (76) and taking $\theta = \frac{1}{2}C\mathcal{N}(\boldsymbol{f},\boldsymbol{g})^{\frac{r-2}{r-1}}$ we obtain:

$$\|\boldsymbol{\xi}_h\|^2_{\boldsymbol{W}^{1,r}(\Omega)} \lesssim C(\mathcal{N}(\boldsymbol{f},\boldsymbol{g}))\left(h^{k_1 r}|\boldsymbol{u}|^r_{\boldsymbol{W}^{k_1+1,r}(\Omega_h)} + h^{2k_2}|\boldsymbol{\sigma}(\cdot,\boldsymbol{\epsilon}(\boldsymbol{u}))|^2_{\mathbb{W}^{k_2,r'}(\Omega_h)} + h^{2k_3+4}|\boldsymbol{f}|^2_{\boldsymbol{W}^{k_3+1,r'}(\Omega_h)}\right)\,. \tag{77}$$

The proof follows by the previous bound, a triangular inequality and Lemma 4. □

Proposition 23. *Let $(\boldsymbol{u}, p)$ be the solution of problem (6) and $(\boldsymbol{u}_h, p_h)$ the solution of problem (49). Assume that $\boldsymbol{u} \in \boldsymbol{W}^{k_1+1,r}(\Omega_h)$, $\boldsymbol{\sigma}(\cdot, \boldsymbol{\epsilon}(\boldsymbol{u})) \in \mathbb{W}^{k_2,r'}(\Omega_h)$, $\boldsymbol{f} \in \boldsymbol{W}^{k_3+1,r'}(\Omega_h)$ $p \in W^{k_4,r'}(\Omega_h)$, for some k_1, k_2, k_3, $k_4 \leq k$. Let the mesh regularity in Assumption 2 hold. Then, we have*

$$\|p - p_h\|_{L^{r'}(\Omega_h)} \lesssim \|\boldsymbol{u} - \boldsymbol{u}_h\|^{2/r'}_{\boldsymbol{W}^{1,r}(\Omega_h)} + h^{k_1(r-1)} R_1^{r-1} + h^{k_2} R_2 + h^{k_3+2} R_3 + h^{k_4} R_4 \,,$$

where the regularity term R_1, R_2, R_3 are defined in (64) and $R_4 := |p|_{W^{k_4,r'}(\Omega_h)}$.

Proof. Let $p_I := \Pi^0_{k-1} p \in \mathbb{P}_{k-1}(\Omega_h)$ and let $\rho_h := p_h - p_I$. Employing (6) and (49), recalling the definition of the form $b(\cdot,\cdot)$ in (7) and combining item (ii) in (23) with the definition of L^2-projection, for all $\boldsymbol{w}_h \in \boldsymbol{U}_h$ we get

$$\begin{aligned} b(\boldsymbol{w}_h, \rho_h) &= a_h(\boldsymbol{u}_h, \boldsymbol{w}_h) - (\boldsymbol{f}_h, \boldsymbol{w}_h) - a(\boldsymbol{u}, \boldsymbol{w}_h) + (\boldsymbol{f}, \boldsymbol{w}_h) + b(\boldsymbol{w}_h, p - p_I) \\ &= a_h(\boldsymbol{u}_h, \boldsymbol{w}_h) - a(\boldsymbol{u}, \boldsymbol{w}_h) + (\boldsymbol{f} - \boldsymbol{f}_h, \boldsymbol{w}_h) \\ &= \int_\Omega \Big((\boldsymbol{\Pi}^0_{k-1} - I)\boldsymbol{\sigma}(\cdot, \boldsymbol{\epsilon}(\boldsymbol{u})) \Big) : \boldsymbol{\epsilon}(\boldsymbol{w}_h) + (\boldsymbol{f} - \boldsymbol{f}_h, \boldsymbol{w}_h) + \\ &+ \int_\Omega \Big(\boldsymbol{\sigma}(\cdot, \boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h)) - \boldsymbol{\sigma}(\cdot, \boldsymbol{\epsilon}(\boldsymbol{u})) \Big) : \boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{w}_h) + S(\widetilde{\boldsymbol{u}_h}, \widetilde{\boldsymbol{w}_h}) \\ &=: T_1 + T_2 + T_3 + T_4 \,. \end{aligned} \tag{78}$$

We estimate separately each term in (78). Employing the Hölder inequality with exponents (r', r) and polynomial approximation properties we infer

$$\begin{aligned} T_1 &\lesssim \|(\boldsymbol{\Pi}^0_{k-1} - I)\boldsymbol{\sigma}(\cdot, \boldsymbol{\epsilon}(\boldsymbol{u}))\|_{\mathbb{L}^{r'}(\Omega_h)} \|\boldsymbol{\epsilon}(\boldsymbol{w}_h)\|_{\mathbb{L}^r(\Omega_h)} \lesssim h^{k_2} |\boldsymbol{\sigma}(\cdot, \boldsymbol{\epsilon}(\boldsymbol{u}))|_{\mathbb{W}^{k_2,r'}(\Omega_h)} \|\boldsymbol{w}_h\|_{\boldsymbol{W}^{1,r}(\Omega_h)} \,, \\ T_2 &= (\boldsymbol{f} - \boldsymbol{f}_h, \boldsymbol{w}_h - \Pi^0_0 \boldsymbol{w}_h) \lesssim h^{k_3+2} |\mathbf{f}|_{\boldsymbol{W}^{k_3+1,r'}(\Omega_h)} \|\boldsymbol{w}_h\|_{W^{1,r}(\Omega_h)} \,. \end{aligned} \tag{79}$$

Employing the Hölder inequality with exponents (r', r), the $W^{1,r}$-stability of $\boldsymbol{\Pi}^0_{k-1}$ and Lemma 25 in the Appendix, we have

$$\begin{aligned} T_3^{r'} &\lesssim \|\boldsymbol{\sigma}(\cdot, \boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h)) - \boldsymbol{\sigma}(\cdot, \boldsymbol{\epsilon}(\boldsymbol{u}))\|^{r'}_{\mathbb{L}^{r'}(\Omega_h)} \|\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{w}_h)\|^{r'}_{\mathbb{L}^r(\Omega_h)} \\ &\lesssim \Big((\boldsymbol{\sigma}(\cdot, \boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h)) - \boldsymbol{\sigma}(\cdot, \boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_I)), \boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{\xi}_h)) + h^{k_1 r} |\boldsymbol{u}|^r_{\boldsymbol{W}^{k_1+1,r}(\Omega_h)} \Big) \|\boldsymbol{w}_h\|^{r'}_{\boldsymbol{W}^{1,r}(\Omega_h)} \end{aligned} \tag{80}$$

where $\boldsymbol{\xi}_h = \boldsymbol{u}_h - \boldsymbol{u}_I$ (cf. proof of Proposition 22). Employing Lemma 14, Corollary 10, the $W^{1,r}$-stability of Π^0_{k-1} and finally Lemma 26 in the Appendix we obtain

$$\begin{aligned} T_4^{r'} &\lesssim \|\widetilde{\boldsymbol{u}}_h\|^r_{\boldsymbol{W}^{1,r}(\Omega_h)} \|\widetilde{\boldsymbol{w}}_h\|^{r'}_{\boldsymbol{W}^{1,r}(\Omega_h)} \lesssim S(\widetilde{\boldsymbol{u}}_h, \widetilde{\boldsymbol{u}}_h) \|\boldsymbol{w}_h\|^{r'}_{\boldsymbol{W}^{1,r}(\Omega_h)} \\ &\lesssim \Big(S(\widetilde{\boldsymbol{u}}_h, \widetilde{\boldsymbol{\xi}}_h) - S(\widetilde{\boldsymbol{u}}_I, \widetilde{\boldsymbol{\xi}}_h) + h^{k_1 r} |\boldsymbol{u}|^r_{\boldsymbol{W}^{k_1+1,r}(\Omega_h)} \Big) \|\boldsymbol{w}_h\|^{r'}_{\boldsymbol{W}^{1,r}(\Omega_h)} \,. \end{aligned} \tag{81}$$

