Research Article

Josefin Ahlkrona* and Malte Braack

Equal-Order Stabilized Finite Element Approximation of the *p*-Stokes Equations on Anisotropic Cartesian Meshes

Abstract: The *p*-Stokes equations with power-law exponent $p \in (1, 2)$ describes non-Newtonian, shear-thinning, incompressible flow. In many industrial applications and natural settings, shear-thinning flow takes place in very thin domains. To account for such anisotropic domains in simulations, we here study an equal-order bi-linear anisotropic finite element discretization of the *p*-Stokes equations, and extend a non-linear Local Projection Stabilization to anisotropic meshes. We prove an a priori estimate and illustrate the results with two numerical examples, one confirming the rate of convergence predicted by the a-priori analysis, and one showing the advantages of an anisotropic stabilization compared to an isotropic one.

Keywords: *p*-Stokes, Anisotropic Meshes, Finite Element Method, Shear-Thinning, Local Projection Stabilization

MSC 2010: 65N30, 65M12, 76A05, 76D07

1 Introduction

Shear-thinning flow in thin domains is common in industry and nature, and numerical simulations is an important tool to understand such problems. The viscosity of shear-thinning (or pseudo-plastic) fluids increases when shear rates decrease, which renders numerical simulations both more computationally demanding, and harder to analyze. Thin domains further increase the complexity of the problem, and tends to reduce the accuracy and stability of numerical methods. Examples of shear thinning flows taking place on thin domains in industry are paint drying on a wall, hot (metal) rolling, and polymer coating processes. An interesting example from nature is the flow of ice in ice sheets, such as the Greenland Ice Sheet or the Antarctic Ice Sheet. There are also non-Newtonian fluids for which the viscosity decreases for decreasing shear, so called shear-thickening (or dilatant) fluids. As such those are much less common than shear-thinning fluids, we here focus on the latter.

The equations that model shear-dependent (steady) flow are the *p*-Stokes equations,

$$\left. \begin{array}{r} -\nabla \cdot \mathbf{S}(\mathbf{Dv}) + \nabla \pi = \mathbf{f}, \\ \nabla \cdot \mathbf{v} = 0, \end{array} \right\} \quad \text{in } \Omega, \tag{1.1}$$

where $\mathbf{v} = (v_x, v_y)$ is the velocity field in a Cartesian coordinate system (x, y), π is the pressure, $\mathbf{S}$ is the shear-dependent extra stress tensor, and $\mathbf{f}$ a given forcing. Here, the two-dimensional domain $\Omega \in \mathbb{R}^2$ is open and bounded with polygonal boundary $\partial\Omega$ (for simplicity). We will consider Dirichlet condition on the entire

*Corresponding author: Josefin Ahlkrona, Mathematisches Seminar, Christian-Albrechts-Universität zu Kiel, Westring 383, 24118 Kiel, Germany. Current address: Department of Mathematics, Stockholm University, Stockholm, Sweden, e-mail: ahlkrona@math.su.se. https://orcid.org/0000-0003-0329-0253
Malte Braack, Mathematisches Seminar, Christian-Albrechts-Universität zu Kiel, Westring 383, 24118 Kiel, Germany, e-mail: braack@math.uni-kiel.de

boundary $\partial\Omega$ and enforce a pressure of zero-mean value,

$$\mathbf{v} = \mathbf{g} \quad \text{on } \partial\Omega, \qquad \int_\Omega \pi = 0.$$

In order for the equations to earn the name *p*-Stokes, the stress tensor **S** must exhibit a so-called (p, ϵ)-structure (defined in Assumption 1 in Section 2.5). A common example of a stress-tensor with (p, ϵ) structure is

$$\mathbf{S}(\nabla\mathbf{v}) = \mu_0(\epsilon^2 + |\mathbf{Dv}|^2)^{\frac{p-2}{2}} \mathbf{Dv}, \tag{1.2}$$

where $\mathbf{Dv} = (\nabla\mathbf{v})^{\text{sym}} = \frac{1}{2}(\nabla\mathbf{v} + (\nabla\mathbf{v})^T)$ is the strain rate tensor. Note that $S(\nabla\mathbf{v}) = S(\mathbf{Dv})$ by definition. If $\epsilon = 0$, equation (1.2) describes a so-called power-law fluid, and for $\epsilon > 0$ it represents a Carreau-fluid. The viscosity-parameter μ_0 scales the magnitude of the viscosity $\mu := \mu_0(\epsilon^2 + |\mathbf{Dv}|^2)^{\frac{p-2}{p}}$, we will set $\mu_0 = 1.0$ for brevity of the presentation. The exponent $p \in (1, \infty)$ determines the non-linearity of the material. If $p = 2$, the viscosity μ is constant ($\mu = \mu_0$), and the system of (1.1) is simply the common steady Stokes problem describing Newtonian fluids. If $p \in (1, 2)$, the fluid is shear-thinning, and for $p > 2$ the fluid is shear-thickening. Hence we consider the range $p \in (1, 2)$ in this work. Parts of the analysis is identical for $p \in (2, \infty)$, while other parts would require extensions.

The finite element discretization of the *p*-Stokes equations on isotropic meshes has been studied since the early 1990s, see [3] and [23]. The analysis is closely related to the analysis of the *p*-Laplacian, and involves interpolation estimates in quasi-norms of the type in [3]. Such estimates can be found if the problem is transferred from classical Sobolev spaces to Orlicz–Sobolev spaces (see [14]). Using this approach, optimal a priori estimates for the shear-thinning case were derived independently by [6] and [18, 19]. In [6], inf-sup stable elements are used, while in [18, 19] stabilized equal-order finite elements are employed. The stabilization there is a Local Projection Stabilization (LPS) [5] modified to suit the *p*-Stokes equations. Further, we like to refer to the work of Diening, Kreuzer and Süli [13], where even more general models of rheology are discretized with inf-sup stable finite elements.

Since we are studying thin domains, the *x*-direction of Ω is assumed to be much wider than the *y*-direction. In order to resolve thin domains in a computationally efficient manner, a finite element discretization should employ anisotropic meshes. Anisotropic meshes are also useful for capturing a type of boundary layer that may occur in shear-thinning free surface flow due to the varying viscosity. These boundary layers are in general not as sharp as boundary layers in convective problems, but nevertheless need to be resolved [1, 20].

The introduction of anisotropic meshes complicates both analysis and computations. Anisotropic meshes are quite well-studied for Newtonian fluids, see e.g. [2, 7, 22]. In [10] the LPS stabilization was extended to anisotropic meshes for Newtonian fluids. However, in the non-Newtonian case, there is to our knowledge no study on anisotropic meshes. In this work we extend the stabilization of [19] to anisotropic meshes by combining the anisotropic stabilization term in [10] with the idea behind the non-linear one in [18]. We focus on (bi-)linear elements, as the regularity of the solutions are limited, which may render high-order elements sub-optimal [18]. In order to derive a priori estimates, we extend the interpolation estimates in the Orlicz–Sobolev spaces of [14] to anisotropic meshes.

An example of an application where a stabilization term especially adapted to both a non-linear rheology and anisotropic meshes would be of benefit is ice sheet modelling. Ice sheets are very thin compared to the vast areas they cover, and under long time scales ice sheet dynamics can be described as a shear-thinning thin film flow with $p \approx 1.3$. Stabilized low-order equal finite elements on anisotropic meshes is a common approach in this field, but the lack of specialized stabilization terms is problematic [17].

The paper is structured as follows. In Section 2 the finite element formulation of the *p*-Stokes equations on anisotropic meshes is introduced. Some fundamental tools for working with the *p*-Stokes system are also described here, such as *N*-functions that are needed to transfer to Orlicz–Sobolev spaces. Section 3 is dedicated to the error analysis. First, a modified interpolation operator is constructed and interpolation estimates are shown for anisotropic meshes, including the estimates in quasi-norms that are needed for the *p*-Stokes system. Then the well-posedness of the problem is shown and a priori estimates for the velocity and pressure are derived. Numerical experiments in Section 4 verify the error analysis and compares the performance of isotropic stabilization to that of anisotropic stabilization. The paper concludes in Section 5 with a summary.

2 Finite Element Formulation

2.1 Notation

The spatial dimension is denoted by d, and as we work in two dimensions $d = 2$, although a generalization to $d = 3$ with almost uniform spatial horizontal mesh sizes is possible (see details below). The scalar product between two vectors $\mathbf{q}$ and $\mathbf{p}$ is $\mathbf{q} \cdot \mathbf{p}$. Between two tensors $\mathbf{Q} : \mathbf{P} \in \mathbb{R}^{d \times d}$ it is defined as $\mathbf{Q} : \mathbf{P} = \sum_{i,j=1}^{d} Q_{ij} P_{ij}$, and we use $|\mathbf{Q}| = (\mathbf{Q} : \mathbf{Q})^{\frac{1}{2}}$. The symmetric part of a tensor $\mathbf{Q}$ is $\mathbf{Q}^{\mathrm{sym}} := \frac{1}{2}(\mathbf{Q} + \mathbf{Q}^T)$. We will sometimes use a multi-index notation with a multi-index $\alpha = (\alpha_1, \alpha_2)$ to write partial derivatives as $\partial^\alpha := \partial_x^{\alpha_1} \partial_y^{\alpha_2}$.

For $v \in [1, \infty]$, $L^v(\Omega)$ denotes the Lebesgue space and $W^{m,v}(\Omega)$ is a Sobolev space of order m. The space $L_0^v(\Omega)$ is defined as $\{q \in L^v(\Omega) : \int_\Omega q = 0\}$ while $W_0^{m,v}(\Omega)$ is the standard Sobolev space of order m with vanishing traces on $\partial\Omega$. The Lebesgue measure of a domain $\omega \subset \Omega$ is written $|\omega|$. The $L^v(\omega)$-norm is $\| \cdot \|_{v;\omega}$ and the notation for the $W^{m,v}(\omega)$-norm is $\| \cdot \|_{m,v;\omega}$. The integral $\int_\omega uv \, d\omega$ is denoted $(u, v)_\omega$. The norm in the vector space $\mathbf{W}^{m,v}(\Omega) := [W^{m,v}(\Omega)]^d$ is $\|\mathbf{q}\|_{m,v,\Omega} = (\sum_{1 \le i \le d} \sum_{0 \le |\alpha| \le m} \|\partial^\alpha q_i\|_{v,\Omega}^v)^{\frac{1}{v}}$. If there exists $C > 0$ so that $f \le Cg$, we will write $f \lesssim g$, and we use the notation $f \sim g$ if there exist constants $c, C > 0$ so that $cf \le g \le Cf$. For brevity of presentation we sometimes write integrals of some function g as $\int g$ instead of $\int g \, dx$.

2.2 Weak Formulation and Discretization

The weak formulation of the p-Stokes system is: Find velocity and pressure $(\mathbf{v}, \pi) \in \mathcal{X} \times \mathcal{Q}$ so that

$$(\mathbf{S}(\mathbf{Dv}), \mathbf{D}\boldsymbol{\omega})_\Omega - (\pi, \nabla \cdot \boldsymbol{\omega})_\Omega = (\mathbf{f}, \boldsymbol{\omega})_\Omega \quad \text{for all } \boldsymbol{\omega} \in \mathcal{X}, \tag{2.1}$$

$$(\nabla \cdot \mathbf{v}, q)_\Omega = 0 \quad \text{for all } q \in \mathcal{Q}. \tag{2.2}$$

The natural spaces for the solution are $\mathcal{X} := \mathbf{W}_0^{1,p}(\Omega)$ and $\mathcal{Q} := L_0^{p'}(\Omega)$, where $p' = \frac{p}{p-1}$. We only state the case of homogeneous Dirichlet condition, because the case of inhomogeneous boundary data can be reformulated in the form (2.1)–(2.2) with modified right-hand side $\mathbf{f}$.

Now let $\mathbb{T}_h$ be a decomposition of Ω into rectangles. For simplicity and brevity of presentation, we assume the mesh to be globally uniform. On the decomposition $\mathbb{T}_h$ we introduce a finite element space of bi-linear, continuous functions

$$X_h := X_h(\mathbb{T}_h) := \{\omega \in C(\overline{\Omega}) : \omega|_K \circ F_K \in \mathbb{Q}_1(\hat{K}) \text{ for all } K \in \mathbb{T}_h\},$$

where the space $\mathbb{Q}_1(\hat{K}) := \mathrm{span}\{\hat{1}, \hat{x}, \hat{y}, \hat{x}\hat{y} : (\hat{x}, \hat{y}) \in \hat{K}\}$ is defined on the reference quadrilateral $\hat{K} := (0, 1)^d$, and F_K maps from the reference element to an element K

$$F_K(\hat{\mathbf{x}}) = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + \begin{pmatrix} h_x & 0 \\ 0 & h_y \end{pmatrix} \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix},$$

where the width of a quadrilateral in the x-direction is h_x, and in the y-direction h_y. We introduce the equal-order finite element spaces

$$\mathcal{X}_h := \mathcal{X} \cap X_h^d \quad \text{and} \quad \mathcal{Q}_h := \mathcal{Q} \cap X_h$$

and formulate a (unstabilized) discretized problem as: Find velocity and pressure $(\mathbf{v}_h, \pi_h) \in \mathcal{X}_h \times \mathcal{Q}_h$, so that

$$(\mathbf{S}(\mathbf{Dv}_h), \mathbf{D}\boldsymbol{\omega}_h)_\Omega - (\pi_h, \nabla \cdot \boldsymbol{\omega}_h)_\Omega = (\mathbf{f}, \boldsymbol{\omega}_h)_\Omega \quad \text{for all } \boldsymbol{\omega}_h \in \mathcal{X}_h, \tag{2.3}$$

$$(\nabla \cdot \mathbf{v}_h, q_h)_\Omega = 0 \quad \text{for all } q_h \in \mathcal{Q}_h. \tag{2.4}$$

The equal-order finite element pair $\mathcal{X}_h \times \mathcal{Q}_h$ violates the inf-sup condition [12], and to avoid spurious oscillations in the pressure, stabilization is needed. The most common argument for using stabilized equal-order elements instead of inf-sup stable elements is that they are easier to implement. Additionally, an inf-sup stable implementation on highly anisotropic meshes requires $Q_k \times Q_{k-2}$-elements in general (rather than $Q_k \times Q_{k-1}$-elements). If a continuous pressure is required, third-order polynomials should thus be used to
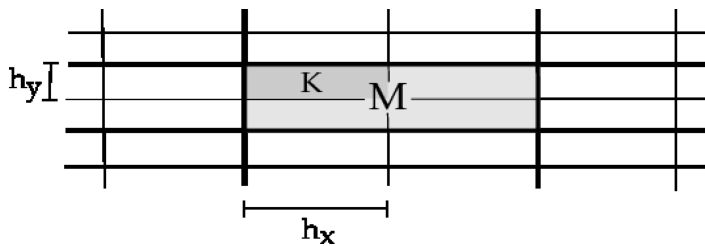
Figure 1: A mesh $\mathbb{T}_h$ consisting of elements K. The element is anisotropic since it is longer in the *x*-direction than in the *y*-direction. Each element belongs to a patch M.

approximate the velocity space. Such a high polynomial order in the velocity is usually sup-optimal because the low-order pressure approximation reduces the overall accuracy, and because of the low regularity of the solution [19]. Generalization to $d = 3$ with horizontal mesh sizes $h_x \sim h_y$ and smaller vertical mesh size $h_z << h_x$ is straightforward. However, we restrict here to the two-dimensional case for ease of presentation.