Therefore recalling the definition of $a_h(\cdot,\cdot)$ in (30), from (80) and (81), and recalling (75) and (77), we have

$$\begin{aligned} T_3^{r'} + T_4^{r'} &\lesssim \Big(a_h(\boldsymbol{u}_h, \boldsymbol{\xi}_h) - a_h(\boldsymbol{u}_I, \boldsymbol{\xi}_h) + h^{k_1 r} |\boldsymbol{u}|^r_{\boldsymbol{W}^{k_1+1,r}(\Omega_h)} \Big) \|\boldsymbol{w}_h\|^{r'}_{\boldsymbol{W}^{1,r}(\Omega_h)} \\ &\lesssim \Big(\|\boldsymbol{\xi}_h\|^2_{\boldsymbol{W}^{1,r}(\Omega_h)} + h^{k_1 r} |\boldsymbol{u}|^r_{\boldsymbol{W}^{k_1+1,r}(\Omega_h)} \Big) \|\boldsymbol{w}_h\|^{r'}_{\boldsymbol{W}^{1,r}(\Omega_h)} \\ &\lesssim \Big(\|\boldsymbol{u} - \boldsymbol{u}_h\|^2_{\boldsymbol{W}^{1,r}(\Omega_h)} + h^{2k_1} |\boldsymbol{u}|^2_{\boldsymbol{W}^{k_1+1,r}(\Omega_h)} + h^{k_1 r} |\boldsymbol{u}|^r_{\boldsymbol{W}^{k_1+1,r}(\Omega_h)} \Big) \|\boldsymbol{w}_h\|^{r'}_{\boldsymbol{W}^{1,r}(\Omega_h)} \,. \end{aligned} \tag{82}$$

Collecting (79) and (82) in (78) we obtain

$$b(\boldsymbol{w}_h, \rho_h) \lesssim \Big(\|\boldsymbol{u} - \boldsymbol{u}_h\|^{2/r'}_{\boldsymbol{W}^{1,r}(\Omega_h)} + h^{k_1(r-1)} R_1^{r-1} + h^{k_2} R_2 + h^{k_3+2} R_3 \Big) \|\boldsymbol{w}_h\|_{W^{1,r}(\Omega_h)} \quad \forall \boldsymbol{w}_h \in \boldsymbol{U}_h.$$

Employing the discrete inf-sup (Lemma 15), from the equation above we have

$$\begin{aligned} \|\rho_h\|_{L^{r'}(\Omega_h)} &\leq \frac{1}{\bar{\beta}(r)} \sup_{\boldsymbol{w}_h \in \boldsymbol{U}_h} \frac{b(\boldsymbol{w}_h, \rho_h)}{\|\boldsymbol{w}_h\|_{W^{1,r}(\Omega_h)}} \\ &\lesssim \frac{1}{\bar{\beta}(r)} \Big(\|\boldsymbol{u} - \boldsymbol{u}_h\|^{2/r'}_{\boldsymbol{W}^{1,r}(\Omega_h)} + h^{k_1(r-1)} R_1^{r-1} + h^{k_2} R_2 + h^{k_3+2} R_3 \Big) \,. \end{aligned}$$

Now the thesis follows from triangular inequality and standard polynomial approximation property. □

Remark 24. (Rates of convergence) Notice that, assuming in Proposition 22 and Proposition 23 maximum regularity for the solution and for the data (that is $k_i = k$ for $i = 1, \dots, 4$) we recover the following asymptotic behaviour

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{W}^{1,r}(\Omega_h)} \lesssim h^{kr/2}, \qquad \|p - p_h\|_{L^{r'}(\Omega_h)} \lesssim h^{k(r-1)}$$

that are in agreement with the classical results presented, e.g., in [9] in the context of finite elements.

5 Numerical results

In this section, we present two sets of numerical experiments that corroborate our theoretical findings and test the practical performances of the proposed scheme (49) in various scenarios. In particular, despite not covered by our theory, we will also explore the case $\delta \neq 0$ in (2).

5.1 Fixed-point iteration

We first describe the linearization strategy based on a fixed-point iteration adopted to solve the non-linear Stokes equation. We consider the Carreau–Yasuda model (2) and for any $r \in (1, 2]$ we introduce the following notation:

$$\begin{aligned} \boldsymbol{\sigma}_r(\boldsymbol{x}, \boldsymbol{\epsilon}(z); \boldsymbol{\epsilon}(\boldsymbol{v})) &:= \mu(\boldsymbol{x})(\delta^\alpha + |\boldsymbol{\epsilon}(z)|^\alpha)^{\frac{r-2}{\alpha}} \boldsymbol{\epsilon}(\boldsymbol{v}) && \text{for all } z, \boldsymbol{v} \in \boldsymbol{U}, \\ a_r(z; \boldsymbol{v}, \boldsymbol{w}) &:= \int_\Omega \boldsymbol{\sigma}_r(\cdot, \boldsymbol{\epsilon}(z); \boldsymbol{\epsilon}(\boldsymbol{v})) : \boldsymbol{\epsilon}(\boldsymbol{w}) && \text{for all } z, \boldsymbol{v}, \boldsymbol{w} \in \boldsymbol{U}, \end{aligned}$$

furthermore $a_{h,r}(\cdot;\cdot,\cdot)$ will denote the VEM counterpart of $a_r(\cdot;\cdot,\cdot)$ obtained generalizing in the linear setting the construction in Subsection 3.3. Consider the discrete Problem (49) associated with the Carreau–Yasuda model with parameter $r \in (1, 2]$. Then, in the light of the notation introduced above, Problem (49) can be formulated as follows: find $(\boldsymbol{u}_h, p_h) \in \boldsymbol{U}_h \times P_h$ such that

$$\begin{aligned} a_{r,h}(\boldsymbol{u}_h; \boldsymbol{u}_h, \boldsymbol{v}_h) + b(\boldsymbol{v}_h, p_h) &= \int_\Omega \boldsymbol{f}_h \cdot \boldsymbol{v}_h + \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{\gamma}(\boldsymbol{v}_h) && \forall \boldsymbol{v}_h \in \boldsymbol{U}_h, \\ -b(\boldsymbol{u}_h, q_h) &= 0 && \forall q_h \in P_h . \end{aligned} \tag{83}$$

To solve the non-linear equation (83) we adopt the following strategy:

`STEP 1.` Let $(\boldsymbol{u}_h^{\mathrm{S}}, p_h^{\mathrm{S}})$ be the solution of the linear Stokes equation associated with (83) (corresponding to $r = 2$).
Let $\overline{r} := \frac{r+2}{2}$. Starting from $(\boldsymbol{u}_h^0, p_h^0) = (\boldsymbol{u}_h^{\mathrm{S}}, p_h^{\mathrm{S}})$, for $n \geq 0$, until convergence solve

$$\begin{aligned} a_{\overline{r},h}(\boldsymbol{u}_h^n; \boldsymbol{u}_h^{n+1}, \boldsymbol{v}_h) + b(\boldsymbol{v}_h, p_h^{n+1}) &= \int_\Omega \boldsymbol{f}_h \cdot \boldsymbol{v}_h + \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{\gamma}(\boldsymbol{v}_h) && \forall \boldsymbol{v}_h \in \boldsymbol{U}_h, \\ -b(\boldsymbol{u}_h^{n+1}, q_h) &= 0 && \forall q_h \in P_h . \end{aligned}$$

`STEP 2.` Let $(\boldsymbol{u}_h^{\overline{r}}, p_h^{\overline{r}})$ be the solution obtained in `STEP 1`.
Starting from $(\boldsymbol{u}_h^0, p_h^0) = (\boldsymbol{u}_h^{\overline{r}}, p_h^{\overline{r}})$, for $n \geq 0$, until convergence solve

$$\begin{aligned} a_{r,h}(\boldsymbol{u}_h^n; \boldsymbol{u}_h^{n+1}, \boldsymbol{v}_h) + b(\boldsymbol{v}_h, p_h^{n+1}) &= \int_\Omega \boldsymbol{f}_h \cdot \boldsymbol{v}_h + \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{\gamma}(\boldsymbol{v}_h) && \forall \boldsymbol{v}_h \in \boldsymbol{U}_h, \\ -b(\boldsymbol{u}_h^{n+1}, q_h) &= 0 && \forall q_h \in P_h . \end{aligned}$$

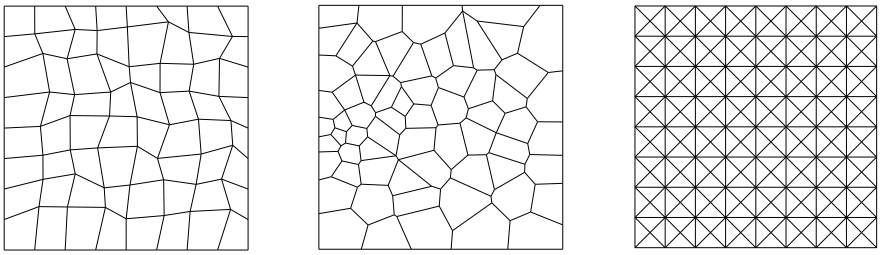


Figure 1: Example of the adopted polygonal meshes.