2.3 Stabilization

In this work we stabilize (2.3) with a non-linear stabilization term $s(\pi_h, q_h)$ of local projection type

$$(\mathbf{S}(\mathbf{Dv}_h), \mathbf{D}\boldsymbol{\omega}_h)_\Omega - (\pi_h, \nabla \cdot \boldsymbol{\omega}_h)_\Omega = (\mathbf{f}, \boldsymbol{\omega}_h)_\Omega \quad \text{for all } \boldsymbol{\omega}_h \in \mathfrak{X}_h, \tag{2.5}$$

$$(\nabla \cdot \mathbf{v}_h, q_h)_\Omega + s(\pi_h, q_h) = 0 \qquad \text{for all } q_h \in \mathfrak{Q}_h. \tag{2.6}$$

To define the stabilization term we introduce some further notations. The mesh is arranged in a patch-wise manner, see Figure 1. Hence, we have a second shape-regular decomposition $\mathbb{M}_h$ of Ω where each element (or patch) $M \in \mathbb{M}_h$ consists of four rectangles of $\mathbb{T}_h$ so that $\mathbb{M}_h = \mathbb{T}_{2h}$. As in [5], we also define a space Y_h consisting of piecewise constants on each patch M,

$$Y_h := \{\omega \in L^2(\Omega) : \omega_M \circ F_M \in \mathbb{Q}_0(\hat{M})\}.$$

The restriction of the space Y_h to a patch M is denoted by $Y_h(M) := \{q_h|_M : q_h \in Y_h\}$ and we denote

$$X_h^0(M) := \{\omega_h|_M : \omega_h \in X_h, \ \omega_h = 0 \text{ on } \Omega \setminus M\}.$$

We define a fluctuation operator as $\theta_h := \mathrm{id} - P_h$, where id is the identity and $P_h : L^v \to Y_h$ is projecting to a smaller space Y_h so that $(P_h q_h)|_M := P_M(q_h|_M)$ for all $M \in \mathbb{M}_h$, where P_M is the local projection $P_M : L^v(M) \to Y_h(M)$. The projection in this setting is simply

$$P_h q_h = \frac{1}{|M|} \int_M q_h.$$

The stabilization term is given by

$$s(\pi_h, q_h) := \alpha_0 \sum_{M \in \mathbb{M}} (\mathbf{M}_s \boldsymbol{\theta}_h \nabla \pi_h, \boldsymbol{\theta}_h \nabla q_h)_M, \tag{2.7}$$

$$\mathbf{M}_s(\partial_x \pi_h, \partial_y \pi_h) := \begin{bmatrix} h_x^2(\frac{1}{\tau}(\tau + |\theta_h \partial_x \pi_h|))^{p'-2} & 0 \\ 0 & h_y^2(\frac{1}{\tau}(\tau + \frac{h_y}{h_x}|\theta_h \partial_y \pi_h|))^{p'-2} \end{bmatrix}, \tag{2.8}$$

where $\tau \geq 0$ and $\alpha_0 \geq 0$ are user-defined parameters and $\boldsymbol{\theta}$ denotes the fluctuation operator applied to vector-valued functions in a component-wise manner.

2.4 Anisotropic Meshes

The important point in this work is that the mesh is allowed to be anisotropic, so that

$$\mathcal{A} := \frac{h_x}{h_y} \gg 1.$$

That is why the factor $\frac{h_y}{h_x}$ and the factor h_y^2 is introduced in (2.8). Without the former one, we will see that the discretization errors are proportional to $(\frac{h_x}{h_y})^{\frac{p'-2}{2}}$. Such a factor can become extremely large as $p \to 1$ (i.e. $p' \to \infty$ and $\frac{h_x}{h_y} \gg 1$. Exchanging the latter h_y^2 factor to the more standard h_x^2 factor increases errors even more. Note that for $h_x = h_y = h$ the stabilization term is identical to that of [19], and for $p = p' = 2$ it is identical to the stabilization of [5] (if $h_x = h_y = h$) or [10] (if $h_x > h_y$).

On these anisotropic meshes, it is important that the following local inf-sup condition holds, with a constant that is independent of the aspect ratio $\mathcal{A}$.

Lemma 1. *For $\nu \geq 1$ there exists $\beta_{\mathrm{loc}} > 0$ independent of h_x and h_y such that*

$$\inf_{q_h \in Y_h(M)} \sup_{\omega_h \in X_h^0(M)} \frac{(\omega_h, q_h)_M}{\|\omega_h\|_{\nu,M} \|q_h\|_{\nu',M}} \geq \beta_{\mathrm{loc}} > 0$$

for all $h_x > 0$, $h_y > 0$ and all $M \in \mathbb{M}$, where $\nu' = \frac{\nu}{\nu-1}$.

Here it is important that β_{loc} is independent of the aspect ratio $\mathcal{A}$. The parameter ν will in practice equal to p.

Proof. The proof is similar to that of [21, Lemma 3.2], which considers isotropic elements and $\nu = \nu' = 2$. The extension from $\nu = 2$ to $\nu \in (1, \infty)$ is straightforward, see [18]. In order to extend the proof to anisotropic meshes, we consider that the proof in [21] is based on transferring back and forth between reference elements and physical elements, using an estimation of the determinant of the Jacobian of the transformation from the reference elements. On our simple anisotropic meshes we can use that we clearly have $|\det DF_K| = h_x h_y$. Following the approach of [21], we have the desired estimate. Details are given in the Appendix. $\square$

The last property that is required for the LPS-stabilization is for the fluctuation operator θ_h to be bounded, so that

$$\|\theta_h q\|_{0,\nu;M} \lesssim \|q\|_{0,\nu;M} \quad \text{for all } M \in \mathbb{M} \text{ and all } q \in L^\nu(\Omega).$$

This property follows directly from [18], also on anisotropic meshes.

2.5 Some Useful Help-Functions and Properties of S

2.5.1 The Help-Functions φ, $\mathcal{F}$ and Their Relation to the Stress Tensor S

Since the stress-tensor is non-linear, some extra functions and properties are required for the error analysis. The extra help functions that will be used in the analysis are the N-function φ, $\varphi^* \in C^1(\mathbb{R}_0^+, \mathbb{R}_0^+)$, the function $\mathcal{F} : \mathbb{R}^{d \times d} \to \mathbb{R}_{\mathrm{sym}}^{d \times d}$, and $\mathcal{G}_1, \mathcal{G}_2 : \mathbb{R} \to \mathbb{R}$. The functions φ, φ^* and $\mathcal{F}$ are related to the continuous formulation, and their properties and their relations to **S** are therefore identical to those in [6, 18, 19]. The N-function φ and the complementary N-function φ^* are defined as

$$\varphi(t) = \frac{1}{p} t^p \quad \text{and} \quad \varphi^*(t) = \frac{1}{p'} t^{p'}.$$

The *shifted* N-functions $\varphi_a \in C^1(\mathbb{R}_0^+, \mathbb{R}_0^+)$ are defined as

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

and similarly for φ^*. In particular, it is useful to apply a shift of ϵ to φ to get the φ_ϵ-function

$$\varphi_\epsilon(t) = \int_0^t (\epsilon + s)^{p-2} s\, ds,$$

which is of interest since it has a structure that is related to the stress-tensor. The function $\mathcal{F}$ also has a structure reminiscent of the stress-tensor. It is defined as

$$\mathcal{F}(\mathbf{P}) = (\epsilon + |\mathbf{P}^{\mathrm{sym}}|)^{\frac{p-2}{2}} \mathbf{P}^{\mathrm{sym}}. \tag{2.9}$$

In the analysis, we will exploit some relations between **S, D**, φ and $\mathcal{F}$ as given in e.g. [19], summarized in Lemmas 13 and 14 in the Appendix. These relations allow for working with the help-functions in the error-analysis instead of with the stress-tensor directly. An example of how this can be advantageous is that since φ is convex one can use Jensen's inequality, in cases where one in a linear analysis would have used the triangle inequality. Also, φ satisfies the Δ_2-condition, i.e. there exists a $c > 0$ such that for all $t \geq 0$, $\varphi(2t) \leq c\varphi(t)$. The smallest constant is called Δ_2. We have that $\Delta_2 < \infty$, implying that $\varphi(t) \sim \varphi(ct)$ (c being any constant), which clearly is a desired property in analysis. Even though we will consider a Carreau-type fluid with a stress tensor **S** of the form (1.2) in our numerical experiments, Lemmas 13 and 14 are valid for any stress-tensor that fulfills the following assumption, with $p \leq 2$.

Assumption 1 (On the Stress Tensor **S**). The extra stress tensor

$$\mathbf{S} : \mathbb{R}^{d \times d} \to \mathbb{R}^{d \times d}_{\text{sym}}$$

belongs to $C^0(\mathbb{R}^{d \times d}, \mathbb{R}^{d \times d}_{\text{sym}}) \cap C^1(\mathbb{R}^{d \times d} \setminus \{0\}, \mathbb{R}^{d \times d}_{\text{sym}})$ and satisfies $S(\mathbf{Q}) = S(\mathbf{Q}^{\text{sym}})$ and $\mathbf{S}(\mathbf{0}) = \mathbf{0}$. The stress tensor is of (p, ϵ)-structure, meaning that there exist $p \in (1, \infty)$, $\epsilon \in [0, \infty)$, and constants $\sigma_0, \sigma_1 > 0$ such that for all $\mathbf{Q}, \mathbf{P} \in \mathbb{R}^{d \times d}$ with $\mathbf{Q} \neq 0$,

$$\sum_{i,j,k,l=1}^{d} \partial_{kl} S_{ij}(\mathbf{Q}) P_{ij} P_{kl} \geq \sigma_0 (\epsilon + |\mathbf{Q}^{\text{sym}}|)^{p-2} |\mathbf{P}^{\text{sym}}|^2,$$

$$\partial_{kl} S_{ij}(\mathbf{Q}) \leq \sigma_1 (\epsilon + |\mathbf{Q}^{\text{sym}}|)^{p-2} \quad \text{for all } i, j, k, l \in \{1, \ldots, d\}.$$

The constants in our error estimates will often depend on σ_0 and σ_1. For brevity of presentation we will, except for in this subsection, not state this in the remainder of this paper.

2.5.2 The Help-Functions $\mathcal{G}_1$, $\mathcal{G}_2$ and Their Relation to the Stabilization Term and φ^*

To analyze the effect of the stabilization, some relations between the non-linear stabilization term $s_h(p, q)$, the help functions $\mathcal{G}_1$ and $\mathcal{G}_2$ and the complementary N-function $\varphi^*(t) = \frac{t^{p'}}{p'}$ are required. We here adapt the analysis of [18, 19] for the isotropic problem to anisotropic meshes. In particular, we use two functions $\mathcal{G}_1$ and $\mathcal{G}_2$

$$\mathcal{G}_1(\mathbf{q}) = (\tau + |\mathbf{q}|)^{\frac{p'-2}{2}} \mathbf{q}, \quad \mathcal{G}_2(\mathbf{q}) = (\tau_2 + |\mathbf{q}|)^{\frac{p'-2}{2}} \mathbf{q},$$

instead of one ($\mathcal{G}$ in [18, 19]), and we revisit [19, Lemmas 3.1–3.3] to state them in an anisotropic setting. Here $\tau_2 = \frac{h_x}{h_y} \tau$. The function $\mathcal{G}_1$ will be used to analyze the x-components of the error, and $\mathcal{G}_2$ will be used for the y-components.

Lemma 2. *Let $U \subseteq \Omega$ be a measurable subset and let $\mathbf{g}, \mathbf{q} \in L^{p'}(\Omega)$, with $p \in (1, 2]$. Then*

$$\|\mathbf{g} - \mathbf{q}\|_{p', U}^{p'} \lesssim \|\mathcal{G}_i(\mathbf{g}) - \mathcal{G}_i(\mathbf{q})\|_{2, U}^2, \quad i = 1, 2,$$

where the constant c only depends on p.

Proof. The proof is based on a vector-valued version of Lemma 13, see [18] for details. $\square$

Lemma 3. *For all $\pi, q \in W^{1,p'}(\Omega)$ there holds*

$$s(\pi, \pi - q) - s(q, \pi - q)$$

$$\sim \alpha_0 \sum_{M \in \mathbb{M}_h} \left(h_x^2 \|\mathcal{G}_1(\theta_h \partial_x \pi) - \mathcal{G}_1(\theta_h \partial_x q)\|_{2;M}^2 + \left(\frac{h_y}{h_x}\right)^{p'-2} h_y^2 \|\mathcal{G}_2(\theta_h \partial_y \pi) - \mathcal{G}_2(\theta_h \partial_y q)\|_{2;M}^2 \right)$$

$$\sim \alpha_0 \sum_{M \in \mathbb{M}} \left(h_x^2 \int_M \varphi^*_{\tau + |\theta_h \partial_x \pi|}(|\theta_h \partial_x (\pi - q)|) \, dx + \left(\frac{h_y}{h_x}\right)^{p'-2} h_y^2 \int_M \varphi^*_{\tau_2 + |\theta_h \partial_y \pi|}(|\theta_h \partial_y (\pi - q)|) \, dx \right),$$

where φ is the complementary N-function. The constants depend on p and τ.

Proof. We have

$$s(\pi, \pi - q) - s(q, \pi - q) = \alpha_0 \sum_{M \in \mathbb{M}_h} (M_{s,11}(\pi)\theta_h \partial_x \pi - M_{s,11}(q)\theta_h \partial_x q, \theta_h \partial_x(\pi - q))$$

$$+ \alpha_0 \sum_{M \in \mathbb{M}_h} (M_{s,22}(\pi)\theta_h \partial_y \pi - M_{s,22}(q)\theta_h \partial_y q, \theta_h \partial_y(\pi - q)).$$

Rearranging terms, applying a scalar-valued version of Lemma 13 with p, ϵ, $\mathcal{F}$ and φ replaced by p', τ, $\mathcal{G}_1$, and φ^* and remembering the definition of $\mathcal{G}_1$ yields for the x-component, for each $M \in \mathbb{M}_h$,

$$(M_{s,11}(\pi)\theta_h \partial_x \pi - M_{s,11}(q)\theta_h \partial_x q, \theta_h \partial_x(\pi - q))_M$$

$$= \left(\frac{1}{\tau}\right)^{p'-2} h_x^2 ((\tau + |\theta_h \partial_x \pi|)^{p'-2}\theta_h \partial_x \pi - (\tau + |\theta_h \partial_x q|)^{p'-2}\theta_h \partial_x q, \theta_h \partial_x(\pi - q))_M$$

$$\sim h_x^2 \|\mathcal{G}_1(\theta_h \partial_x \pi) - \mathcal{G}_1(\theta_h \partial_x q)\|_{2;M}^2$$

$$\sim h_x^2 \int_M \varphi^*_{\tau + |\theta_h \partial_x \pi|}(|\theta_h \partial_x \pi - \theta_h \partial_x q|)\, dx.$$

We use the same approach to bound the y-component, but first we rewrite the factor $M_{s,22}(\pi)$ so that

$$\left(\tau + \frac{h_y}{h_x}|\theta_h \partial_y \pi|\right)^{p'-2} = \left(\frac{h_y}{h_x}\right)^{p'-2}(\tau_2 + |\theta_h \partial_y \pi|)^{p'-2}.$$

The y-component is then

$$(M_{s,22}(\pi)\theta_h \partial_y \pi - M_{s,22}(q)\theta_h \partial_y q, \theta_h \partial_y(\pi - q))_M$$

$$= \left(\frac{1}{\tau}\right)^{p'-2} h_y^2 \left(\frac{h_y}{h_x}\right)^{p'-2}((\tau_2 + |\theta_h \partial_y \pi|)^{p'-2}\theta_h \partial_y \pi - (\tau_2 + |\theta_h \partial_y q|)^{p'-2}\theta_h \partial_y q, \theta_h \partial_y(\pi - q))_M$$

$$\sim h_y^2 \left(\frac{h_y}{h_x}\right)^{p'-2}\|\mathcal{G}_2(\theta_h \partial_y \pi) - \mathcal{G}_2(\theta_h \partial_y q)\|_{2;M}^2$$

$$\sim h_y^2 \left(\frac{h_y}{h_x}\right)^{p'-2}\int_M \varphi^*_{\tau_2 + |\theta_h \partial_y \pi|}(|\theta_h \partial_y \pi - \theta_h \partial_y q|)\, dx. \qquad \square$$

Lemma 4. *For all $\delta > 0$ there exists a constant $c = c(\delta, p, \alpha_0, \tau)$ such that for all $\pi, q \in W^{1,p'}(\Omega)$,*

$$|s(\pi, \pi - q)| \lesssim \alpha_0 c_\delta h_x^2 \|\tau + |\partial_x \pi|\|_{p';\Omega}^{p'} + \alpha_0 c_\delta h_y^2 \left(\frac{h_y}{h_x}\right)^{p'-2}\|\tau_2 + |\partial_y \pi|\|_{p';\Omega}^{p'}$$

$$+ \alpha_0 \delta \sum_{M \in \mathbb{M}_h} \left(h_x^2 \|\mathcal{G}_1(\theta_h \partial_x \pi) - \mathcal{G}_1(\theta_h \partial_x q)\|_{2;M}^2 \right.$$

$$\left. + h_y^2 \left(\frac{h_y}{h_x}\right)^{p'-2}\|\mathcal{G}_2(\theta_h \partial_y \pi) - \mathcal{G}_2(\theta_h \partial_y q)\|_{2;M}^2\right).$$

Proof. Again rewriting

$$\left(\tau + \frac{h_y}{h_x}|\theta_h \partial_y \pi|\right)^{p'-2} = \left(\frac{h_y}{h_x}\right)^{p'-2}(\tau_2 + |\theta_h \partial_y \pi|)^{p'-2},$$

working component-wise and applying a scalar-valued version of Lemma 13 with p, φ, and ϵ replaced by p', φ^*, and τ, we have

$$|s(\pi, \pi - q)| \lesssim \alpha_0 \sum_{M \in \mathbb{M}} h_x^2 \int_M (\varphi^*)'_{\tau + \theta_h \partial_x \pi}(|\theta_h \partial_x \pi|)|\theta_h \partial_x \pi - \theta_h \partial_x q|\, dx$$

$$+ \left(\frac{h_y}{h_x}\right)^{p'-2} h_y^2 \int_M (\varphi^*)'_{\tau_2 + \theta_h \partial_y \pi}(|\theta_h \partial_y \pi|)|\theta_h \partial_y \pi - \theta_h \partial_y q|\, dx.$$

Now we apply Young's inequality for N-functions from Lemma 14. For the x-component this gives

$$\sum_{M \in \mathbb{M}} h_x^2 \int_M (\varphi^*)'_{\tau + \theta_h \partial_x \pi}(|\theta_h \partial_x \pi|)|\theta_h \partial_x \pi - \theta_h \partial_x q|\, dx$$

$$\lesssim c_\delta h_x^2 \int_\Omega \varphi^*_{\tau + \theta_h \partial_x \pi}(|\theta_h \partial_x \pi|)\, dx + \delta \sum_{M \in \mathbb{M}} h_x^2 \|\mathcal{G}_1(\theta_h \partial_x \pi) - \mathcal{G}_1(\theta_h \partial_x q)\|_{2;M}^2.$$

The integral on the left-hand side can be bounded with Lemma 3, and Lemma 13 (change of shift), and using that $\varphi^*(t) := \frac{t^{p'}}{p'}$, and lastly the stability of the fluctuation operator,

$$\int_\Omega \varphi^*_{\tau+\theta_h\partial_x\pi}(|\theta_h\partial_x\pi|)\,dx \lesssim \int_\Omega \varphi^*(\tau+|\theta_h\partial_x\pi|)\,dx \lesssim \|\tau+|\partial_x\pi|\|^{p'}_{p';\Omega}.$$

Similarly, for the *y*-component,

$$\sum_{M\in\mathbb{M}}\left(\frac{h_y}{h_x}\right)^{p'-2}h_y^2\int_M(\varphi^*)'_{\tau_2+\theta_h\partial_y\pi}(|\theta_h\partial_y\pi|)|\theta_h\partial_y\pi-\theta_h\partial_yq|\,dx$$

$$\lesssim c_\delta\left(\frac{h_y}{h_x}\right)^{p'-2}h_y^2\int_\Omega(\varphi^*)_{\tau_2+\theta_h\partial_y\pi}(|\theta_h\partial_y\pi|)\,dx + \delta\sum_{M\in\mathbb{M}}\left(\frac{h_y}{h_x}\right)^{p'-2}h_y^2\|\mathcal{G}_2(\theta_h\partial_y\pi)-\mathcal{G}_2(\theta_h\partial_yq)\|^2_{2;M}$$

and

$$\int_\Omega(\varphi^*)_{\tau_2+\theta_h\partial_y\pi}(|\theta_h\partial_y\pi|)\,dx \lesssim \int_\Omega(\varphi^*)(\tau_2+|\theta_h\partial_y\pi|)\,dx \lesssim \|\tau_2+|\partial_y\pi|\|^{p'}_{p';\Omega}. \qquad \square$$

3 Error Analysis

3.1 Modified Interpolation Operator

3.1.1 Standard Properties in an Anisotropic *p*-Stokes Setting

The idea behind the error analysis of LPS-methods is to construct an interpolation operator $j_h : W^{1,p}(\Omega) \to X_h$ (and $\mathbf{j}_h : \mathbf{W}_0^{1,\nu}(\Omega) \to \mathcal{X}_h$) so that the interpolation error is orthogonal to Y_h (or $\mathbf{Y}_h := [Y_h]^d$). This interpolation operator should also fulfill some useful approximation and stability properties. The following lemma states these standard LPS properties in an anisotropic *p*-Stokes setting. The parameter ν will in practice be equal either p or p'. For brevity of presentation, we only present the properties for the scalar j_h but it also extends to $\mathbf{j}_h$ and the vector-valued case.

Lemma 5. *Let $\nu \geq 1$ and let the two spaces X_h and Y_h satisfy Lemma 1. Then there exist an interpolation operator $j_h : W^{1,\nu}(\Omega) \to X_h$ satisfying the following properties:*
(i) *Orthogonality with respect to Y_h: For all $\omega \in W^{1,\nu}(\Omega)$,*

$$(\omega - j_h\omega, q_h)_\Omega = 0 \quad \text{for all } q_h \in Y_h.$$

(ii) *Approximation properties: For all $M \in \mathbb{M}_h$ and all $\omega \in W^{1,\nu}(\Omega)$,*

$$\|\omega - j_h\omega\|_{0,\nu;M} \lesssim h_x\|\partial_x\omega\|_{0,\nu;\Lambda(M)} + h_y\|\partial_y\omega\|_{0,\nu;\Lambda(M)}, \tag{3.1}$$

$$\|\nabla(\omega - j_h\omega)\|_{\nu,M} \lesssim \|\omega\|_{1,\nu;\Lambda(M)} + \frac{h_x}{h_y}\|\partial_x\omega\|_{0,\nu;\Lambda(M)} + \|\partial_y\omega\|_{0,\nu;\Lambda(M)}, \tag{3.2}$$

where $\Lambda(M) = \text{int}\bigcup\{\overline{M'} : M' \in \mathbb{M}_h, \overline{M}\cap\overline{M'} \neq \emptyset\}$ is a neighborhood of the patch M. For all $\omega \in W^{2,\nu}(\Omega)$,

$$\|\omega - j_h\omega\|_{0,\nu;M} \lesssim h_x^2\|\partial_x^2\omega\|_{0,\nu;\Lambda(M)} + h_y^2\|\partial_y^2\omega\|_{0,\nu;\Lambda(M)} + h_xh_y\|\partial_{xy}\omega\|_{0,\nu;\Lambda(M)}, \tag{3.3}$$

$$\|\nabla(\omega - j_h\omega)\|_{\nu,M} \lesssim h_x\|\partial_x\omega\|_{1,\nu;\Lambda(M)} + h_y\|\partial_y\omega\|_{1,\nu;\Lambda(M)} + \frac{h_x^2}{h_y}\|\partial_x^2\omega\|_{0,\nu;\Lambda(M)}. \tag{3.4}$$

The constants depend on the inf-sup constant β_{loc} of Lemma 1.
(iii) *Stability properties: For all $\omega \in W^{1,\nu}(\Omega)$,*

$$\|\nabla j_h\omega\|_{\nu,M} \lesssim \left(1+\frac{h_x}{h_y}\right)\|\nabla\omega\|_{\nu,M}, \tag{3.5}$$

$$\|\partial_x j_h\omega\|_{\nu,M} \lesssim \|\partial_x\omega\|_{\nu;\Lambda(M)} + h_x^{-1}h_y\|\partial_y\omega\|_{\nu,\Lambda(M)}, \tag{3.6}$$

$$\|\partial_y j_h\omega\|_{\nu,M} \lesssim \|\partial_y\omega\|_{\nu;\Lambda(M)} + h_y^{-1}h_x\|\partial_x\omega\|_{\nu,\Lambda(M)}. \tag{3.7}$$

Proof. The interpolation operator j_h (as well as $\mathbf{j}_h$) is constructed from two parts, so that $j_h\omega|_M = i_h\omega + z_h(\omega)$. The base interpolator i_h ensures approximation properties, and $z_h(\omega)$ is added to ensure the orthogonality property. Hence, $\Lambda(M)$ is a patch of patches similarly defined as for the anisotropic Scott–Zhang interpolator in [2].

On isotropic meshes, the Scott–Zhang interpolation is a possible choice for i_h, see e.g. [18] and [21]. On anisotropic elements, we instead use the anisotropic Scott–Zhang-type interpolators described in [2, 4, 8] which all coincide for the non-sheared elements that we study here. The anisotropic Scott–Zhang interpolation operator i_h is constructed by integrating only along the long edges in the x-direction, instead of along any edge as is the case for the original Scott–Zhang interpolation operator. Various approximation and stability properties of i_h have been shown in [2, 4, 8] and we choose the, for our purpose, most useful ones (see Appendix).

For z_h we know that due to Lemma 1 there exists a unique $z_h(\omega) \in Z_h(M)^\perp$ which ensures that

$$(\omega - i_h\omega - z_h(\omega), q_h)_M = 0$$

(where $Z_h(M) := \{\omega_h \in X_h^0(M) : (\omega_h, q_h)_M = 0 \text{ for all } q_h \in Y_h(M)\}$), see [16, 18, 21]. This $z_h(\omega)$ also fulfills

$$\|z_h(\omega)\|_{v;M} \leq \frac{1}{\beta_{\mathrm{loc}}}\|\omega - i_h\omega\|_{v;M}. \tag{3.8}$$

Here β_{loc} is the inf-sup constant from Lemma 1. Using the triangle inequality, the approximation property of z_h (3.8) and the approximation properties of i_h, we arrive at the approximation properties of j_h in the L^v-norm, i.e. (3.1) and (3.3). To bound the gradient of the interpolation error we need to bound the gradient of z_h. We do this by combining (3.8) with the inverse inequality from [9], which yields

$$\|\partial_x z_h(\omega)\|_{v,M} \lesssim \frac{1}{h_x}\|z_h(\omega)\|_{v,M} \lesssim \frac{1}{\beta_{\mathrm{loc}} h_x}\|\omega - i_h\omega\|_{v,M}, \tag{3.9}$$

$$\|\partial_y z_h(\omega)\|_{v,M} \lesssim \frac{1}{h_y}\|z_h(\omega)\|_{v,M} \lesssim \frac{1}{\beta_{\mathrm{loc}} h_y}\|\omega - i_h\omega\|_{v,M}. \tag{3.10}$$

Again combining with the triangle inequality

$$\|\nabla(\omega - j_h\omega)\|_{v,M} \leq \|\nabla(\omega - i_h\omega)\|_{v,M} + \|\nabla z_h(\omega)\|_{v,M}$$

and the approximation properties of i_h, we have the desired estimates (3.2) and (3.4). Using that

$$\|j_h\omega\|_{v,M} = \|j_h\omega + \omega - \omega\|_{v,M} \leq \|\omega\|_{v,M} + \|\omega - j_h\omega\|_{v,M}$$

together with the approximation property (3.1), we have also L^v-stability.

For $W^{1,v}$-stability we use the stability of the base interpolator i_h of equation (C.2) and (C.3), or alternatively (C.4) and (C.5) together with the inverse estimates (3.9) and (3.10). □

Given j_h, it is easy to construct a corresponding vector valued interpolation operator $\mathbf{j}_h : \mathbf{W}_0^{1,v}(\Omega) \to \mathcal{X}_h$ which renders the interpolation error orthogonal to $\mathbf{Y}_h$.

3.1.2 Approximation Property in Orlicz Spaces on Anisotropic Meshes

The approximation and stability properties of j_h described above are not fundamentally different from those needed in the linear Stokes case. However, for the p-Stokes case, we need an additional approximation property in Orlicz spaces, more specifically in terms of the help function $\mathcal{F}$. On isotropic meshes such an interpolation operator was shown to exist in [18]. Here we construct operators $j_h, \mathbf{j}_h$ suitable for anisotropic meshes.

Lemma 6. *Let $v \geq 1$, let $\mathcal{X}_h$ and $\mathbf{Y}_h$ satisfy Lemma 1 and let $\mathcal{F}$ be defined by (2.9). If $\mathcal{F}(\mathbf{D}\boldsymbol{\omega}) \in [W^{1,2}(\Omega)]^{d \times d}$, then there exists an interpolation operator $\mathbf{j} : \mathbf{W}_0^{1,v}(\Omega) \to \mathcal{X}_h$ such that for all $M \in \mathbb{M}$,*

$$\|\mathcal{F}(\mathbf{D}\boldsymbol{\omega}) - \mathcal{F}(\mathbf{D}\mathbf{j}_h\boldsymbol{\omega}_h)\|_{2;M} \lesssim \sqrt{\frac{h_x}{h_y}}\left(h_x\|\partial_x\mathcal{F}(\mathbf{D}\boldsymbol{\omega})\|_{2,\Lambda(M)} + h_y\|\partial_y\mathcal{F}(\mathbf{D}\boldsymbol{\omega})\|_{2,\Lambda(M)}\right). \tag{3.11}$$

The remainder of this subsection is dedicated to prove Lemma 6. For this purpose we transfer from working with $\mathcal{F}$, to working in Orlicz–Sobolev spaces with N-functions. We denote classical Orlicz spaces $L^{\varphi}(\Omega)$ and Orlicz–Sobolev spaces with $W^{1,\varphi}(\Omega)$. We have that $f \in L^{\varphi}(\Omega)$ if and only if $\int_{\Omega} \varphi(|f|)\,dx < \infty$ and $f \in W^{1,\varphi(\Omega)}$ if and only if $\nabla^i f \in L^{\varphi}(\Omega)$ for $0 \le i \le 1$. We then need interpolation estimates in terms of N-functions, which requires *Orlicz-stability*, *Orlicz-approximability* and *Orlicz-continuity*. For the isotropic case, these properties are shown in [14], and we here extend them to anisotropic meshes (for low-order approximations).

Orlicz-Stability. We start by proving Orlicz-stability. We first need to show that j_h fulfills an auxiliary stability property corresponding to [14, Assumption 4.1].

Lemma 7. *The interpolation operator j_h satisfies the following $W^{1,1}$-stability:*

$$\fint_M |\mathbf{j}_h \boldsymbol{\omega}| + \fint_M h_x |\nabla \mathbf{j}_h \boldsymbol{\omega}| \lesssim \fint_{\Lambda(M)} |\boldsymbol{\omega}| + \frac{h_x}{h_y} h_x \fint_{\Lambda(M)} |\nabla \boldsymbol{\omega}|,$$

where $\fint_U \omega := \frac{1}{|U|} \int_U \omega$ denotes the mean value over U.