5.2 Error quantities

To compute the VEM error between the exact solution $(u_{\text{ex}}, p_{\text{ex}})$ and the VEM solution $(\boldsymbol{u}_h, p_h)$ we consider the computable error quantities

$$\begin{aligned}
\texttt{err}(\boldsymbol{u}_h, W^{1,r}) &:= \frac{\|\boldsymbol{\nabla}\boldsymbol{u}_{\text{ex}} - \boldsymbol{\Pi}^0_{k-1}\boldsymbol{\nabla}\boldsymbol{u}_h\|_{L^r(\Omega)}}{\|\boldsymbol{\nabla}\boldsymbol{u}_{\text{ex}}\|_{\mathbb{L}^r(\Omega)}}\,, \\
\texttt{err}(p_h, L^{r'}) &:= \frac{\|p_{\text{ex}} - p_h\|_{L^{r'}(\Omega)}}{\|p_{\text{ex}}\|_{L^{r'}(\Omega)}} \\
\texttt{err}(\boldsymbol{\sigma}, L^{r'}) &:= \frac{\|\boldsymbol{\sigma}(\cdot, \boldsymbol{\epsilon}(\boldsymbol{u}_{\text{ex}})) - \boldsymbol{\sigma}(\cdot, \boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h))\|_{L^{r'}(\Omega)}}{\|\boldsymbol{\sigma}(\cdot, \boldsymbol{\epsilon}(\boldsymbol{u}_{\text{ex}}))\|_{\mathbb{L}^{r'}(\Omega)}}\,.
\end{aligned}$$

In particular, we will compare the observed experimental convergence rates with the ones expected from the theory. In this last respect, we refer, for the case $\delta = 0$, to Propositions 22 and 23 (see also Remark 24), while we expect, for the case $\delta \neq 0$, optimal convergence rates (cf. [9, Theorem 4.1]), i.e.

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{W}^{1,r}(\Omega_h)} \lesssim h^k\,, \qquad \|p - p_h\|_{L^{r'}(\Omega_h)} \lesssim h^k.$$

In the forthcoming tests, we consider $k = 2$.

Test 1. Performance w.r.t. $\mathtt{r}$ and δ. In this test we examine the performance and the convergence properties of the proposed scheme (49) in the light of Proposition 22 and Proposition 23.
The aim of this test is to check the actual performance of the virtual element method assuming a Carreau–Yasuda model for different values of r and δ in (2). For simplicity we take $\mu = 1$ and $\alpha = 2$. We consider Problem (1) with full Dirichlet boundary conditions (i.e. $\Gamma_D = \partial\Omega$) on the unit square $\Omega = (0,1)^2$. The load terms $\boldsymbol{f}$ (depending on r and δ in (2)) and the Dirichlet boundary conditions are chosen according to the analytical solution

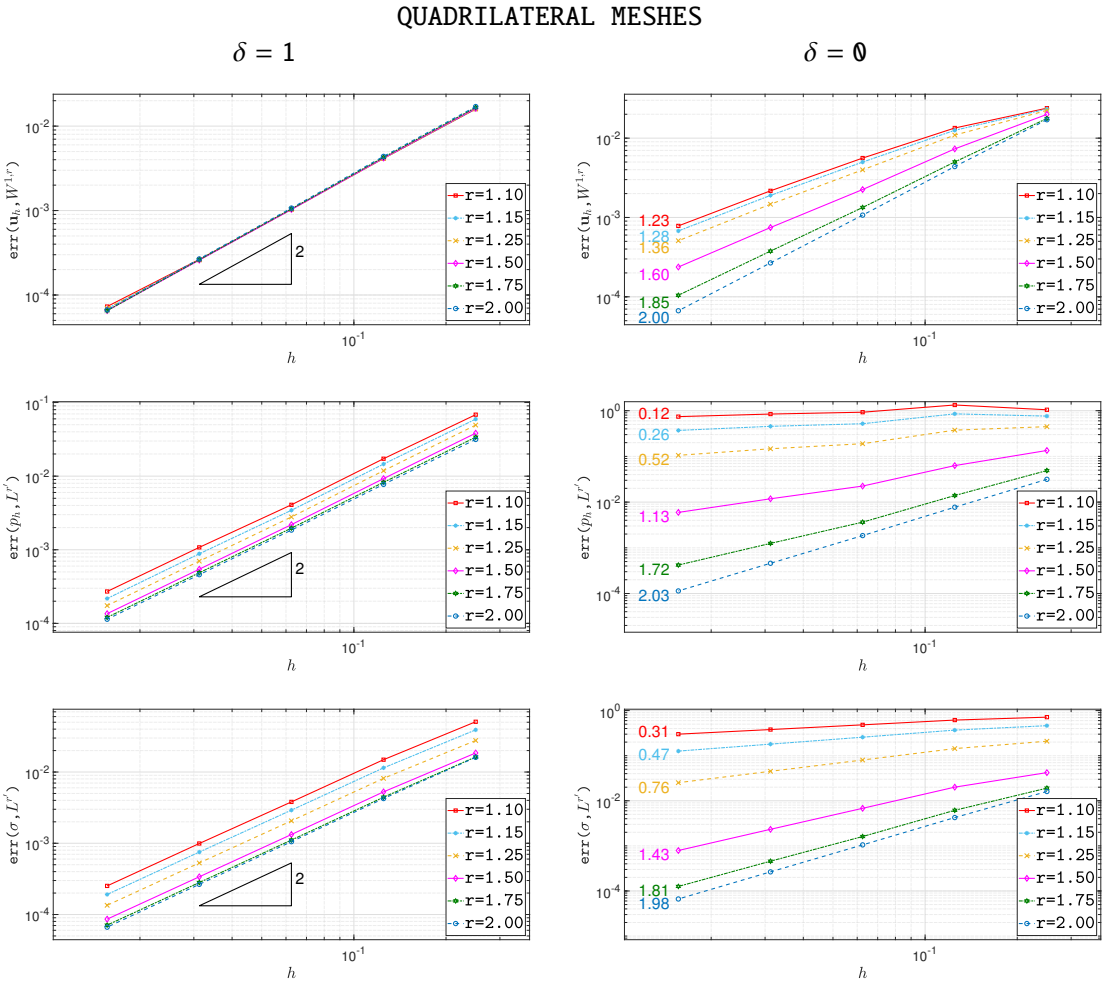
$$\boldsymbol{u}_{\text{ex}}(x_1, x_2) = \begin{bmatrix} \sin\left(\frac{\pi}{2}x_1\right)\cos\left(\frac{\pi}{2}x_2\right) \\ -\cos\left(\frac{\pi}{2}x_1\right)\sin\left(\frac{\pi}{2}x_2\right) \end{bmatrix}\,, \qquad p_{\text{ex}}(x_1, x_2) = -\sin\left(\frac{\pi}{2}x_1\right)\sin\left(\frac{\pi}{2}x_2\right) + \frac{4}{\pi^2}\,.$$

The domain Ω is partitioned with the following sequences of polygonal meshes: `QUADRILATERAL` distorted meshes and `RANDOM` Voronoi meshes (see Fig. 1). For the generation of the Voronoi meshes we used the code `Polymesher` [68]. For each family of meshes, we take the sequence with diameter `h = 1/4, 1/8, 1/16, 1/32, 1/64`. To highlight the behavior of the method for different situations we consider the following cases

$$\mathtt{r} = \mathtt{1.10},\ \mathtt{1.15},\ \mathtt{1.25},\ \mathtt{1.50},\ \mathtt{1.75},\ \mathtt{2.00}, \qquad \delta = \mathtt{1},\ \mathtt{0}\,.$$

In Fig. 1 and Fig. 2 we display the errors $\texttt{err}(\boldsymbol{u}_h, W^{1,r})$ (first row), $\texttt{err}(p_h, L^{r'})$ (second row), $\texttt{err}(\boldsymbol{\sigma}, L^{r'})$ (third row), introduced in Subsection 5.2 for the sequences of aforementioned meshes. We notice that the theoretical predictions are confirmed: for $\delta = 1$ we recover the optimal rate of convergence, i.e. order 2, whereas for $\delta = 0$, we report on the plot the averaged order of convergence that is in agreement with the theory.