Proof. The interpolation estimate (3.1) with $\nu = 1$ bounds the first term

$$\fint_M |\mathbf{j}_h \boldsymbol{\omega}| \le \fint_M |\mathbf{j}_h \boldsymbol{\omega} - \boldsymbol{\omega}| + \fint_M |\boldsymbol{\omega}| \lesssim \left(1 + \frac{1}{\beta_{\mathrm{loc}}}\right) \fint_{\Lambda(M)} \sum_{|\alpha|=1} h^{\alpha} |D^{\alpha}\boldsymbol{\omega}| + \fint_M |\boldsymbol{\omega}| \lesssim \sum_{|\alpha|\le 1} \fint_{\Lambda(M)} h^{\alpha} |D^{\alpha}\boldsymbol{\omega}|. \qquad (3.12)$$

Here, α is a multi-index. For the gradient terms we use the directional stability estimates (3.6) and (3.7) to get

$$\fint_M |\partial_x \mathbf{j}_h \boldsymbol{\omega}| + \fint_M |\partial_y \mathbf{j}_h \boldsymbol{\omega}| \lesssim \fint_{\Lambda(M)} |\partial_x \boldsymbol{\omega}| + \frac{h_y}{h_x} \fint_{\Lambda(M)} |\partial_y \boldsymbol{\omega}| + \frac{h_x}{h_y} \fint_{\Lambda(M)} |\partial_x \boldsymbol{\omega}| + \fint_{\Lambda(M)} |\partial_y \boldsymbol{\omega}|$$

$$= \left(1 + \frac{h_x}{h_y}\right) \fint_{\Lambda(M)} |\partial_x \boldsymbol{\omega}| + \left(1 + \frac{h_y}{h_x}\right) \fint_{\Lambda(M)} |\partial_y \boldsymbol{\omega}|. \qquad (3.13)$$

Since M is as anisotropic as $\Lambda(M)$, combining (3.12) and (3.13) we have

$$\fint_M |\mathbf{j}_h \boldsymbol{\omega}| + \fint_M h_x |\partial_x \mathbf{j}_h \boldsymbol{\omega}| + \fint_M h_y |\partial_y \mathbf{j}_h \boldsymbol{\omega}| \lesssim \fint_{\Lambda(M)} |\boldsymbol{\omega}| + \left(h_x + \frac{h_x^2}{h_y}\right) \fint_{\Lambda(M)} |\partial_x \boldsymbol{\omega}| + \left(h_y + \frac{h_y^2}{h_x}\right) \fint_{\Lambda(M)} |\partial_y \boldsymbol{\omega}|.$$

Rearranging and using that $\frac{h_x}{h_y} \ge h_x \ge h_y$ completes the proof. $\qquad \square$

Lemma 8 (First-Order Anisotropic Orlicz-Stability). *For a j_h that satisfies Lemma 7 it holds*

$$\fint_M \varphi(|\mathbf{j}_h \boldsymbol{\omega}|) + \fint_M \varphi(h_x |\nabla \mathbf{j}_h \boldsymbol{\omega}|) \lesssim \fint_{\Lambda(M)} \varphi(|\boldsymbol{\omega}|) + \fint_{\Lambda(M)} \varphi\left(\frac{h_x^2}{h_y} |\nabla \boldsymbol{\omega}|\right).$$

Proof. Since it holds that

$$|\boldsymbol{\omega}_h(x)| \lesssim \fint_K |\boldsymbol{\omega}_h(y)|, \quad |\nabla \boldsymbol{\omega}_h(x)| \lesssim \fint_K |\nabla \boldsymbol{\omega}_h(y)|,$$

we have

$$\fint_M \varphi(|\mathbf{j}_h \boldsymbol{\omega}|) + \fint_M \varphi(h_x |\nabla \mathbf{j}_h \boldsymbol{\omega}|) \lesssim \fint_M \varphi\left(\fint_M |\mathbf{j}_h \boldsymbol{\omega}|\right) + \fint_M \varphi\left(\fint_M h_x |\nabla \mathbf{j}_h \boldsymbol{\omega}|\right).$$

Using Lemma 7, the convexity of φ to exchange the order of summation and integration, the Δ_2-condition to move constants, and Jensen's inequality for integrals, we get

$$\fint_M \varphi(|\mathbf{j}_h \boldsymbol{\omega}|) + \fint_M \varphi(h_x |\nabla \mathbf{j}_h \boldsymbol{\omega}|) \lesssim \fint_M \varphi\left(\fint_{\Lambda(M)} |\boldsymbol{\omega}| + \fint_{\Lambda(M)} \frac{h_x^2}{h_y} |\nabla \boldsymbol{\omega}|\right) \lesssim c \fint_M \varphi\left(\fint_{\Lambda(M)} |\boldsymbol{\omega}|\right) + \fint_M \varphi\left(\fint_{\Lambda(M)} \frac{h_x^2}{h_y} |\nabla \boldsymbol{\omega}|\right)$$

$$\lesssim c \fint_{\Lambda(M)} \varphi(|\boldsymbol{\omega}|) + \fint_{\Lambda(M)} \varphi\left(\frac{h_x^2}{h_y} |\nabla \boldsymbol{\omega}|\right). \qquad \square$$

Orlicz-Approximability. To prove Orlicz-approximability, we need the Orlicz-stability proven above, but also first-order, anisotropic versions of [14, Lemma 3.1, Theorem 3.2 and Corollary 3.3], collected in the following lemma.

Lemma 9. *For all $\boldsymbol{\omega} \in (W^{1,\varphi}(K))^d$ there exists $\mathbf{Q}\boldsymbol{\omega} \in ((\mathcal{P}_0)(K))^d$ such that*

$$\fint_K \varphi(|\boldsymbol{\omega} - \mathbf{Q}\boldsymbol{\omega}|) + \fint_K \varphi(h_x|\nabla(\boldsymbol{\omega} - \mathbf{Q}\boldsymbol{\omega})|) \lesssim \fint_K \varphi(h_x|\nabla\boldsymbol{\omega}|).$$

It also holds on patches M and neighborhoods of elements and patches $\Lambda(K), \Lambda(M)$.

Proof. The bound of the second term is trivial since $\mathbf{Q}\mathbf{v} \in ((\mathcal{P}_0)(K))^N$ so that $\nabla(\mathbf{Q}\mathbf{v}) = 0$. In the isotropic setting in [14] the first term is bounded by setting $\mathbf{Q}\mathbf{v}$ as an averaged Taylor polynomial, but on anisotropic elements this will result in a constant which depends heavily on the aspect ratio, see [11]. As we do not have higher-order derivatives (as in [14]) we can instead simply use that $\mathbf{Q}\mathbf{v}$ is a constant, and apply an anisotropic version of the Poincaré-Wirtinger inequality (see e.g. [8, Lemma 2]), which states that there exists a constant c so that

$$\|\mathbf{v} - \mathbf{c}\|_\nu \lesssim h_x\|\partial_x\mathbf{v}\|_\nu + h_y\|\partial_y\mathbf{v}\|_\nu, \tag{3.14}$$

for $\nu \geq 1$. The first term is thus bounded by

$$\fint_K \varphi(|\mathbf{v} - \mathbf{Q}\mathbf{v}|) \lesssim \fint_K \varphi(h_x|\partial_x\mathbf{v}| + h_y|\partial_y\mathbf{v}|) \lesssim \fint_K \varphi(h_x|\nabla\mathbf{v}|). \qquad \square$$

Given the above lemma and Orlicz-stability, Orlicz-approximability follows.

Lemma 10 (Anisotropic Orlicz-Approximability).

$$\fint_M \varphi(|\boldsymbol{\omega} - \mathbf{j}_h\boldsymbol{\omega}|) + \fint_M \varphi(h_x|\nabla(\boldsymbol{\omega} - \mathbf{j}_h\boldsymbol{\omega})|) \lesssim \fint_{\Lambda(M)} \varphi\Big(\frac{h_x^2}{h_y}|\nabla\boldsymbol{\omega}|\Big) \quad \text{for all } \boldsymbol{\omega} \in (W^{1,\varphi}(M))^N.$$

Proof. Adding and subtracting a $\mathfrak{p} \in V_h$, using the convexity of φ, the Δ_2-condition and that $\mathbf{j}_h$ is a projection, we have

$$\text{LHS} := \fint_M \varphi(|\boldsymbol{\omega} - \mathbf{j}_h\boldsymbol{\omega}|) + \fint_M \varphi(h_x|\nabla(\boldsymbol{\omega} - \mathbf{j}_h\boldsymbol{\omega})|)$$

$$= \fint_M \varphi(|\boldsymbol{\omega} - \mathfrak{p} + \mathfrak{p} - \mathbf{j}_h\boldsymbol{\omega}|) + \fint_M \varphi(h_x|\nabla(\boldsymbol{\omega} - \mathfrak{p} + \mathfrak{p} - \mathbf{j}_h\boldsymbol{\omega})|)$$

$$\lesssim \fint_M \varphi(|\boldsymbol{\omega} - \mathfrak{p}|) + \fint_M \varphi(|\mathfrak{p} - \mathbf{j}_h\boldsymbol{\omega}|) + \fint_M \varphi(h_x|\nabla(\boldsymbol{\omega} - \mathfrak{p})|) + \fint_M \varphi(h_x|\nabla(\mathfrak{p} - \mathbf{j}_h\boldsymbol{\omega})|).$$

Applying Lemma 8 (Orlicz-stability) on the second and fourth term with $\boldsymbol{\omega}$ replaced with $\boldsymbol{\omega} - \mathfrak{p}$, and then using Lemma 9 ($\mathfrak{p}$ is arbitrary), we get

$$\text{LHS} \lesssim \fint_M \varphi(h_x|\nabla\boldsymbol{\omega}|) + \fint_{\Lambda(M)} \varphi(h_x|\nabla\boldsymbol{\omega}|) + \fint_M \varphi(h_x|\nabla\boldsymbol{\omega}|) + \fint_{\Lambda(M)} \varphi\Big(\frac{h_x^2}{h_y}|\nabla\boldsymbol{\omega}|\Big).$$

Rearranging, using that $\frac{h_x^2}{h_y} > h_x$ and the Δ_2-condition, and shifting all integration areas from M to $\Lambda(M)$, we get the desired result. $\qquad \square$

Corollary 1 (Orlicz-Continuity). *For all $\boldsymbol{\omega} \in (W^{1,\varphi}(M))^N$,*

$$\fint_M \varphi(h_x|\nabla\mathbf{j}_h\boldsymbol{\omega}|) \lesssim \fint_{\Lambda(M)} \varphi\Big(\frac{h_x^2}{h_y}|\nabla\boldsymbol{\omega}|\Big).$$

Proof. By ignoring the low-order term on the left-hand side in Lemma 10 and using the triangle equality and convexity of φ, we have the result. $\qquad \square$

Corollary 2. *For all* $\mathbf{v} \in (W^{1,\varphi}(M))^N$,

$$\fint_M \varphi_a(|\nabla \mathbf{j}_h \mathbf{v}|) \lesssim \frac{h_x}{h_y} \fint_{\Lambda(M)} \varphi_a(|\nabla \mathbf{v}|).$$

Proof. The result can be derived from Corollary 1 by choosing $\mathbf{v} = h_x \boldsymbol{\omega}$ and using the Δ_2-condition. $\square$

Orlicz-Continuity and Orlicz-Approximability in Terms of $\mathcal{F}$ We derive the approximation property of j_h in terms of $\mathcal{F}$, i.e. we prove Lemma 6. The proof mostly follows the (lengthy) isotropic proof of [18, Lemma 4.4]. Here we only give details on the parts that differs, i.e. when the anisotropic Orlicz-continuity and Poincaré's inequality is used.

Proof of Lemma 6. We start by transferring from $\mathcal{F}$ to working with φ (see [18, p. 66]) for details on this step,

$$\fint_M |\mathcal{F}(\mathbf{D}\boldsymbol{\omega}) - \mathcal{F}(\mathbf{D}\mathbf{j}_h\boldsymbol{\omega})|^2\, d\mathbf{x} \lesssim \underbrace{\fint_M \varphi_{\epsilon+|\mathbf{Dq}|}(|\mathbf{Dj}_h(\mathbf{q}-\boldsymbol{\omega})|)}_{=:T_1} + \underbrace{\fint_M \varphi_{\epsilon+|\mathbf{Dq}|}(|\mathbf{D}\boldsymbol{\omega} - \mathbf{Dq}|)}_{=:T_2},$$

where $\mathbf{q}$ is arbitrary linear polynomial. The first term is bounded using Corollary 2

$$T_1 \lesssim \fint_M \varphi_{\epsilon+|\mathbf{Dq}|}(|\mathbf{Dj}_h(\mathbf{q}-\boldsymbol{\omega})|) \lesssim \frac{h_x}{h_y} \fint_{\Lambda(M)} \varphi_{\epsilon+|\mathbf{Dq}|}(|\nabla(\mathbf{q}-\boldsymbol{\omega})|).$$

Choosing $\nabla q := \fint_{\Lambda(M)} \nabla \boldsymbol{\omega}$, and using a version of Korn's inequality for N-functions (see [15, Theorem 6.13]), we have

$$\frac{h_x}{h_y} \fint_{\Lambda(M)} \varphi_{\epsilon+|\mathbf{Dq}|}(|\nabla(\mathbf{q}-\boldsymbol{\omega})|) = \frac{h_x}{h_y} \fint_{\Lambda(M)} \varphi_{\epsilon+|\mathbf{Dq}|}\left(\left|\nabla\boldsymbol{\omega} - \fint_{\Lambda(M)}\nabla\boldsymbol{\omega}\right|\right)$$

$$\lesssim \frac{h_x}{h_y} \fint_{\Lambda(M)} \varphi_{\epsilon+|\mathbf{Dq}|}\left(\left|\mathbf{D}\boldsymbol{\omega} - \fint_{\Lambda(M)}\mathbf{D}\boldsymbol{\omega}\right|\right).$$

The second term T_2 is also bounded by $\fint_{\Lambda(M)} \varphi_{\epsilon+|\mathbf{Dq}|}(|\mathbf{D}\boldsymbol{\omega} - \fint_{\Lambda(M)}\mathbf{D}\boldsymbol{\omega}|)$ since $D\mathbf{q} = \fint_{\Lambda(M)}\mathbf{D}\boldsymbol{\omega}$. Adding $T_1 + T_2$ and using the relation between φ and $\mathcal{F}$ (Lemma 13), we have

$$\fint_M |\mathcal{F}(\mathbf{D}\boldsymbol{\omega}) - \mathcal{F}(\mathbf{D}\mathbf{j}_h\boldsymbol{\omega})|^2 \lesssim \frac{h_x}{h_y} \fint_{\Lambda(M)} \varphi_{\epsilon+|\mathbf{Dq}|}\left(\left|\mathbf{D}\boldsymbol{\omega} - \fint_{\Lambda(M)}\mathbf{D}\boldsymbol{\omega}\right|\right)$$

$$\lesssim \frac{h_x}{h_y} \fint_{\Lambda(M)} \left|\mathcal{F}(\mathbf{D}\boldsymbol{\omega}) - \mathcal{F}\left(\fint_{\Lambda(M)}\mathbf{D}\boldsymbol{\omega}\right)\right|^2. \tag{3.15}$$

Furthermore, it is possible to show that

$$\fint_{\Lambda(M)} \left|\mathcal{F}(\mathbf{D}\boldsymbol{\omega}) - \mathcal{F}\left(\fint_{\Lambda(M)}\mathbf{D}\boldsymbol{\omega}\right)\right|^2 \lesssim \fint_{\Lambda(M)} \left|\mathcal{F}(\mathbf{D}\boldsymbol{\omega}) - \fint_{\Lambda(M)}\mathcal{F}(\mathbf{D}\boldsymbol{\omega})\right|^2, \tag{3.16}$$

see details on this step on [18, pp. 66–67] (the anisotropy of the elements does not alter these steps). Inserting (3.16) into (3.15), we get

$$\fint_M |\mathcal{F}(\mathbf{D}\boldsymbol{\omega}) - \mathcal{F}(\mathbf{D}\mathbf{j}_h\boldsymbol{\omega})|^2 \lesssim \frac{h_x}{h_y} \fint_{\Lambda(M)} \left|\mathcal{F}(\mathbf{D}\boldsymbol{\omega}) - \fint_{\Lambda(M)}\mathcal{F}(\mathbf{D}\boldsymbol{\omega})\right|^2,$$

or equivalently since $\Lambda(M)$ and M are equally anisotropic,

$$\|\mathcal{F}(\mathbf{D}\boldsymbol{\omega}) - \mathcal{F}(\mathbf{D}\mathbf{j}_h\boldsymbol{\omega})\|_2^2 \lesssim \frac{h_x}{h_y} \left\|\mathcal{F}(\mathbf{D}\boldsymbol{\omega}) - \fint_{\Lambda(M)}\mathcal{F}(\mathbf{D}\boldsymbol{\omega})\right\|_2^2.$$

Note that $\fint_{\Lambda(M)}\mathcal{F}(\mathbf{D}\boldsymbol{\omega})$ is a patch-wise constant. Thus the anisotropic Poincaré–Wirtinger inequality (3.14) applies and we have the desired result. $\square$

3.2 Well-Posedness

In this subsection we show that there exists a (unique) solution to the discrete problem (2.3), which is bounded by the data of the problem. In order to do so, we first show that a discrete inf-sup condition holds.