Furthermore, we display the number of iterations for the fixed-point procedure described in Subsection 5.1 where N_1|N_2 denotes N_1 iterations for STEP 1 and N_2 iterations for STEP 2 are needed. We observe, as expected, that smaller values of r correspond to larger numbers of iterations, especially for $\delta = 0$.



1/h	r 1.10	1.15	1.25	1.50	1.75
4	5|6	5|6	4|5	4|5	3|4
8	5|6	4|6	4|5	4|5	3|4
16	5|6	4|6	4|5	4|5	3|4
32	4|6	4|6	4|5	4|5	3|4
64	4|6	4|6	4|5	4|5	3|4

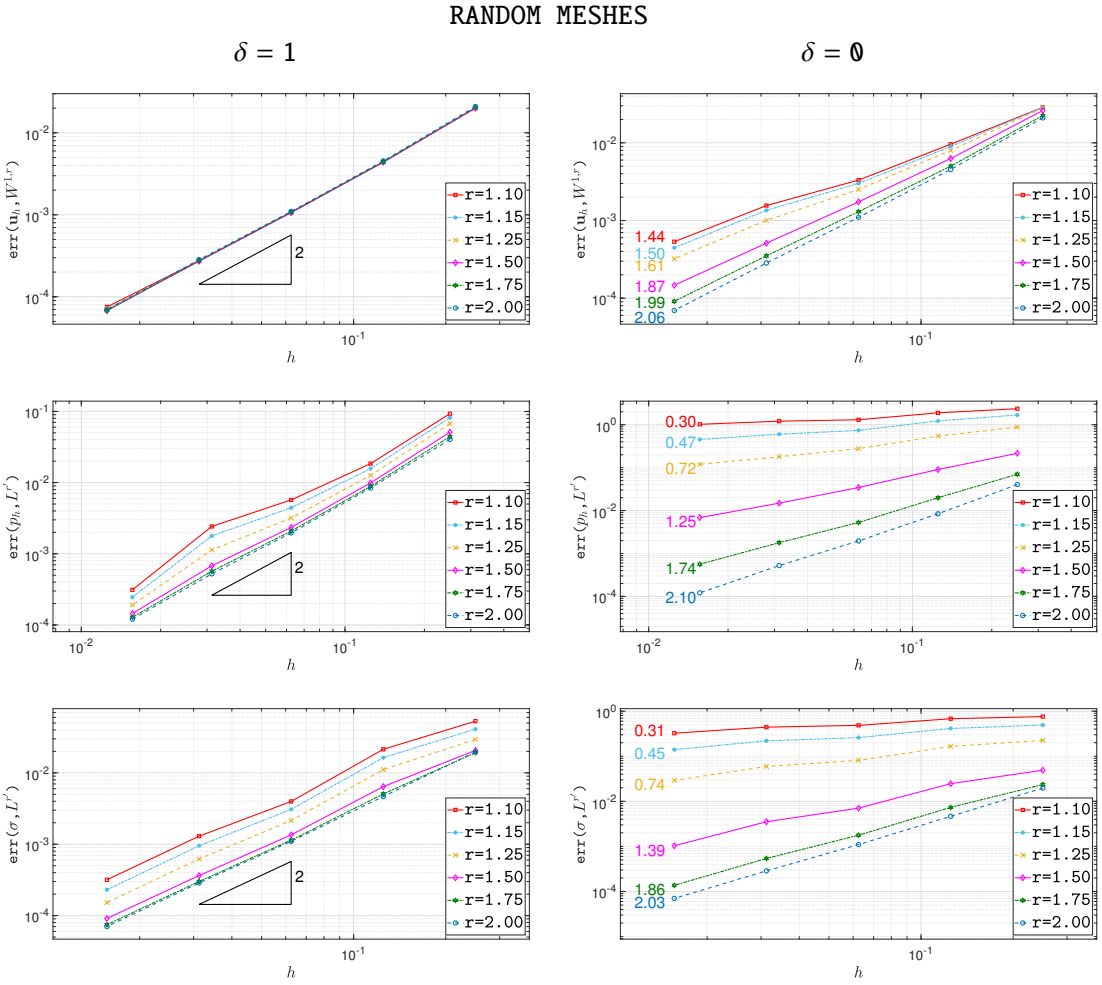
1/h	r 1.10	1.15	1.25	1.50	1.75
4	7|29	7|24	7|16	5|7	4|5
8	7|31	7|23	6|14	5|7	3|4
16	7|28	7|20	6|12	5|6	3|4
32	7|24	7|17	6|10	4|6	3|4
64	7|19	7|14	6|09	4|6	3|4

Table 1: Test 1. Convergence histories of the VEM errors (cf. beginning of Subsection 5.2) and the number of iterations of the fixed-point procedure for the mesh family QUADRILATERAL.

Test 2. Singular solution. The scope of the present test is to show the performance of the method in case of low regular solutions. We investigate the behavior of the proposed method for the test proposed in [26, Section 7].

We consider Problem (1) on the square domain $\Omega = (-1, 1)^2$ where $\boldsymbol{f}$ and the Dirichlet boundary conditions (posed on the whole $\partial\Omega$) are chosen in accordance with the analytical solution

$$\boldsymbol{u}_{\mathrm{ex}}(x_1, x_2) = |\boldsymbol{x}|^{0.01} \begin{bmatrix} x_2 \\ -x_1 \end{bmatrix}, \qquad p_{\mathrm{ex}}(x_1, x_2) = -|\boldsymbol{x}|^{\gamma} + c_{\gamma}\,.$$



1/h	r 1.10	1.15	1.25	1.50	1.75
4	5|6	5|6	4|6	4|5	3|4
8	5|6	4|6	4|6	4|5	3|4
16	5|6	4|6	4|5	4|5	3|4
32	4|6	4|6	4|5	3|4	3|4
64	4|6	4|6	4|5	3|4	3|4

1/h	r 1.10	1.15	1.25	1.50	1.75
4	7|22	7|17	7|12	5|7	4|5
8	7|25	7|18	6|12	5|6	4|4
16	7|20	7|15	6|10	5|6	4|4
32	7|17	7|13	6|09	5|6	4|4
64	7|14	7|11	6|07	5|6	4|4

Table 2: Test 1. Convergence histories of the VEM errors (cf. beginning of Subsection 5.2) and number of iterations of the fixed-point procedure for the mesh family RANDOM.

where $\gamma = \frac{2}{r} - 1 + 0.01$ and c_γ is s.t. $p_{\rm ex}$ is zero averaged. Notice that for all $r \in (1, 2]$

$$\boldsymbol{u}_{\rm ex} \in \boldsymbol{W}^{2/r+1,r}(\Omega)\,, \quad \boldsymbol{\sigma}(\cdot, \boldsymbol{\epsilon}(\boldsymbol{u}_{\rm ex})) \in \mathbb{W}^{2/r',r'}(\Omega)\,, \quad \boldsymbol{f} \in \boldsymbol{W}^{2/r'-1,r'}(\Omega)\,, \quad p_{\rm ex} \in W^{1,r'}(\Omega)\,,$$

therefore, with the notation of Proposition 22 and Proposition 23:

$$k_1 = \frac{2}{r}\,, \quad k_2 = \frac{2}{r'}\,, \quad k_3 = \frac{2}{r'} - 2\,, \quad k_4 = 1\,.$$

We analyse the Carreau–Yasuda model (2) with $\mu = 1$, $\alpha = 1$ and

$$\texttt{r} = \texttt{1.25, 1.33, 1.50, 1.67, 1.75, 2.00, .}$$

We consider $\delta = \texttt{0}$ (differently from [26, Section 7] where $\delta = \texttt{1e-04}$). The domain Ω is partitioned with a sequence of TRIANGULAR meshes with diameter h =1/2, 1/4, 1/8, 1/16 (see Fig.1). In

Tab. 3 we show the errors $\texttt{err}(\boldsymbol{u}_h, W^{1,r})$ for the proposed values of r. Notice that, accordingly with Proposition 22 the expected rate of convergence is $2/r'$, nevertheless we numerically observe linear convergence. This means that, at least in the pre-asymptotic regime, the error velocity errors is dominated by the term $h^{k_1 r/2} = h$ (cf. equation (63)). In Tab. 4 and Tab. 5 we exhibit the errors $\texttt{err}(p_h, L^{r'})$, $\texttt{err}(\boldsymbol{\sigma}, L^{r'})$, respectively. Notice that the pressure errors are in accordance with Proposition 23 and the previous numerical evidence, that is $\|p - p_h\|_{L^{r'}(\Omega)} \lesssim \|\boldsymbol{u} - \boldsymbol{u}_h\|^{2/r'}_{W^{1,r}(\Omega_h)} + h^{2/r'} + h \lesssim h^{2/r'}$.

	r					
1/h	1.25	1.33	1.50	1.67	1.75	2.00
2	7.5101e-04	7.7297e-04	8.2290e-04	8.7303e-04	8.9993e-04	9.9850e-04
4	2.8701e-04	3.0353e-04	3.4469e-04	3.8999e-04	4.1395e-04	4.9925e-04
8	1.1626e-04	1.2176e-04	1.4217e-04	1.7106e-04	1.8778e-04	2.4837e-04
16	7.4789e-05	5.9046e-05	6.1538e-05	7.5020e-05	8.4862e-05	1.2339e-04
a.c.r.	1.1093e+00	1.2368e+00	1.2470e+00	1.1802e+00	1.1355e+00	1.0055e+00
$\frac{2}{r'}$	0.40	0.50	0.66	0.80	0.86	1.00

Table 3: Test 2. Errors $\texttt{err}(\boldsymbol{u}_h, W^{1,r})$.

	r					
1/h	1.25	1.33	1.50	1.67	1.75	2.00
2	1.2116e-01	1.0603e-01	9.6989e-02	9.9739e-02	1.0303e-01	1.7563e-01
4	8.7515e-02	6.7873e-02	5.2335e-02	5.0615e-02	5.1764e-02	8.7824e-02
8	6.4955e-02	4.5430e-02	2.9065e-02	2.5820e-02	2.6045e-02	4.3701e-02
16	4.8749e-02	3.1194e-02	1.6625e-02	1.3260e-02	1.3135e-02	2.1711e-02
a.c.r.	4.3785e-01	5.8840e-01	8.4813e-01	9.7035e-01	9.9055e-01	1.0053e+00
$\frac{2}{r'}$	0.40	0.50	0.66	0.80	0.86	1.00

Table 4: Test 2. Errors $\texttt{err}(p_h, L^{r'})$.

	r					
1/h	1.25	1.33	1.50	1.67	1.75	2.00
2	1.4319e-01	1.4386e-01	1.4352e-01	1.4369e-01	1.4400e-01	1.4592e-01
4	1.0815e-01	1.0136e-01	9.0112e-02	8.2299e-02	7.9320e-02	7.3008e-02
8	8.1752e-02	7.1451e-02	5.6597e-02	4.7136e-02	4.3642e-02	3.6328e-02
16	6.1782e-02	5.0421e-02	3.5648e-02	2.7099e-02	2.4086e-02	1.8048e-02
a.c.r.	4.0425e-01	5.0419e-01	6.6980e-01	8.0220e-01	8.5995e-01	1.0051e+00
$\frac{2}{r'}$	0.40	0.50	0.66	0.80	0.86	1.00

Table 5: Test 2. Errors $\texttt{err}(\boldsymbol{\sigma}, L^{r'})$.

Acknowledgements

This research has been partially funded by the European Union (ERC, NEMESIS, project number 101115663). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. PFA, LBdV and MV have been partially funded by PRIN2020 n. 20204LN5N5 *"Advanced polyhedral discretisations of heterogeneous PDEs for multiphysics problems"* research grant, funded by the

Italian Ministry of Universities and Research (MUR). GV has been partially funded by PRIN2022 n. 2022MBY5JM*"FREYA - Fault REactivation: a hYbrid numerical Approach"* research grant, and PRIN2022PNRR n. P2022M7JZW*"SAFER MESH - Sustainable mAnagement oF watEr Resources: ModEls and numerical MetHods"* research grant, funded by the Italian Ministry of Universities and Research (MUR). PFA is partially supported by ICSC – Centro Nazionale di Ricerca in High Performance Computing, Big Data, and Quantum Computing funded by European Union – NextGenerationEU. The present research is part of the activities of "Dipartimento di Eccelllenza 2023-2027". The authors are members of INdAM-GNCS.

References

[1] D. Adak, D. Mora, S. Natarajan, and A. Silgado. A virtual element discretization for the time dependent Navier-Stokes equations in stream-function formulation. *ESAIM Math. Model. Numer. Anal.*, 55(5):2535–2566, 2021.

[2] D. Adak, D. Mora, and A. Silgado. The Morley-type virtual element method for the Navier-Stokes equations in stream-function form. *Comput. Methods Appl. Mech. Engrg.*, 419:Paper No. 116573, 28, 2024.

[3] P. F. Antonietti, C. Facciolà, P. Houston, I. Mazzieri, G. Pennesi, and M. Verani. High-order discontinuous Galerkin methods on polyhedral grids for geophysical applications: seismic wave propagation and fractured reservoir simulations. In *Polyhedral methods in geosciences*, volume 27 of *SEMA SIMAI Springer Ser.*, pages 159–225. Springer, Cham, [2021] ©2021.

[4] P. F. Antonietti, S. Giani, and P. Houston. hp-version composite discontinuous Galerkin methods for elliptic problems on complicated domains. *SIAM J. Sci. Comput.*, 35(3):A1417–A1439, 2013.

[5] P. F. Antonietti, G. Vacca, and M. Verani. Virtual element method for the Navier-Stokes equation coupled with the heat equation. *IMA J. Numer. Anal.*, 43(6):3396–3429, 2023.

[6] P.F. Antonietti, L. Beirão da Veiga, D. Mora, and M. Verani. A stream virtual element formulation of the Stokes problem on polygonal meshes. *SIAM J. Numer. Anal.*, 52(1):386–404, 2014.

[7] P.F. Antonietti, L.B. Beirao da Veiga, and G. Manzini. *The Virtual Element Method and its Applications*. SEMA SIMAI Springer series 31, Springer International Publishing, 2022.

[8] J. Baranger and K. Najib. Analyse numérique des écoulements quasi-newtoniens dont la viscosité obéit à la loi puissance ou la loi de carreau. *Numer. Math.*, 58(1):35–49, 1990.

[9] 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.

[10] J.W. Barrett and W.B. Liu. Finite element error analysis of a quasi-Newtonian flow obeying the Carreau or power law. *Numer. Math.*, 64(4):433–453, 1993.

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

[12] L. Beirão da Veiga, F. Brezzi, L. D. Marini, and A. Russo. The virtual element method. *Acta Numer.*, 32:123–202, 2023.

[13] L. Beirão da Veiga, F. Dassi, D. A. Di Pietro, and J. Droniou. Arbitrary-order pressure-robust DDR and VEM methods for the Stokes problem on polyhedral meshes. *Comput. Methods Appl. Mech. Engrg.*, 397:Paper No. 115061, 31, 2022.

[14] L. Beirão da Veiga, F. Dassi, and A. Russo. High-order virtual element method on polyhedral meshes. *Comput. Math. Appl.*, 74(5):1110–1122, 2017.

[15] L. Beirão da Veiga, F. Dassi, and G. Vacca. The Stokes complex for virtual elements in three dimensions. *Math. Models Methods Appl. Sci.*, 30(3):477–512, 2020.