Lemma 11 (Discrete inf-sup condition). *For all $q_h \in \mathcal{Q}_h$ and all $v' \in]2, \infty[$ it holds that*

$$\beta_c \|q_h\|_{v';\Omega} \le \frac{h_x}{h_y} \left(\sup_{\boldsymbol{\omega}_h \in \mathbf{X}_h} \frac{(\nabla \cdot \boldsymbol{\omega}_h, q_h)_\Omega}{\|\nabla \boldsymbol{\omega_h}\|_{v;\Omega}} + \sum_{M \in \mathbb{M}} \sum_{|\alpha|=1} h^\alpha \|\theta_h(D^\alpha q_h)\|_{v',M} \right), \tag{3.17}$$

where β_c is the constant of a continuous inf-sup condition satisfied by the spaces $\mathcal{X}$ and $\mathcal{Q}$.

Proof. We know that the pair $\mathcal{X} \times \mathcal{Q}$ satisfies the continuous inf-sup condition (see e.g. [18]), i.e. there exists a positive constant β_c so that

$$\beta_c \|q\|_{v';\Omega} \le \sup_{\boldsymbol{\omega} \in \mathbf{W}_0^{1,v}(\Omega)} \frac{(\nabla \cdot \boldsymbol{\omega}, q)_\Omega}{\|\nabla \boldsymbol{\omega}\|_{v;\Omega}} \quad \text{for all } q \in L_0^{v'}(\Omega),$$

and since $\mathcal{Q}_h \subset L_0^{v'}(\Omega)$, this holds also for all $q_h \in \mathcal{Q}_h$. Now we split the right-hand side into two terms by adding and subtracting $\mathbf{j}_h \boldsymbol{\omega}$. The first term is then multiplied by $\frac{\|\nabla \mathbf{j}_h \boldsymbol{\omega}\|}{\|\nabla \mathbf{j}_h \boldsymbol{\omega}\|} = 1$ to get

$$\beta_c \|q_h\|_{v'} \le \sup_{\boldsymbol{\omega} \in \mathbf{W}_0^{1,v}(\Omega)} \frac{(\nabla \cdot \mathbf{j_h} \boldsymbol{\omega}, q_h)_\Omega \|\nabla \mathbf{j_h} \boldsymbol{\omega}\|_v}{\|\nabla \mathbf{j_h} \boldsymbol{\omega}\|_v \|\nabla \boldsymbol{\omega}\|_v} + \sup_{\boldsymbol{\omega}_h \in \mathbf{W}_0^{1,p}(\Omega)} \frac{(\nabla \cdot (\boldsymbol{\omega} - \mathbf{j_h} \boldsymbol{\omega}), q_h)_\Omega}{\|\nabla \boldsymbol{\omega}\|_v} \quad \text{for all } q_h \in L_0^{v'}(\Omega).$$

The first term can be bounded by using the stability property of the interpolation operator (3.5):

$$\sup_{\boldsymbol{\omega} \in \mathbf{W}_0^{1,v}(\Omega)} \frac{(\nabla \cdot \mathbf{j_h} \boldsymbol{\omega}, q_h)_\Omega \|\nabla \mathbf{j_h} \boldsymbol{\omega}\|_v}{\|\nabla \mathbf{j_h} \boldsymbol{\omega}\|_v \|\nabla \boldsymbol{\omega}\|_v} \lesssim \frac{h_x}{h_y} \sup_{\boldsymbol{\omega}_h \in \mathbf{X}_h} \frac{(\nabla \cdot \boldsymbol{\omega_h}, q_h)_\Omega}{\|\nabla \boldsymbol{\omega}_h\|_v}.$$

Let us now estimate the nominator of the second term. Integrating by parts (remembering the Dirichlet boundary conditions, and that $\mathbf{j}_h$ preserves boundary conditions) followed by subtracting $P_h \nabla q_h$ from ∇q_h (exploiting the orthogonality property of $\mathbf{j}_h$) yields

$$|(\nabla \cdot (\boldsymbol{\omega} - \mathbf{j_h} \boldsymbol{\omega}), q_h)_\Omega| = |(\boldsymbol{\omega} - \mathbf{j_h} \boldsymbol{\omega}), \nabla q_h)_\Omega| = |((\boldsymbol{\omega} - \mathbf{j_h} \boldsymbol{\omega}), \boldsymbol{\theta}_h(\nabla q_h))_\Omega|.$$

Separation of the components gives

$$|(\nabla \cdot (\boldsymbol{\omega} - \mathbf{j_h} \boldsymbol{\omega}), q_h)_\Omega| \lesssim \sum_{M \in \mathbb{M}} \left(\frac{1}{h_x} \|\omega_x - \mathbf{j}_h \omega_x\|_{v,M} h_x \|\theta_h(\partial_x q_h)\|_{v',M} + \frac{1}{h_y} \|\omega_y - \mathbf{j}_h \omega_y\|_{v,M} h_y \|\theta_h(\partial_y q_h)\|_{v',M} \right).$$

Applying Hölder's inequality and adding extra $\frac{1}{h_x^v} \|\omega_x - j_h \omega_x\|_{v,M}^v$ and $\frac{1}{h_y^v} \|\omega_y - j_h \omega_y\|_{v,M}^v$, we get

$$|(\boldsymbol{\omega} - \mathbf{j_h} \boldsymbol{\omega}), \nabla q_h)_\Omega| \lesssim \left(\sum_{M \in \mathbb{M}} \frac{1}{h_x^v} \|\omega_x - j_h \omega_x\|_{v,M}^v \right)^{\frac{1}{v}} \left(\sum_{M \in \mathbb{M}} h_x^{v'} \|\theta_h(\partial_x q_h)\|_{v',M}^{v'} \right)^{\frac{1}{v'}}$$

$$+ \left(\sum_{M \in \mathbb{M}} \frac{1}{h_y^v} \|\omega_y - j_h \omega_y\|_{v,M}^v \right)^{\frac{1}{v}} \left(\sum_{M \in \mathbb{M}} h_y^{v'} \|\theta_h(\partial_y q_h)\|_{v',M}^{v'} \right)^{\frac{1}{v'}}$$

$$\lesssim \left(\sum_{M \in \mathbb{M}} \frac{1}{h_x^v} \|\omega_x - j_h \omega_x\|_{v,M}^v + \frac{1}{h_y^v} \|\omega_y - j_h \omega_y\|_{v,M}^v \right)^{\frac{1}{v}} \left(\sum_{M \in \mathbb{M}} \sum_{|\alpha|=1} h^\alpha \|\theta_h(D^\alpha q_h)\|_{v',M} \right).$$

Using that $h_x > h_y$, and the approximation properties of the $\mathbf{j}_h$ we can bound the terms in the sum of the first factor on the right-hand side as

$$\frac{1}{h_x^v} \|\omega_x - j_h \omega_x\|_{v,M}^v + \frac{1}{h_y^v} \|\omega_y - j_h \omega_y\|_{v,M}^v \lesssim \frac{h_x^v}{h_y^v} \left(\|\partial_x \boldsymbol{\omega}\|_{v,M} + \|\partial_y \boldsymbol{\omega}\|_{v,M} + \|\partial_x \omega_y\|_{v,M} + \|\partial_y \omega_y\|_{v,M} \right)^v$$

$$\le \left(\frac{h_x}{h_y} \|\nabla \boldsymbol{\omega}\|_{v,M} \right)^v.$$

Dividing by $\|\nabla \boldsymbol{\omega}\|_{v,M}$ completes the proof. □

Lemma 12 (Existence and Stability). *The solution of (2.5) exists, is unique and satisfies*

$$\|\mathbf{v}_h\|_{1,p}^p + \|\mathbf{S}(\mathbf{D}\mathbf{v}_h)\|_{p'}^{p'} + s_h(\pi_h, \pi_h) \le C_1, \tag{3.18}$$

$$\beta_c\|\pi_h\|_{p'} \lesssim \frac{h_x}{h_y}C_2, \tag{3.19}$$

where the constant C_1 depend on Ω, $\mathbf{f}$, p, ϵ, and C_2 additionally on α_0 and τ.

Proof. The proof of existence and uniqueness of the solution is identical to the proof on isotropic meshes. The proof of the velocity estimate (3.18) is not altered although the mesh is anisotropic, and we refer the interested reader to the proof of [19, Lemma 3.6]. To bound the pressure, we start by inserting the discrete equation into the discrete inf-sup condition of Lemma 11:

$$\beta_c\|\pi_h\|_{p'} \lesssim \frac{h_x}{h_y}\sup_{\boldsymbol{\omega_h}\in\mathbf{X}_h}\frac{(\mathbf{S}(\mathbf{D}\mathbf{v}_h),\mathbf{D}\boldsymbol{\omega}_h)_\Omega - (\mathbf{f},\boldsymbol{\omega}_h)_\Omega}{\|\nabla\boldsymbol{\omega_h}\|_p} + \frac{h_x}{h_y}\sum_{M\in\mathbb{M}}\sum_{|\alpha|=1}h^\alpha\|\theta_h(D^\alpha\pi_h)\|_{p',M}. \tag{3.20}$$

We now treat the three terms of (3.20), i.e. the first term and the two components of the second term, separately. Using Hölder's and Poincaré's inequality, that $\|Du\| < C\|\nabla u\|$ and $\|\mathbf{D}\mathbf{v}_h\|_p^p \gtrsim \|\mathbf{S}(\mathbf{D}\mathbf{v}_h)\|_{p'}^{p'}$, we bound the first term by

$$\sup_{\boldsymbol{\omega_h}\in\mathbf{X}_h}\frac{(\mathbf{S}(\mathbf{D}\mathbf{v}_h),\mathbf{D}\boldsymbol{\omega}_h)_\Omega - (\mathbf{f},\boldsymbol{\omega}_h)_\Omega}{\|\nabla\boldsymbol{\omega_h}\|_p} \lesssim \sup_{\boldsymbol{\omega_h}\in\mathbf{X}_h}\left(\frac{\|\mathbf{S}(\mathbf{D}\mathbf{v}_h)\|_{p'}\|\nabla\boldsymbol{\omega}_h\|_p}{\|\nabla\boldsymbol{\omega_h}\|_p} + C(\Omega,p)\frac{\|\mathbf{f}\|_{p'}\|\nabla\boldsymbol{\omega}_h\|_p}{\|\nabla\boldsymbol{\omega_h}\|_p}\right)$$
$$\lesssim \|\mathbf{D}\mathbf{v}_h\|_p^{p-1} + C(\Omega,p)\|\mathbf{f}\|_{p'}.$$

For the x-component of the second term of (3.20) we have considering that $p' > 2$,

$$\sum_{M\in\mathbb{M}}h_x\|\theta_h(\partial_x\pi_h)\|_{p',M} \lesssim \left(\sum_{M\in\mathbb{M}}h_x^{p'}\|\theta_h(\partial_x\pi_h)\|_{p',M}^{p'}\right)^{\frac{1}{p'}}$$
$$\lesssim h_x^{1-\frac{2}{p'}}\left(\underbrace{\sum_{M\in\mathbb{M}}\int_M h_x^2(\tau + |\theta_h(\partial_x\pi_h)|)^{p'-2}|\theta_h(\partial_x\pi_h)|^2}_{x\text{-component of }s(\pi_h,\pi_h)}\right)^{\frac{1}{p'}}.$$

For the y-component we have

$$\left(\sum_{M\in\mathbb{M}}h_y^{p'}\|\theta_h(\partial_y\pi_h)\|_{p',M}^{p'}\right)^{\frac{1}{p'}} \lesssim \left(h_y^{p'-2}\sum_{M\in\mathbb{M}}h_y^2\int_M(\tau_2 + |\theta_h(\partial_y\pi_h)|^{p'-2})|\theta_h(\partial_y\pi_h)|^2\right)^{\frac{1}{p'}}$$
$$\lesssim h_y^{1-\frac{2}{p'}}\left(\frac{h_x}{h_y}\right)^{1-\frac{2}{p'}}\left(\underbrace{\int_M h_y^2\left(\tau + \frac{h_x}{h_y}|\theta_h(\partial_y\pi_h)|^{p'-2}\right)|\theta_h(\partial_y\pi_h)|_{p',M}^2}_{y\text{-component of }s(\pi_h,\pi_h)}\right)^{\frac{1}{p'}}.$$

By combining the three terms while considering that $p' > 2$ and inserting (3.18), we finally arrive at

$$\frac{h_y}{h_x}\beta_c\|\pi_h\|_{p'} \lesssim \|\mathbf{D}\mathbf{v}_h\|_p^{p-1} + \|\mathbf{f}\|_{p'} + s(\pi_h,\pi_h)^{\frac{1}{p'}}.$$

Estimate (3.19) now follows as a consequence of (3.18). $\square$

3.3 A Priori Error Estimate

Theorem 1 (A Priori Error Estimates). *For $p \in\ ]1, 2]$, $\epsilon \in [0, \epsilon_0]$, $\mathcal{F}(\mathbf{D}\mathbf{v}) \in W^{1,2}(\Omega)^{d\times d}$ and $\pi \in W^{1,p'}(\Omega)$ the following a priori error estimates hold:*

$$\|\mathbf{v} - \mathbf{v}_h\|_{1,p} \lesssim \|\mathcal{F}(\mathbf{D}\mathbf{v}) - \mathcal{F}(\mathbf{D}\mathbf{v}_h)\|_2 \le C_v h_x \mathcal{A}^{\frac{p}{2}}, \tag{3.21}$$

$$\|\pi - \pi_h\|_{p'} \le C_p h_x^{\frac{2}{p'}}\mathcal{A}^p, \tag{3.22}$$

with the aspect ratio $\mathcal{A} = \frac{h_x}{h_y} \gg 1$ and $p' = \frac{p}{p-1}$. The constants depend on α_0, p, ϵ_0, Ω, f, τ, $\|\partial_x \pi\|_{p'}$, $\|\partial_y(\pi)\|_{p'}$, $\|\partial_x \mathcal{F}(\mathbf{D}\mathbf{v})\|_2$, $\|\partial_y \mathcal{F}(\mathbf{D}\mathbf{v})\|_2$, and $\|\partial^\alpha v_x\|_p$, $\|\partial^\alpha v_y\|_p$ with $|\alpha| = 2$. In the pressure estimate (3.22) *the constant also depends on β_c.*

Proof. We start with proving the velocity estimate (3.21). By Poincaré's inequality, generalized Korn's inequality (see e.g. [18, equation (2.1)]) and Lemma 13 in the Appendix the velocity error is can be bounded by

$$\|\mathbf{v} - \mathbf{v}_h\|_{1,p} \lesssim \|\mathbf{D}\mathbf{v} - \mathbf{D}\mathbf{v}_h\|_p$$
$$\lesssim \|\mathcal{F}(\mathbf{D}\mathbf{v}) - \mathcal{F}(\mathbf{D}\mathbf{v}_h)\|_2$$
$$\lesssim \|\mathcal{F}(\mathbf{D}\mathbf{v}) - \mathcal{F}(\mathbf{Dj}_h\mathbf{v})\|_2 + \|\mathcal{F}(\mathbf{Dj}_h\mathbf{v}) - \mathcal{F}(\mathbf{D}\mathbf{v}_h)\|_2.$$