[16] L. Beirão da Veiga, F. Dassi, and G. Vacca. Vorticity-stabilized virtual elements for the Oseen equation. *Mathematical Models and Methods in Applied Sciences*, 31(14):3009–3052, 2021.

[17] L. Beirão da Veiga, K. Lipnikov, and G. Manzini. *The mimetic finite difference method for elliptic problems*, volume 11 of *MS&A. Modeling, Simulation and Applications*. Springer, Cham, 2014.

[18] 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.

[19] 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.*, 51(2):509–535, 2017.

[20] 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.

[21] L. Beirão da Veiga, D. Mora, and G. Vacca. The Stokes complex for virtual elements with application to Navier-Stokes flows. *J. Sci. Comput.*, 81(2):990–1018, 2019.

[22] L. Beirão da Veiga and G. Vacca. Sharper error estimates for virtual elements and a bubble-enriched version. *SIAM J. Numer. Anal.*, 60(4):1853–1878, 2022.

[23] L. Beirão da Veiga, F. Brezzi, A. Cangiani, G. Manzini, L. D. Marini, and A. Russo. Basic principles of virtual element methods. *Math. Models Methods Appl. Sci.*, 23(1):199–214, 2013.

[24] L. Beirão da Veiga, C. Lovadina, and A. Russo. Stability analysis for the virtual element method. *Math. Mod.and Meth. in Appl. Sci.*, 27(13):2557–2594, 2017.

[25] L. Beirao da Veiga, D. Mora, and A. Silgado. A fully-discrete virtual element method for the nonstationary boussinesq equations in stream-function form. *Computer Methods in Applied Mechanics and Engineering*, 408:115947, 2023.

[26] 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.

[27] L.C. Berselli, L. Diening, and M. Růžička. Existence of strong solutions for incompressible fluids with shear dependent viscosities. *J. Math. Fluid Mech.*, 12(1):101–132, 2010.

[28] T. Bevilacqua, F. Dassi, S. Zampini, and S. Scacchi. BDDC Preconditioners for Virtual Element Approximations of the Three-Dimensional Stokes Equations. *SIAM J. Sci. Comput.*, 46(1):A156–A178, 2024.

[29] T. Bevilacqua and S. Scacchi. BDDC preconditioners for divergence free virtual element discretizations of the Stokes equations. *J. Sci. Comput.*, 92(2):Paper No. 63, 27, 2022.

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

[31] 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.

[32] M. Botti, D. Castanon Quiroz, D. A. Di Pietro, and A. Harnist. A Hybrid High-Order method for creeping flows of non-Newtonian fluids. *ESAIM: Math. Model Numer. Anal.*, 55(5):2045–2073, 2021.

[33] S. C. Brenner, Q. Guan, and L. Y. Sung. Some estimates for virtual element methods. *Comput. Methods Appl. Math.*, 17(4):553–574, 2017.

[34] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*, volume 15 of *Texts in Applied Mathematics*. Springer, New York, third edition, 2008.

[35] S. C. Brenner and L.Y. Sung. Virtual element methods on meshes with small edges or faces. *Math. Models Methods Appl. Sci.*, 28(7):1291–1336, 2018.

[36] H. Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Universitext. Springer New York, 2010.

[37] E. Cáceres and G. N. Gatica. A mixed virtual element method for the pseudostress-velocity formulation of the Stokes problem. *IMA J. Numer. Anal.*, 37(1):296–331, 2017.

[38] E. Cáceres, G. N. Gatica, and F. A. Sequeira. A mixed virtual element method for quasi-Newtonian Stokes flows. *SIAM J. Numer. Anal.*, 56(1):317–343, 2018.

[39] A. Cangiani, E. H. Georgoulis, and P. Houston. hp-version discontinuous Galerkin methods on polygonal and polyhedral meshes. *Math. Models Methods Appl. Sci.*, 24(10):2009–2041, 2014.

[40] A. Cangiani, V. Gyrya, and G. Manzini. The nonconforming virtual element method for the Stokes equations. *SIAM J. Numer. Anal.*, 54(6):3411–3435, 2016.

[41] D. Castanon Quiroz, D.A. Di Pietro, and A. Harnist. A Hybrid High-Order method for incompressible flows of non-newtonian fluids with power-like convective behaviour. *IMA Journal of Numerical Analysis*, 43(1):144–186, 12 2021.

[42] A. Chernov, C. Marcati, and L. Mascotto. p- and hp-virtual elements for the Stokes problem. *Adv. Comput. Math.*, 47(2):Paper No. 24, 31, 2021.

[43] P. G. Ciarlet and P. Ciarlet. Another approach to linearized elasticity and a new proof of korn's inequality. *Mathematical Models and Methods in Applied Sciences*, 15(02):259–271, 2005.

[44] B. Cockburn, J. Gopalakrishnan, and R. Lazarov. Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems. *SIAM J. Numer. Anal.*, 47(2):1319–1365, 2009.

[45] B. Cockburn, J. Gopalakrishnan, and F. J. Sayas. A projection-based error analysis of HDG methods. *Mathematics of Computation*, 79(271):1351–1367, 2010.

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

[47] D. A. Di Pietro and J. Droniou. *The hybrid high-order method for polytopal meshes*, volume 19 of *MS&A. Modeling, Simulation and Applications*. Springer, Cham, [2020] ©2020. Design, analysis, and applications.

[48] D. A. Di Pietro and A. Ern. A hybrid high-order locking-free method for linear elasticity on general meshes. *Comput. Meth. Appl. Mech. Engrg.*, 283:1–21, 2015.

[49] D. A. Di Pietro, A. Ern, and S. Lemaire. An arbitrary-order and compact-stencil discretization of diffusion on general meshes based on local reconstruction operators. *Comput. Methods Appl. Math.*, 14(4):461–472, 2014.

[50] L. Diening and F. Ettwein. Fractional estimates for non-differentiable elliptic systems with general growth. *Forum Math.*, 20(3):523–556, 2008.

[51] Q. Du and M.D. Gunzuburger. Finite-element approximation of a ladyzhenskaya model for stationary incompressible viscous flows. *SIAM J. Numer. Analysis*, 27:1–19, 1990.

[52] D. Frerichs and C. Merdon. Divergence-preserving reconstructions on polygons and a really pressure-robust virtual element method for the Stokes problem. *IMA J. Numer. Anal.*, 42(1):597–619, 2022.

[53] G. N. Gatica, M. Munar, and F. A. Sequeira. A mixed virtual element method for the Navier-Stokes equations. *Math. Models Methods Appl. Sci.*, 28(14):2719–2762, 2018.

[54] G. N. Gatica and F. A. Sequeira. An L^p spaces-based mixed virtual element method for the two-dimensional Navier-Stokes equations. *Math. Models Methods Appl. Sci.*, 31(14):2937–2977, 2021.

[55] G.N. Gatica and F. A. Sequeira. Analysis of an augmented HDG method for a class of quasi-newtonian Stokes flows. *J. Sci. Comput.*, 65(3):1270–1308, 2015.

[56] G. Geymonat and P. M. Suquet. Functional spaces for Norton-Hoff materials. *Math. Methods Appl. Sci.*, 8(2):206–222, 1986.

[57] A. Hirn. Approximation of the p-Stokes equations with equal-order finite elements. *J. Math. Fluid Mech.*, 15(1):65–88, 2013.

[58] A. Kaltenbach and M. Růžička. A local discontinuous Galerkin approximation for the p-Navier-Stokes system, part i: Convergence analysis. *SIAM Journal on Numerical Analysis*, 61(4):1613–1640, 2023.

[59] C. Kreuzer and E. Süli. Adaptive finite element approximation of steady flows of incompressible fluids with implicit power-law-like rheology. *ESAIM: Math. Model. Numer. Anal.*, 50(5):1333–1369, 2016.

[60] C. Kreuzer, R. Verfürth, and P. Zanotti. Quasi-optimal and pressure robust discretizations of the stokes equations by moment- and divergence-preserving operators. *Computational Methods in Applied Mathematics*, 21(2):423–443, 2021.

[61] F. Lepe and G. Rivera. A virtual element approximation for the pseudostress formulation of the Stokes eigenvalue problem. *Comput. Methods Appl. Mech. Engrg.*, 379:Paper No. 113753, 21, 2021.

[62] P. Lewintan and P. Neff. Nečas–Lions lemma revisited: An L^p-version of the generalized Korn inequality for incompatible tensor fields. *Math. Models Methods Appl. Sci.*, 44(14):11392–11403, 2021.

[63] Y. Li, C. Hu, and M. Feng. A Least-Squares stabilization virtual element method for the Stokes problem on polygonal meshes. *Int. J. Numer. Anal. Model.*, 19(5):685–708, 2022.

[64] Y. Li, C. Hu, and M. Feng. On stabilized equal-order virtual element methods for the Navier-Stokes equations on polygonal meshes. *Comput. Math. Appl.*, 154:267–286, 2024.

[65] T. Malkmus, M. Růžička, S. Eckstein, and I. Toulopoulos. Generalizations of SIP methods to systems with p-structure. *IMA Journal of Numerical Analysis*, 38(3):1420–1451, 09 2017.

[66] J. Meng, L. Beirão da Veiga, and L. Mascotto. Stability and interpolation properties for Stokes-like virtual element spaces. *J. Sci. Comput.*, 94(56), 2023.

[67] 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.

[68] C. Talischi, G. H. Paulino, A. Pereira, and I-F. M. Menezes. `PolyMesher`: a general-purpose mesh generator for polygonal elements written in Matlab. *Struct. Multidiscip. Optim.*, 45(3):309–328, 2012.

[69] G. Vacca. An H^1-conforming virtual element for Darcy and Brinkman equations. *Math. Models Methods Appl. Sci.*, 28(1):159–194, 2018.

[70] G. Wang and Y. Wang. Least-Squares Virtual Element Method for Stokes Problems on Polygonal Meshes. *J. Sci. Comput.*, 98(2):Paper no. 46, 2024.

[71] L. Wang. On Korn's inequality. *Journal of Computational Mathematics*, 21(3):321–324, 2003.

[72] 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.

A Appendix

We here show the proof of two technical lemmas.

Lemma 25. *Under the same assumptions of Proposition 22, let $\boldsymbol{u}_I \in \boldsymbol{U}_h$ be the interpolant of $\boldsymbol{u}$ (cf. Lemma 4) and $\boldsymbol{\xi}_h := \boldsymbol{u}_h - \boldsymbol{u}_I$, then the following holds*

$$\|(\boldsymbol{\sigma}(\cdot,\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h)) - \boldsymbol{\sigma}(\cdot,\boldsymbol{\epsilon}(\boldsymbol{u}))\|^{r'}_{\mathbb{L}^{r'}(\Omega)} \lesssim h^{k_1 r} R_1^r + \\ + \int_\Omega \Big(\boldsymbol{\sigma}(\cdot,\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h)) - \boldsymbol{\sigma}(\cdot,\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_I))\Big) : \boldsymbol{\Pi}^0_{k-1}\epsilon(\boldsymbol{\xi}_h)\,.$$

Proof. Employing (62a) with $\delta = 0$ we have

$$\begin{aligned}
&\|(\boldsymbol{\sigma}(\cdot,\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h)) - \boldsymbol{\sigma}(\cdot,\boldsymbol{\epsilon}(\boldsymbol{u}))\|^{r'}_{\mathbb{L}^{r'}(\Omega)} \lesssim \int_{\Omega_h} (|\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h)| + |\boldsymbol{\epsilon}(\boldsymbol{u})|)^{(r-2)r'} |\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h) - \boldsymbol{\epsilon}(\boldsymbol{u})|^{r'} \\
&= \int_\Omega \Big(\Big[(|\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h)| + |\boldsymbol{\epsilon}(\boldsymbol{u})|)^{(r-2)(r'-1)} |\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h) - \boldsymbol{\epsilon}(\boldsymbol{u})|^{r'-2} \Big] \times \\
&\qquad \times \Big[(|\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h)| + |\boldsymbol{\epsilon}(\boldsymbol{u})|)^{(r-2)} |\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h) - \boldsymbol{\epsilon}(\boldsymbol{u})|^2 \Big] \Big) =: \int_\Omega T_1 \times T_2\,,
\end{aligned} \tag{84}$$

Following [9], employing $r - 2 \le 0$ and $r' \ge 2$, together with $|X| + |Y| \ge |X - Y|$ for $X, Y \in \mathbb{R}^{d\times d}$ and noticing that $(r-2)(r'-1) + r' - 2 = (r-1)r' - r = 0$, we have

$$T_1 \lesssim |\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h) - \boldsymbol{\epsilon}(\boldsymbol{u})|^{(r-2)(r'-1)+(r'-2)} = 1\,. \tag{85}$$

We now estimate the term T_2 in (84). In the following C will denote a generic positive constant independent of h that may change at each occurrence, whereas the parameter ε adopted in (87) and (88) will be specified later. Using (62b) we have

$$\begin{aligned}
T_2 &\le C\big(\boldsymbol{\sigma}(\cdot,\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h)) - \boldsymbol{\sigma}(\cdot,\boldsymbol{\epsilon}(\boldsymbol{u}))\big) : \big(\Pi^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h) - \boldsymbol{\epsilon}(\boldsymbol{u})\big) \\
&= C\big(\boldsymbol{\sigma}(\cdot,\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h)) - \boldsymbol{\sigma}(\cdot,\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_I))\big) : \big(\Pi^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h) - \boldsymbol{\epsilon}(\boldsymbol{u})\big) + \\
&\quad + C\big(\boldsymbol{\sigma}(\cdot,\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_I)) - \boldsymbol{\sigma}(\cdot,\boldsymbol{\epsilon}(\boldsymbol{u}))\big) : \big(\Pi^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h) - \boldsymbol{\epsilon}(\boldsymbol{u})\big) =: T_2^A + T_2^B\,.
\end{aligned} \tag{86}$$

Using Lemma 21 and Lemma 20 we obtain

$$\begin{aligned} T_2^A &\le C\varphi'_{|\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h)|}(|\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{\xi}_h)|)\,|\Pi^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h)-\boldsymbol{\epsilon}(\boldsymbol{u})| && \text{(by (62a))}\\ &\le \varepsilon\varphi_{|\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h)|}(|\Pi^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h)-\boldsymbol{\epsilon}(\boldsymbol{u})|)+C(\varepsilon)\varphi_{|\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h)|}(|\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{\xi}_h)|) && \text{(by (60))}\\ &\le \gamma\varepsilon T_2 + C(\varepsilon)\big((\boldsymbol{\sigma}(\cdot,\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h))-\boldsymbol{\sigma}(\cdot,\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_I)))\,:\,\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{\xi}_h)\big) && \text{(by (62b))} \end{aligned} \tag{87}$$

where in the last line γ denotes the associated uniform hidden positive constant. Using analogous arguments we have

$$\begin{aligned} T_2^B &\le C\varphi'_{|\boldsymbol{\epsilon}(\boldsymbol{u})|}(|\boldsymbol{\epsilon}(\boldsymbol{u})-\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_I)|)\,|\boldsymbol{\epsilon}(\boldsymbol{u})-\Pi^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h)| && \text{(by (62a))}\\ &\le \varepsilon\varphi_{|\boldsymbol{\epsilon}(\boldsymbol{u})|}(|\boldsymbol{\epsilon}(\boldsymbol{u})-\Pi^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h)|)+C(\varepsilon)\varphi_{|\boldsymbol{\epsilon}(\boldsymbol{u})|}(|\boldsymbol{\epsilon}(\boldsymbol{u})-\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_I)|) && \text{(by (60))}\\ &\le \gamma\varepsilon T_2 + C(\varepsilon)(|\boldsymbol{\epsilon}(\boldsymbol{u})|+|\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_I)|)^{r-2}|\boldsymbol{\epsilon}(\boldsymbol{u})-\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_I)|^2 && \text{(by (62b))}\\ &\le \gamma\varepsilon T_2 + C(\varepsilon)|\boldsymbol{\epsilon}(\boldsymbol{u})-\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_I)|^r && (r-2\le 0) \end{aligned} \tag{88}$$