The first term of the right-hand side is bounded by the interpolation estimate of Lemma 6. In order to bound the second term involving the projection error, we introduce the following quantity:

$$\mathfrak{E}_{\mathrm{lps}}^2 := \|\mathcal{F}(\mathbf{Dj}_h\mathbf{v}) - \mathcal{F}(\mathbf{D}\mathbf{v}_h)\|_2^2 + \sum_{M \in \mathbb{M}} \left(h_x^2 \|\mathcal{G}_1(\theta_h \partial_x j_h \pi) - \mathcal{G}_1(\theta_h \partial_x \pi_h)\|_{2;M}^2 \right.$$
$$\left. + \left(\frac{h_y}{h_x}\right)^{p'-2} h_y^2 \|\mathcal{G}_2(\theta_h \partial_y j_h \pi) - \mathcal{G}_2(\theta_h \partial_y \pi_h)\|_{2;M}^2 \right),$$

which due to Lemma 4 is equivalent to

$$\mathfrak{E}_{\mathrm{lps}}^2 \sim \|\mathcal{F}(\mathbf{Dj}_h\mathbf{v}) - \mathcal{F}(\mathbf{D}\mathbf{v}_h)\|_2^2 + \underbrace{s(j_h \pi, j_h \pi - \pi_h)}_{=:I_1} - s(\pi_h, j_h \pi - \pi_h). \tag{3.23}$$

Using Lemma 13 and that

$$\mathbf{S}(\mathbf{Dj}_h\mathbf{v}) - \mathbf{S}(\mathbf{D}\mathbf{v}_h) = \mathbf{S}(\mathbf{Dj}_h\mathbf{v}) - \mathbf{S}(\mathbf{D}\mathbf{v}) + \mathbf{S}(\mathbf{D}\mathbf{v}) - \mathbf{S}(\mathbf{D}\mathbf{v}_h)$$

together with Galerkin-orthogonality, we get

$$\|\mathcal{F}(\mathbf{Dj}_h\mathbf{v}) - \mathcal{F}(\mathbf{D}\mathbf{v}_h)\|_2^2 \lesssim (\mathbf{S}(\mathbf{Dj}_h\mathbf{v}) - \mathbf{S}(\mathbf{D}\mathbf{v}_h), \mathbf{D}(j_h\mathbf{v} - \mathbf{v}_h))$$
$$= \underbrace{(\mathbf{S}(\mathbf{Dj}_h\mathbf{v}) - \mathbf{S}(\mathbf{D}\mathbf{v}), \mathbf{D}(j_h\mathbf{v} - \mathbf{v}_h))}_{=:I_2} + (\pi - \pi_h, \nabla \cdot (j_h\mathbf{v} - \mathbf{v}_h))$$
$$+ (\nabla \cdot (\mathbf{v} - \mathbf{v}_h), j_h \pi - \pi_h) + s(\pi_h, j_h \pi - \pi_h).$$

Some algebraic manipulations on the second and third term (exploiting that $\nabla \cdot \mathbf{v} = 0$) and inserting into (3.23) yields

$$\mathfrak{E}_{\mathrm{lps}}^2 \lesssim I_1 + I_2 + \underbrace{(\pi - j_h \pi, \nabla \cdot (j_h\mathbf{v} - \mathbf{v}_h))}_{=:I_3} + \underbrace{(j_h \pi - \pi_h, \nabla \cdot (j_h\mathbf{v} - \mathbf{v}))}_{=:I_4}.$$

We now proceed to bound each of the four terms I_1, I_2, I_3, and I_4.

For the term I_1 we use Lemma 4 and the definition of the $\mathfrak{E}_{\mathrm{lps}}$ (considering that $h_x \geq h_y$ and $p' \geq 2$):

$$I_1 \leq \delta_1 \mathfrak{E}_{\mathrm{lps}}^2 + c_{\delta_1} \left(h_x^2 \|\tau + |\partial_x(j_h\pi)|\|_{p';\Omega}^{p'} + h_y^2 \left(\frac{h_y}{h_x}\right)^{p'-2} \|\tau_2 + |\partial_y(j_h\pi)|\|_{p';\Omega}^{p'} \right).$$

Note that, in order for Lemma 4 to be valid, the additional regularity of the pressure, $\pi \in W^{1,p'}$, is required. The stability estimates (3.6)–(3.7) and $\tau_2 = \frac{h_x}{h_y}\tau$ yields

$$I_1 \leq \delta_1 \mathfrak{E}_{\mathrm{lps}}^2 + c_{\delta_1} \left(h_x^2 \tau^{p'} |\Omega| + h_x^2 \|\partial_x(j_h\pi)\|_{p';\Omega}^{p'} + h_y^2 \left(\frac{h_y}{h_x}\right)^{p'-2} \|\partial_y(j_h\pi)\|_{p';\Omega}^{p'} \right)$$

$$\leq \delta_1 \mathfrak{E}_{\mathrm{lps}}^2 + c_{\delta_1} \left(h_x^2 \tau^{p'} |\Omega| + \left(h_x^2 + h_y^2 \left(\frac{h_y}{h_x}\right)^{p'-2} \left(\frac{h_x}{h_y}\right)^{p'} \right) \|\partial_x \pi\|_{p'}^{p'} \right) + c_{\delta_1} \left(h_x^2 \left(\frac{h_y}{h_x}\right)^{p'} + h_y^2 \left(\frac{h_y}{h_x}\right)^{p'-2} \right) \|\partial_y \pi\|_{p'}^{p'}$$

$$\leq \delta_1 \mathfrak{E}_{\mathrm{lps}}^2 + 2c_{\delta_1} \left(h_x^2 \tau^{p'} |\Omega| + h_x^2 \|\partial_x \pi\|_{p'}^{p'} + h_x^2 \left(\frac{h_y}{h_x}\right)^{p'} \|\partial_y \pi\|_{p'}^{p'} \right)$$

$$\leq \delta_1 \mathfrak{E}_{\mathrm{lps}}^2 + c_{\delta_1} h_x^2 (1 + \mathcal{A}^{-p'}).$$

The factors underlined with a wiggly line stems from the inverse aspect ratio introduced in the non-linear part of the stabilization, see (2.7). These balances the aspect ratio factor arising from the stability estimates of j_h, to prevent the error from being proportional to $\mathcal{A}^{p'-2}$. The terms underlined by a straight line stems from the h_y^2 factor in (2.7), which clearly prevents terms containing $h_x^2 \|\partial_y \pi\|^{p'}$ to enter.

To bound I_2, we follow the steps in the proof of [19, Theorem 3.1], namely applying Lemma 13 and using Young's inequality – these steps are identical on anisotropic meshes. We then apply the interpolation property (3.11) for $\mathcal{F}$, assuming that $\mathcal{F}(\mathbf{Dv}) \in W^{1,2}(\Omega)^{d \times d}$ to get

$$I_2 \lesssim c_{\delta_2} \|\mathcal{F}(\mathbf{D}j_h\mathbf{v}) - \mathcal{F}(\mathbf{Dv})\|_2^2 + \delta_2 \|\mathcal{F}(\mathbf{D}j_h\mathbf{v}) - \mathcal{F}(\mathbf{Dv}_h)\|_2^2$$

$$\lesssim \delta_2 \mathfrak{E}_{\text{lps}}^2 + c_{\delta_2} \left(\frac{h_x}{h_y} h_x^2 \|\partial_x \mathcal{F}(\mathbf{Dv})\|_2^2 + \frac{h_x}{h_y} h_y^2 \|\partial_y \mathcal{F}(\mathbf{Dv})\|_2^2 \right)$$

$$\lesssim \delta_2 \mathfrak{E}_{\text{lps}}^2 + c_{\delta_2} h_x^2 (\mathcal{A} + \mathcal{A}^{-1}).$$

For bounding the third term I_3, we first use generalized Korn's inequality, Lemma 13 and the definition of $\mathfrak{E}_{\text{lps}}$ to obtain the following estimate:

$$\|\nabla(j_h\mathbf{v} - \mathbf{v}_h)\|_p \leq \|\mathcal{F}(\mathbf{D}j_h\mathbf{v}) - \mathcal{F}(\mathbf{Dv}_h)\|_2 \|\epsilon + |\mathbf{D}j_h\mathbf{v}| + |\mathbf{Dv}_h|\|_p^{1-\frac{p}{2}} \leq \mathfrak{E}_{\text{lps}} \|\epsilon + |\mathbf{D}j_h\mathbf{v}| + |\mathbf{Dv}_h|\|_p^{1-\frac{p}{2}}. \qquad (3.24)$$

Applying Hölder's inequality and Young's inequality to I_3, and then inserting (3.24) and the interpolation property on the pressure term (assuming $\pi \in W^{1,p'}$), we have

$$I_3 \leq \|j_h\pi - \pi\|_{p'} \|\nabla(j_h\mathbf{v} - \mathbf{v}_h)\|_p$$

$$\leq \delta_3 \mathfrak{E}_{\text{lps}}^2 + c_{\delta_3} \|\epsilon + |\mathbf{D}j_h\mathbf{v}| + |\mathbf{Dv}_h|\|_p^{2-p} \|j_h\pi - \pi\|_{p'}^2$$

$$\leq \delta_3 \mathfrak{E}_{\text{lps}}^2 + c_{\delta_3} \left(\epsilon |\Omega|^{\frac{1}{p}} + \|\mathbf{D}j_h\mathbf{v}\|_p + \|\mathbf{Dv}_h\|_p \right)^{2-p} \left(h_x \|\partial_x \pi\|_{p'} + h_y \|\partial_y \pi\|_{p'} \right)^2$$

$$\leq \delta_3 \mathfrak{E}_{\text{lps}}^2 + c_{\delta_3} \left(\epsilon |\Omega|^{\frac{1}{p}} + \left(1 + \frac{h_x}{h_y} \right) \|\nabla\mathbf{v}\|_p + \|\nabla\mathbf{v}_h\|_p \right)^{2-p} \left(h_x \|\partial_x \pi\|_{p'} + h_y \|\partial_y \pi\|_{p'} \right)^2.$$

The stability of the interpolation operator, the stability of the continuous solution, and Lemma 12 give $\|\nabla j_h\mathbf{v}\|_p \lesssim \frac{h_x}{h_y} C_1(\mathbf{f}, \epsilon, \Omega, p)$ and $\|\nabla\mathbf{v}_h\|_p \leq C_2(\mathbf{f}, \epsilon, \Omega, p)$. Using these bounds, we get

$$I_3 \leq \delta_3 \mathfrak{E}_{\text{lps}}^2 + c_{\delta_3} \left(\frac{h_x}{h_y} \right)^{2-p} \left(h_x \|\partial_x \pi\|_{p'} + h_y \|\partial_y \pi\|_{p'} \right)^2$$

$$\lesssim \delta_3 \mathfrak{E}_{\text{lps}}^2 + c_{\delta_3} h_x^2 (\mathcal{A}^{2-p} + \mathcal{A}^{-p}).$$

For the last term, I_4, we use integration by parts, the orthogonality of $\mathbf{j}_h$, and the triangle inequality to split into components

$$I_4 = |(\mathbf{j}_h\mathbf{v} - \mathbf{v}, \nabla(j_h\pi - \pi_h))|$$

$$= |(\mathbf{j}_h\mathbf{v} - \mathbf{v}, \boldsymbol{\theta}_h(\nabla(j_h\pi - \pi_h)))|$$

$$\lesssim |(j_h v_x - v_x, \theta_h(\partial_x(j_h\pi - \pi_h)))| + |(j_h v_y - v_y, \theta_h(\partial_y(j_h\pi - \pi_h)))|.$$

The x-component is bounded by applying Hölder's inequality followed by Young's inequality

$$I_{4x} := |(j_h v_x - v_x, \theta_h(\partial_x(j_h\pi - \pi_h)))|$$

$$\lesssim \sum_{M \in \mathbb{M}_h} h_x^{-\frac{2}{p'}} \|j_h v_x - v_x\|_{p;M} h_x^{\frac{2}{p'}} \|\theta_h(\partial_x(j_h\pi - \pi_h))\|_{p';M}$$

$$\lesssim \sum_{M \in \mathbb{M}_h} c_{\delta_4} h_x^{-2(p-1)} \|j_h v_x - v_x\|_{p;M}^p + \delta_4 \sum_{M \in \mathbb{M}_h} h_x^2 \|\theta_h(\partial_x(j_h\pi - \pi_h))\|_{p';M}^{p'}.$$

Using the interpolation properties of j_h, the regularity $\mathbf{v} \in W^{2,p}$, and the fact that

$$\|\theta_h(\partial_x(j_h\pi - \pi_h))\|_{p';M}^{p'} \leq \|\mathcal{G}_1(\theta_h(\partial_x j_h\pi)) - \mathcal{G}_1(\theta_h(\partial_x \pi_h))\|_{2;M}^2,$$

which is in turn bounded by $\mathfrak{E}_{\text{lps}}$, we arrive at

$$I_{4x} \leq \delta_4 \mathfrak{E}_{\text{lps}}^2 + c_{\delta_4} h_x^{2-2p} \left(h_x^2 \|\partial_{xx}^2 v_x\|_{p;\Omega} + h_y^2 \|\partial_{yy}^2 v_x\|_{p;\Omega} + h_x h_y \|\partial_{xy}^2 v_x\|_{p;\Omega} \right)^p$$

$$\lesssim \delta_4 \mathfrak{E}_{\text{lps}}^2 + c_{\delta_4} h_x^2 (1 + \mathcal{A}^{-2p} + \mathcal{A}^{-p}).$$

Similarly for the y-component, $I_{4y} := |(j_h v_y - v_y, \theta_h(\partial_y(j_h\pi - \pi_h)))|$, we have

$$
\begin{aligned}
I_{4y} &\lesssim \sum_{M\in\mathbb{M}_h} h_y^{-\frac{2}{p'}}\left(\frac{h_y}{h_x}\right)^{\frac{2-p'}{p'}} \|j_h v_y - v_y\|_{p;M}\, h_y^{\frac{2}{p'}}\left(\frac{h_y}{h_x}\right)^{\frac{p'-2}{p'}}\|\theta_h(\partial_y(j_h\pi-\pi_h))\|_{p';M}\\
&\lesssim c_{\delta_4}\sum_{M\in\mathbb{M}_h} \underline{h_y^{-2(p-1)}\left(\frac{h_y}{h_x}\right)^{p-2}}\|j_h v_y - v_y\|_{p;M}^p + \delta_4 \sum_{M\in\mathbb{M}_h} h_y^2 \left(\frac{h_y}{h_x}\right)^{p'-2}\underwave{\|\theta_h(\partial_y(j_h\pi-\pi_h))\|_{p';M}^{p'}}\\
&\lesssim \delta_4 \mathfrak{E}_{\mathrm{lps}}^2 + c_{\delta_4}\underwave{h_y^{-2(p-1)}\left(\frac{h_y}{h_x}\right)^{p-2}}\left(h_x^2\|\partial_{xx}^2 v_y\|_{p;\Omega} + \underline{h_y^2\|\partial_{yy}^2 v_y\|_{p;\Omega}} + h_x h_y \|\partial_{xy}^2 v_y\|_{p;\Omega}\right)^p\\
&\lesssim \delta_4 \mathfrak{E}_{\mathrm{lps}}^2 + c_{\delta_4} h_y^{-p} h_x^{2-p}\left(h_x^2\|\partial_{xx}^2 v_y\|_{p;\Omega} + h_y^2\|\partial_{yy}^2 v_y\|_{p;\Omega} + h_x h_y \|\partial_{xy}^2 v_y\|_{p;\Omega}\right)^p\\
&\lesssim \delta_4 \mathfrak{E}_{\mathrm{lps}}^2 + c_{\delta_4} h_x^2(\mathcal{A}^p + \mathcal{A}^{-p} + 1).
\end{aligned}
$$

Again the factors underlined with a wiggly line stems from the inverse aspect ratio introduced in the non-linear part of the stabilization and the terms with a straight lines stems from choosing h_y^2 instead of h_x^2 in the y-component of (2.7).