Taking $\epsilon = \frac{1}{4\gamma}$ in (87) and (88), from (86) we obtain

$$T_2 \lesssim \big(\boldsymbol{\sigma}(\cdot,\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_h))-\boldsymbol{\sigma}(\cdot,\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_I))\,:\,\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{\xi}_h)\big)+|\boldsymbol{\epsilon}(\boldsymbol{u})-\boldsymbol{\Pi}^0_{k-1}\boldsymbol{\epsilon}(\boldsymbol{u}_I)|^r\,.$$

The proof follows combining the bound above with (85) in (84) and employing Lemma 4. □

Lemma 26. *Under the same assumptions of Proposition 22, let $\boldsymbol{u}_I \in \boldsymbol{U}_h$ be the interpolant of $\boldsymbol{u}$ (cf. Lemma 4) and $\boldsymbol{\xi}_h := \boldsymbol{u}_h - \boldsymbol{u}_I$, then the following holds*

$$S(\widetilde{\boldsymbol{u}}_h,\widetilde{\boldsymbol{u}}_h) \lesssim S(\widetilde{\boldsymbol{u}}_h,\widetilde{\boldsymbol{\xi}}_h) - S(\widetilde{\boldsymbol{u}}_I,\widetilde{\boldsymbol{\xi}}_h) + h^{k_1 r}R_1^r\,,$$

where $\widetilde{\boldsymbol{v}} := (I-\Pi^0_k)\boldsymbol{v}$ for any $\boldsymbol{v}\in\boldsymbol{L}^2(\Omega)$.

Proof. Let $\widehat{\boldsymbol{\sigma}}(\boldsymbol{x}) := |\boldsymbol{x}|^{r-2}\boldsymbol{x}$ for any $\boldsymbol{x}\in\mathbb{R}^{N_E}$, where N_E denotes the number of local DoFs (cf. Section 3.3). Then simple computations yield

$$\begin{aligned} S(\widetilde{\boldsymbol{u}}_h,\widetilde{\boldsymbol{u}}_h) &= \sum_{E\in\Omega_h} h_E^{2-r}|\boldsymbol{\chi}(\widetilde{\boldsymbol{u}}_h)|^r = \sum_{E\in\Omega_h} h_E^{2-r}\widehat{\boldsymbol{\sigma}}(\boldsymbol{\chi}(\widetilde{\boldsymbol{u}}_h))\cdot\boldsymbol{\chi}(\widetilde{\boldsymbol{u}}_h)\\ &= \sum_{E\in\Omega_h} h_E^{2-r}\big(\widehat{\boldsymbol{\sigma}}(\boldsymbol{\chi}(\widetilde{\boldsymbol{u}}_h))-\widehat{\boldsymbol{\sigma}}(\boldsymbol{\chi}(\widetilde{\boldsymbol{u}}_I))\big)\cdot\boldsymbol{\chi}(\widetilde{\boldsymbol{u}}_h) + \sum_{E\in\Omega_h} h_E^{2-r}\widehat{\boldsymbol{\sigma}}(\boldsymbol{\chi}(\widetilde{\boldsymbol{u}}_I))\cdot\boldsymbol{\chi}(\widetilde{\boldsymbol{u}}_h)\\ &=: T_1+T_2\,. \end{aligned} \tag{89}$$

In the following C will denote a generic positive constant independent of h that may change at each occurrence, whereas the parameter ε adopted in (90) and (91) will be specified later. Using Lemma 20 and Lemma 21 we infer

$$\begin{aligned} T_1 &\le C\sum_{E\in\Omega_h} h_E^{2-r}\varphi'_{|\boldsymbol{\chi}(\widetilde{\boldsymbol{u}}_h)|}(|\boldsymbol{\chi}(\widetilde{\boldsymbol{\xi}}_h)|)\,|\boldsymbol{\chi}(\widetilde{\boldsymbol{u}}_h)| && \text{(by (62a))}\\ &\le \varepsilon\sum_{E\in\Omega_h} h_E^{2-r}\varphi_{|\boldsymbol{\chi}(\widetilde{\boldsymbol{u}}_h)|}(|\boldsymbol{\chi}(\widetilde{\boldsymbol{u}}_h)|) + C(\varepsilon)\sum_{E\in\Omega_h} h_E^{2-r}\varphi_{|\boldsymbol{\chi}(\widetilde{\boldsymbol{u}}_h)|}(|\boldsymbol{\chi}(\widetilde{\boldsymbol{\xi}}_h)|) && \text{(by (60))}\\ &\le \gamma\varepsilon\sum_{E\in\Omega_h} h_E^{2-r}|\boldsymbol{\chi}(\widetilde{\boldsymbol{u}}_h)|^r + C(\varepsilon)\sum_{E\in\Omega_h} h_E^{2-r}(|\boldsymbol{\chi}(\widetilde{\boldsymbol{u}}_h)|+|\boldsymbol{\chi}(\widetilde{\boldsymbol{u}}_I)|)^{r-2}\,|\boldsymbol{\chi}(\widetilde{\boldsymbol{\xi}}_h)|^2 && \text{(by (62b))}\\ &= \gamma\varepsilon S(\widetilde{\boldsymbol{u}}_h,\widetilde{\boldsymbol{u}}_h) + C(\varepsilon)\sum_{E\in\Omega_h} h_E^{2-r}(|\boldsymbol{\chi}(\widetilde{\boldsymbol{u}}_h)|+|\boldsymbol{\chi}(\widetilde{\boldsymbol{u}}_I)|)^{r-2}\,|\boldsymbol{\chi}(\widetilde{\boldsymbol{\xi}}_h)|^2\,, \end{aligned} \tag{90}$$

where in the last line γ denotes the uniform hidden positive constant from Lemma 21. Using analogous

arguments we have

$$
\begin{aligned}
T_2 &\le C \sum_{E\in\Omega_h} h_E^{2-r}\varphi'(|\chi(\widetilde{\boldsymbol{u}}_I)|)\,|\chi(\widetilde{\boldsymbol{u}}_h)| && \text{(by (62a))}\\
&\le \varepsilon \sum_{E\in\Omega_h} h_E^{2-r}\varphi(|\chi(\widetilde{\boldsymbol{u}}_h)|) + C(\varepsilon)\sum_{E\in\Omega_h} h_E^{2-r}\varphi(|\chi(\widetilde{\boldsymbol{u}}_I)|) && \text{(by (60))}\\
&\le \gamma\varepsilon \sum_{E\in\Omega_h} h_E^{2-r}|\chi(\widetilde{\boldsymbol{u}}_h)|^r + C(\varepsilon)\sum_{E\in\Omega_h} h_E^{2-r}|\chi(\widetilde{\boldsymbol{u}}_I)|^r && \text{(by (62b))}\\
&= \gamma\varepsilon S(\widetilde{\boldsymbol{u}}_h,\widetilde{\boldsymbol{u}}_h) + C(\varepsilon)\sum_{E\in\Omega_h} h_E^{2-r}|\chi(\widetilde{\boldsymbol{u}}_I)|^r\,,
\end{aligned}
\tag{91}
$$

Taking in (90) and (91) $\epsilon = \frac{1}{4\gamma}$, from (89) we obtain

$$
S(\widetilde{\boldsymbol{u}}_h,\widetilde{\boldsymbol{u}}_h) \lesssim \sum_{E\in\Omega_h} h_E^{2-r}(|\chi(\widetilde{\boldsymbol{u}}_h)| + |\chi(\widetilde{\boldsymbol{u}}_I)|)^{r-2}\,|\chi(\widetilde{\boldsymbol{\xi}}_h)|^2 + \sum_{E\in\Omega_h} h_E^{2-r}|\chi(\widetilde{\boldsymbol{u}}_I)|^r\,.
$$

We now notice that the right-hand side in the previous bound coincides, up to a multiplicative constant, with the right-hand side in (70). Therefore employing (71) and (72) we finally obtain

$$
S(\widetilde{\boldsymbol{u}}_h,\widetilde{\boldsymbol{u}}_h) \lesssim S(\widetilde{\boldsymbol{u}}_h,\widetilde{\boldsymbol{\xi}}_h) - S(\widetilde{\boldsymbol{u}}_I,\widetilde{\boldsymbol{\xi}}_h) + h^{k_1 r}R_1^r\,.
$$

□