Collecting $I_1 - I_4$ and combining constants independent of the element aspect ratio and stability parameter, and choosing $\delta_1, \delta_2, \delta_3, \delta_4$ small enough, we get

$$
\mathfrak{E}_{\mathrm{lps}}^2 \lesssim h_x^2(1 + \mathcal{A} + \mathcal{A}^{-1} + \mathcal{A}^{-p'} + \mathcal{A}^{2-p} + \mathcal{A}^{-p} + \mathcal{A}^{-2p} + \mathcal{A}^{-p} + \mathcal{A}^p).
$$

Since we assumed in this work an aspect ratio $\mathcal{A} > 1$, the dominant term is $\mathcal{A}^p$. Hence, we arrived at

$$
\mathfrak{E}_{\mathrm{lps}} \lesssim h_x \mathcal{A}^{\frac{p}{2}}, \tag{3.25}
$$

where the constant depends on $\alpha_0, p, \epsilon_0, \Omega, f$ and τ. This is exactly the assertion (3.21).

To arrive at the pressure error estimate (3.22), we first notice that the projection error $j_h\pi - \pi_h$ is in $\mathcal{Q}_h$, so that we can use the discrete inf-sup condition of (3.17) to get

$$
\begin{aligned}
\mathcal{A}^{-1}\beta_c\|j_h\pi - \pi_h\|_{p'} &\lesssim \sup_{\boldsymbol{\omega}_h\in\mathbf{X}_h}\frac{(\nabla\cdot\boldsymbol{\omega}_h, (j_h\pi - \pi_h))_\Omega}{\|\nabla\boldsymbol{\omega}_h\|_p} + \left(\sum_{M\in\mathbb{M}} h_x^{p'}\|\theta_h(\partial_x(j_h\pi-\pi_h))\|_{p',M}^{p'}\right)^{\frac{1}{p'}}\\
&\quad + \left(\sum_{M\in\mathbb{M}} h_y^{p'}\|\theta_h(\partial_y(j_h\pi-\pi_h))\|_{p',M}^{p'}\right)^{\frac{1}{p'}}\\
&=: J_1 + J_2 + J_3.
\end{aligned}
$$

By inserting the continuous and discrete equation, using Hölder's inequality, followed by (B.1) and the approximation properties of j_h (3.1), we get

$$
\begin{aligned}
J_1 &\lesssim \sup_{\boldsymbol{\omega}_h\in\mathbf{X}_h}\frac{(\mathbf{S}(\mathbf{Dv}) - \mathbf{S}(\mathbf{Dv}_h), \mathbf{D}\omega_h)_\Omega}{\|\nabla\boldsymbol{\omega}_h\|_p} + \sup_{\boldsymbol{\omega}_h\in\mathbf{X}_h}\frac{(j_h\pi - \pi, \nabla\cdot\boldsymbol{\omega}_h)_\Omega}{\|\nabla\boldsymbol{\omega}_h\|_p}\\
&\lesssim \|\mathbf{S}(\mathbf{Dv}) - \mathbf{S}(\mathbf{Dv}_h)\|_{p'} + \|j_h\pi - \pi\|_p'\\
&\lesssim \|\mathcal{F}(\mathbf{Dv}) - \mathcal{F}(\mathbf{Dv}_h)\|_2^{\frac{2}{p'}} + h_x\|\partial_x\pi\|_{p'} + h_y\|\partial_y\pi\|_{p'}\\
&\lesssim \mathfrak{E}_{\mathrm{lps}}^{\frac{2}{p'}} + h_x\|\partial_x\pi\|_{p'} + h_y\|\partial_y\pi\|_{p'}.
\end{aligned}
$$

The second and third term J_2 and J_3 are estimated by considering that $h_x \geq h_y$, and $p' \geq 2$, using Lemma 2, and the definition of $\mathfrak{E}_{\mathrm{lps}}^{2/p'}$:

$$
\begin{aligned}
J_2 &\lesssim h_x^{1-\frac{2}{p'}}\left(\sum_{M\in\mathbb{M}} h_x^2\|\theta_h(\partial_x(j_h\pi-\pi_h))\|_{p',M}^{p'}\right)^{\frac{1}{p'}}\\
&\lesssim h_x^{1-\frac{2}{p'}}\left(\sum_{M\in\mathbb{M}} h_x^2\|\mathcal{G}_1(\theta_h(\partial_x j_h\pi)) - \mathcal{G}_1(\theta_h(\partial_x\pi_h))\|_{2;M}^2\right)^{\frac{1}{p'}}\\
&\lesssim h_x^{1-\frac{2}{p'}}\mathfrak{E}_{\mathrm{lps}}^{\frac{2}{p'}}
\end{aligned}
$$

and

$$J_3 \lesssim h_y^{1-\frac{2}{p'}} \left(\sum_{M \in \mathbb{M}} h_y^2 \|\theta_h(\partial_y(j_h\pi - \pi_h))\|_{p',M}^{p'} \right)^{\frac{1}{p'}}$$

$$\lesssim h_x^{1-\frac{2}{p'}} \left(\sum_{M \in \mathbb{M}} h_y^2 \left(\frac{h_y}{h_x}\right)^{p'-2} \|\mathcal{G}_2(\theta_h(\partial_y j_h\pi)) - \mathcal{G}_2(\theta_h(\partial_y\pi_h))\|_{2;M}^2 \right)^{\frac{1}{p'}}$$

$$\lesssim h_x^{1-\frac{2}{p'}} \mathfrak{E}_{\mathrm{lps}}^{\frac{2}{p'}}.$$

Combining the estimates of J_1 and J_2 with the interpolation error, we get the following estimate of the discretization-error in the pressure:

$$\|\pi - \pi_h\|_{p'} \leq \|j_h\pi - \pi\|_{p'} + \|j_h\pi - \pi_h\|_{p'}$$

$$\lesssim \left(1 + \frac{\mathcal{A}}{\beta_c}\right)\left(h_x\|\partial_x\pi\|_{p'} + h_y\|\partial_y\pi\|_{p'}\right) + \frac{\mathcal{A}}{\beta_c}\left(1 + h_x^{1-\frac{2}{p'}}\right)\mathfrak{E}_{\mathrm{lps}}^{\frac{2}{p'}}.$$

Inserting (3.25) gives (3.22):

$$\|\pi - \pi_h\|_{p'} \lesssim \mathcal{A}h_x + \mathcal{A}(1 + h_x^{1-\frac{2}{p'}})(h_x\mathcal{A}^{\frac{p}{2}})^{\frac{2}{p'}}$$

$$\lesssim \mathcal{A}h_x + \mathcal{A}^{1+p/p'}(h_x^{\frac{2}{p'}} + h_x)$$

$$\lesssim \mathcal{A}^p h_x^{\frac{2}{p'}}. \qquad \qquad \square$$

Remark. For a stabilization without the $\frac{h_y}{h_x}$-factor in $\mathbf{M}_{s,22}$ in (2.8), i.e. for a stabilization on the form

$$s(\pi_h, q_h) := \alpha_0 \sum_{M \in \mathbb{M}} (\mathbf{M}_s \boldsymbol{\theta}_h \nabla\pi_h, \boldsymbol{\theta}_h \nabla q_h)_M),$$

$$\mathbf{M}_s(\partial_x\pi_h, \partial_x\pi_h) := \begin{bmatrix} h_x^2(\frac{1}{\tau}(\tau + |\theta_h\partial_x\pi_h|))^{p'-2} & 0 \\ 0 & h_y^2(\frac{1}{\tau}(\tau + |\theta_h\partial_y\pi_h|))^{p'-2}, \end{bmatrix},$$

we get, by setting the factors marked by awiggly line equal to one

$$\mathfrak{E}_{\mathrm{lps}} \lesssim h_x(\mathcal{A}^{\frac{p'-2}{2}} + \mathcal{A}^{p-1}).$$

The first term dominates for $p < \frac{3}{2}$, and as $p \to 1$ we have that $\frac{1}{2}(p' - 2) \to \infty$. In this sense, the inaccuracies related to the anisotropy of the mesh is amplified by the non-linearity of the material. For an isotropic stabilization, i.e. with

$$\mathbf{M}_s(\partial_x\pi_h, \partial_x\pi_h) = \begin{bmatrix} h_x^2(\frac{1}{\tau}(\tau + |\theta_h\partial_x\pi_h|))^{p'-2} & 0 \\ 0 & h_x^2(\frac{1}{\tau}(\tau + |\theta_h\partial_y\pi_h|))^{p'-2} \end{bmatrix}, \qquad (3.26)$$

we replace h_y by h_x in the terms underlined with a straight line (as well as setting the factors underlined by wiggly lines equal to one) in the estimates of I_1 and I_4 to see that

$$\mathfrak{E}_{\mathrm{lps}} \lesssim h_x(\mathcal{A}^{\frac{p'}{2}}).$$

Also terms with $h_x\|\partial_y\pi\|_p^{\prime p'/2}$ will now arise from I_1, where $\|\partial_y\pi\|_p'$ is typically large in a thin domain. The anisotropic stabilization instead involves more well-balanced terms $h_y\|\partial_y\pi\|_p^{\prime p'/2}$.

4 Numerical Experiments

We present numerical results for two different configurations. The first experiment is designed to show how the convergence rate depends on p. For this purpose, we have chosen a problem for which the projection error, not the interpolation error, dominates the pressure error. The second experiment is designed to compare the robustness and accuracy of the isotropic and anisotropic stabilization. It is constructed to qualitatively mimic the solution of a typical thin film flow scenario seen in e.g. glaciology but still be regular enough to allow for a wide spectrum of stabilization parameters α_0 to be tested. For both experiments, we have chosen the highest

aspect ratio $\mathcal{A}$ that the (non-specialized) linear solver allows for. The first configuration has an aspect ratio of the cell sizes of $\mathcal{A} = 100$, the second one of $\mathcal{A} = 1000$. The regularization parameter in (1.2) is chosen as $\epsilon = 10^{-5}$.

4.1 Experiment 1 – Convergence

We study how the numerical error decreases as the resolution increases. The experiment is similar to the convergence experiment in [18], but adapted to suit anisotropic problems. The domain is $\Omega := (-\frac{L}{2}, \frac{L}{2}) \times (-\frac{H}{2}, \frac{H}{2})$, with $L = 1$ and $H = 0.01$, and the parameter scaling the viscosity is $\mu_0 = 1.0$. The problem is chosen so that the exact solution is

$$\mathbf{v} = \begin{pmatrix} L\left(\left(\frac{x}{L}\right)^2 + \left(\frac{y}{H}\right)^2\right)^{\frac{a-1}{2}} \frac{y}{H} \\ -H\left(\left(\frac{x}{L}\right)^2 + \left(\frac{y}{H}\right)^2\right)^{\frac{a-1}{2}} \frac{x}{L} \end{pmatrix}, \quad \pi = -\left(\left(\frac{x}{L}\right)^2 + \left(\frac{y}{H}\right)^2\right)^{\frac{b}{2}} \frac{x}{L} \frac{y}{H}.$$

In order for the velocity to fulfill the regularity assumption $\mathcal{F}(\mathbf{Dv}) \in L_2$, it is required that $a > 1$. Following [18], we set $a = 1.01$. For the pressure to be regular enough, we need $b > -1 - \frac{p'}{2}$. We set $b = 0.1$, which fulfills the regularity requirement but is low enough to ensure that the pressure error is dominated by the projection error $(\sim h^{\frac{2}{p'}})$, not the interpolation error $(\sim h^2)$ for the mesh resolutions tested here. The right-hand side is given as $\mathbf{f} = -\nabla \cdot \mathbf{S}(\mathbf{Dv}) + \nabla\pi$ and the Dirichlet boundary condition is $\mathbf{g} = \mathbf{v}|_{\partial\Omega}$. We set $\alpha_0 = 0.01$, and $\tau = 1.0$.

The number of elements in each dimension is N, so that the element aspect ratio $\mathcal{A} = \frac{h_x}{h_y} = 100$ is independent of N and equals the domain aspect ratio. As the convergence rate of the pressure is expected to depend on the parameter p, we test three different cases, namely $p = 1.1$, $p = 1.5$, and $p = 1.9$ on seven different meshes, with $N = \{16, 64, 256, 1034, 4096, 16384, 65536\}$. The errors $\|p - p_h\|_{p'}$, $\|v_x - v_{x,h}\|_p$, $\|v_y - v_{y,h}\|_p$ and their respective order of convergence are shown in Table 1. For the velocity error we observe convergence

	$N \times N$	$\|p - p_h\|_{p'}$	conv.	$\|v_x - v_{x,h}\|_{1,p}$	conv.	$\|v_y - v_{y,h}\|_{1,p}$	conv.
$p = 1.1$	16	$2.94e-03$	5.06	$3.57e-03$	0.98	$1.50e-05$	1.10
	64	$3.03e-04$	3.28	$1.78e-03$	1.01	$7.07e-06$	1.09
	256	$7.32e-05$	2.05	$8.75e-04$	1.02	$3.58e-06$	0.98
	1024	$1.77e-05$	2.05	$4.33e-04$	1.02	$1.79e-06$	1.00
	4096	$1.14e-05$	0.63	$2.15e-04$	1.01	$8.47e-07$	1.08
	16384	$1.01e-05$	0.17	$1.07e-04$	1.01	$4.29e-07$	0.98
	65536	$8.96e-06$	0.18	$5.33e-05$	1.00	$2.32e-07$	0.89
	theory		0.18		1		1
$p = 1.5$	16	$2.74e-03$	2.88	$1.31e-02$	0.90	$5.32e-05$	0.95
	64	$1.22e-04$	4.49	$6.80e-03$	0.94	$2.73e-05$	0.96
	256	$9.03e-05$	0.43	$3.48e-03$	0.97	$1.45e-05$	0.92
	1024	$1.67e-05$	2.43	$1.77e-03$	0.98	$7.67e-06$	0.91
	4096	$8.48e-06$	0.98	$8.94e-04$	0.98	$4.00e-06$	0.94
	16384	$5.11e-06$	0.73	$4.50e-04$	0.99	$1.98e-06$	1.01
	65536	$3.14e-06$	0.70	$2.26e-04$	0.99	$9.92e-07$	1.00
	theory		0.67		1		1
$p = 1.9$	16	$6.32e-03$	0.52	$2.89e-02$	0.83	$1.16e-04$	0.83
	64	$3.77e-04$	4.07	$1.58e-02$	0.87	$6.42e-05$	0.85
	256	$2.62e-04$	0.53	$8.49e-03$	0.90	$3.57e-05$	0.85
	1024	$4.70e-05$	2.48	$4.50e-03$	0.92	$1.97e-05$	0.86
	4096	$1.79e-05$	1.39	$2.36e-03$	0.93	$1.08e-05$	0.86
	16384	$8.81e-06$	1.02	$1.23e-03$	0.94	$5.80e-06$	0.90
	65536	$4.45e-06$	0.98	$6.38e-04$	0.95	$2.98e-06$	0.96
	theory		0.95		1		1

Table 1: Results of Experiment 1, verification of Theorem 1.

order close to $\mathcal{O}(h)$ for all choices of p and all meshes which corresponds quite well with the theoretical order. The error estimate in Theorem 1 with respect to the pressure π gives $\|p - p_h\|_{p'} = \mathcal{O}(h^{\frac{2}{p'}}) = \mathcal{O}(h^{2(\frac{1}{p}-1)})$. In the worst case, $p = 1.1$, the numerical order decays on finer meshes, but stabilizes at $O(h^{0.18})$ which exactly corresponds with the theoretical order. For moderate $p = 1.9$, the observed order of convergence in the pressure is on the finest mesh $O(h^{0.98})$ which is slightly above the theoretical order $O(h^{0.95})$. Hence, we state that, as the mesh is refined, the convergence order approaches the expected, theoretical convergence order.

4.2 Experiment 2 – Robustness

The anisotropic stabilization was designed to balance the order of magnitude of the x- and y-component of the stabilization term. We thus expect that the anisotropic stabilization increases the range of stabilization parameters α_0 that does not lead to under-nor over-stabilization, compared to an isotropic stabilization. To verify this, we run simulations with the anisotropic stabilization (2.8), a fully isotropic stabilization (3.26), and a isotropic stabilization which still includes the factor $\mathcal{A}^{2-p'}$ that was introduced to suppress error terms proportional to $\mathcal{A}^{p'}$ or $\mathcal{A}^{p'-2}$, i.e. a stabilization on the form

$$s(\pi_h, q_h) := \alpha_0 \sum_{M \in \mathbb{M}} (\mathbf{M}_s \boldsymbol{\theta}_h \nabla \pi_h, \boldsymbol{\theta}_h \nabla q_h)_M), \tag{4.1}$$

$$\mathbf{M}_s(\partial_x \pi_h, \partial_x \pi_h) := \begin{bmatrix} h_x^2(\frac{1}{\tau}(\tau + \frac{h_y}{h_x}|\theta_h \partial_x \pi_h|))^{p'-2} & 0 \\ 0 & h_x^2(\frac{1}{\tau}(\tau + \frac{h_y}{h_x}|\theta_h \partial_y \pi_h|))^{p'-2} \end{bmatrix}. \tag{4.2}$$

We will call this *the semi-isotropic stabilization*. For each method, we vary α_0 and measure the error in velocity and pressure.

The domain is $\Omega := (0, L) \times (0, H)$, with $L = 1$ and two different choices $H = 0.001$ and $H = 0.01$. We choose a right-hand side $\mathbf{f} = -\nabla \cdot \mathbf{S}(\mathbf{D v}) + \nabla \pi$ and Dirichlet boundary conditions $\mathbf{g} = \mathbf{v}|_{\partial\Omega}$ so that the exact solution $(\mathbf{v}, \pi)$ is

$$\mathbf{v} = \begin{pmatrix} \sin(f\frac{x}{L})\cos(f\frac{y}{H}) \\ -\frac{H}{L}\cos(f\frac{x}{L})\sin(f\frac{y}{H}) \end{pmatrix}, \quad \pi = c_p \sin\left(\frac{\pi x}{L}\right)\cos\left(\pi\frac{\pi y}{H}\right),$$

where $c_p = 100$ and $f = 0.01\pi$. The large pressure component and small vertical component is typical of many non-Newtonian thin-film problems, found for instance in glaciology. We set $p = 1.4$, $\mu_0 = 0.1$, and $\tau = c_p$. The number of intervals in each direction is $N = 32$. The mesh and problem is chosen so that the solver converges for a large range of α_0 for both, isotropic and anisotropic stabilization.

The fully isotropic stabilization (3.26) does not converge. This is indeed the reason why we implemented the semi-isotropic stabilization (4.1). The results for both the semi-isotropic and the anisotropic stabilizations are shown in Figures 2 and 3 for $H = 10^{-3}$ and $H = 10^{-2}$ ($\mathcal{A} = 1000$ and $\mathcal{A} = 100$), respectively. For both methods a too small stabilization parameter α_0 leads to under-stabilization of the pressure, while a too large α_0 leads to errors in the velocity components, especially in the vertical component. The errors in the pressure are due to the spurious oscillations associated to element choices that do not fulfill the inf-sup condition. The errors in the velocity components are due to the consistency error caused by the extra stabilization term. This consistency error is present to some degree for all $\alpha_0 > 0$, but we only call the problem over-stabilized when it is the dominant error. The vertical velocity component is evidently more sensitive to over-stabilization than the horizontal velocity component. This sensitivity of the vertical component is typical of flows in thin domains, and are observed in applications such as glaciology [17].

For the semi-isotropic stabilization, there is no choice of α_0 that does not lead to neither under-stabilization of the pressure nor over-stabilization of the velocity, meaning that there is no α_0 which is high enough to suppress spurious oscillations in the pressure at the same time as it is low enough to avoid significant consistency errors in the velocity components. In contrast to this, the anisotropic stabilization is less sensitive to the choice of α_0 and it is possible to choose an α_0 which yields a stable pressure solution while at the same time avoiding large over-stabilization errors in the velocity. The optimal stabilization parameter for the
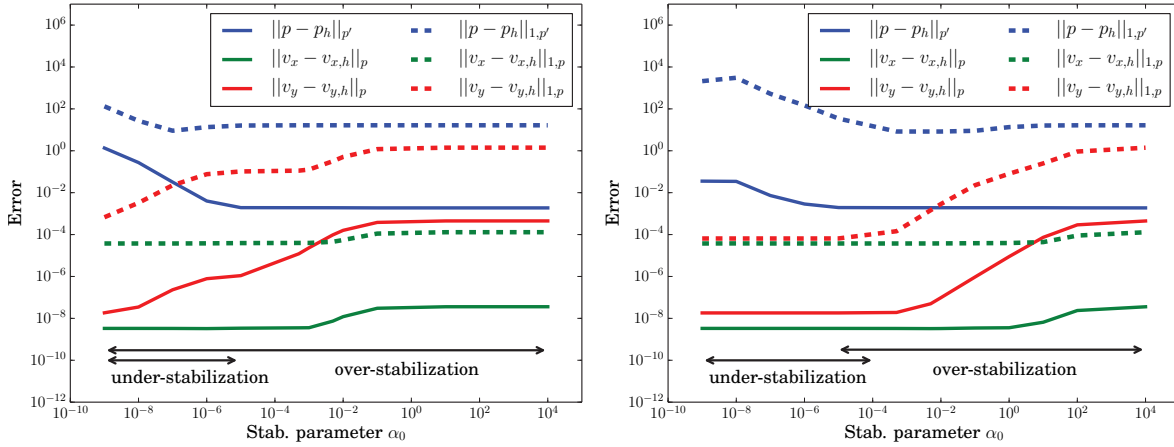
Figure 2: Results of Experiment 2 with aspect ratio 1000. The error as a function of the stabilization parameter α_0 for both semi-isotropic (left) as in (3.26) and anisotropic stabilization (right) as in (2.8). The fully isotropic stabilization did not converge. The range of α_0 for which the problem is under-stabilized or over-stabilized is indicated by $\leftrightarrow$.
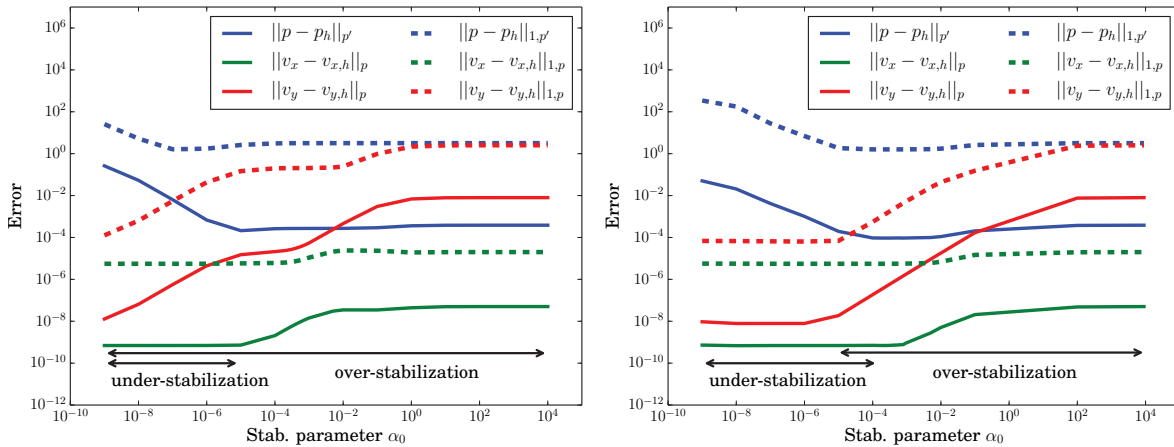

Figure 3: Results of Experiment 2 with aspect ratio 100. The error as a function of the stabilization parameter α_0 for both semi-isotropic (left) as in (3.26) and anisotropic stabilization (right) as in (2.8). The fully isotropic stabilization did not converge. The range of α_0 for which the problem is under-stabilized or over-stabilized is indicated by $\leftrightarrow$.

anisotropic stabilization parameter is rather small (around 10^{-4}), while for the semi-isotropic stabilization it is even smaller (around 10^{-6}). For anisotropic stabilization, the optimal value of α_0 seems to be fairly independent of the aspect ratio. Note that the computational domains differ between the two cases $\mathcal{A} = 1000$ and $\mathcal{A} = 100$. In practice, the optimal stabilization parameter is of course not known, but the moderate value for the anisotropic stabilization is more reasonable. For common choices of $\alpha_0 = 0.1$ to $\alpha_0 = 1$, the error in the vertical velocity is reduced by two to three order of magnitudes by using the anisotropic stabilization instead of the semi-isotropic one. The problem of choosing optimal values of α_0 with the (semi-)isotropic stabilization methods (e.g. LPS, Galerkin-Least-Squares (GLS), Interior Penalty (IP)) has also been reported in [17], where the accuracy in pressure is quite sensitive with respect to α_0 even for moderate aspect ratios (≈ 80) for a different configuration. Hence, anisotropic stabilization is beneficial to be used.

5 Summary and Conclusion

We have derived a priori estimates for an equal-order bi-linear finite element discretization of the *p*-Stokes equations on anisotropic meshes, with power-law exponent $p \in (1, 2)$. To stabilize the inf-sup unstable equal-

order discretization, we designed an anisotropic local projection stabilization for the *p*-Stokes equations. Without a specially designed stabilization term errors related to the anisotropy of the problem is amplified by the non-linearity of the *p*-Stokes equations.

Our error estimates where confirmed in numerical experiments. They show in particular that our anisotropic stabilization is more robust with respect to user defined parameters than an semi-isotropic stabilization, while a fully isotropic standard stabilization did not converge. For typical parameter choices, an anisotropic stabilization reduces errors the velocity by several orders of magnitudes. Such results are relevant to applications in e.g. glaciology and industrial material processing.

A Proof of Lemma 1

Proof. The proof is an extension of the proof of [21, Lemma 3.2]. For the meshes we consider $|\det DF_K| = h_x h_y$ and $|\det DF_M| \sim h_x h_y$. We then have

$$\|q_h\|_{v',M}^{v'} = \int_M q(x)^{v'}\,dx = \int_{\hat{M}} \hat{q}_h(\hat{x})^{v'}|\det DF_K|\,d\hat{x} \sim h_x h_y \|\hat{q}_h\|_{v',M}^{v'}.$$

To estimate $(\omega_h, q_h)_M$, let $\hat{b}$ be the hat function associated with the node in the middle of a patch M and choose $\omega_h(x) := (\hat{q}_h \cdot \hat{b}) \circ F_M^{-1}(x)$, where $q_h \in Y_h(M)$ and $\hat{q}_h \in Q_0(\hat{M})$. Then $\hat{\omega}_h(x) := (\hat{q}_h \cdot \hat{b}) \in \hat{X}_h^0(\hat{M})$ since $(\hat{q}_h \cdot \hat{\omega}_h)$ is continuous on the closure of $\hat{M}$, $(\hat{q}_h \cdot \hat{\omega}_h)|_{\hat{K}} \in Q_1(\hat{K})$ when $\hat{K}$ belongs to the patch $\hat{M}$, and $\hat{b}|_{\partial \hat{M}} = 0$. Therefore $(\hat{\omega}_h, \hat{q}_h)_{\hat{M}} = (\hat{q}_h(\hat{x}), \hat{b}(\hat{x})\hat{q}_h(\hat{x}))_{\hat{M}}$. Since b is positive $(\hat{q}_h(\hat{x}), \hat{q}_h(\hat{x})\hat{b}(\hat{x}))_{\hat{M}}^{\frac{1}{2}}$ is a norm. Because all norms on finite-dimensional spaces are equivalent $\|\hat{q}_h\sqrt{(\hat{b})}\|_{\hat{M}} \gtrsim \|\hat{q}_h\|_{v',\hat{M}}$ and we are working on a reference patch the constant does not depend on the cell sizes. Therefore

$$(\omega_h, q_h)_M \gtrsim h_x h_y \int_{\hat{M}} \hat{q}_h(\hat{x})^2 \hat{b}(\hat{x})\,d\hat{x} \gtrsim h_x h_y \|\hat{q}_h\|_{v',\hat{M}}^2.$$

Now we estimate $\|\omega_h\|_M$. We have $|\hat{b}| \leq 1$ for all $\hat{x} \in \hat{M}$ since it is a hat function on the reference element. Then

$$\|\omega_h\|_{v,M}^v = \int_{\hat{M}} \hat{b}^v(\hat{x})\hat{q}_h{}^v|\det DF_K|\,d\hat{x} \leq 1 \int_{\hat{M}} \hat{q}_h{}^v|\det DF_K|\,d\hat{x} \lesssim h_x h_y \|\hat{q}_h\|_{v,\hat{M}}^v.$$

Again using the argument that all norms on finite-dimensional spaces are equivalent and we are working on a reference patch, we have that $\|\hat{q}_h\|_{v,\hat{M}} \sim \|\hat{q}_h\|_{v',\hat{M}}$, where the constant is independent of the cell sizes. Combining everything, we have the following: for all $q_h \in Y_h(M)$ there exists $\omega_h \in X_h^0(M)$ such that

$$\frac{(q_h, \omega_h)_M}{\|\omega_h\|_{v,M}\|q_h\|_{v',M}} \geq \beta_{\text{loc}} \frac{h_x h_y \|\hat{q}_h\|_{v',\hat{M}}^2}{(h_x h_y \|\hat{q}_h\|_{v',\hat{M}}^v)^{\frac{1}{v}}(h_x h_y \|\hat{q}_h\|_{v',M}^{v'})^{\frac{1}{v'}}} = \beta_{\text{loc}} \frac{h_x h_y}{(h_x h_y)^{\frac{1}{v}}(h_x h_y)^{\frac{1}{v'}}} = \beta_{\text{loc}},$$

where the constant β_{loc} is independent of the aspect ratio. □

B Relations Between S, D and the Help Functions φ, $\mathcal{F}$

Lemma 13 (Relations Between **S**, **D**, φ and $\mathcal{F}$). *Given that **S** is of (p, ϵ)-structure, i.e. that Assumption 1 is fulfilled, the following relations are valid:*
(i) *It holds uniformly for all* **P**, **Q** $\in \mathbb{R}^{d \times d}$ *that*

$$(\mathbf{S}(\mathbf{P}) - \mathbf{S}(\mathbf{Q})) : (\mathbf{P} - \mathbf{Q}) \sim (\epsilon + |\mathbf{P}^{\text{sym}}| + |\mathbf{Q}^{\text{sym}}|)^{p-2}|\mathbf{P}^{\text{sym}} - \mathbf{Q}^{\text{sym}}|^2$$

$$\sim \varphi_{\epsilon+|\mathbf{P}^{\text{sym}}|}(|\mathbf{P}^{\text{sym}} - \mathbf{Q}^{\text{sym}}|)$$

$$\sim |\mathcal{F}(\mathbf{P}) - \mathcal{F}(\mathbf{Q})|^2,$$

$$|\mathbf{S}(\mathbf{P}) - \mathbf{S}(\mathbf{Q})| \sim (\epsilon + |\mathbf{P}^{\text{sym}}| + |\mathbf{Q}^{\text{sym}}|)^{p-2}|\mathbf{P}^{\text{sym}} - \mathbf{Q}^{\text{sym}}|$$

$$\sim \varphi'_{\epsilon+|\mathbf{P}^{\text{sym}}|}(|\mathbf{P}^{\text{sym}} - \mathbf{Q}^{\text{sym}}|).$$

(ii) For $\mathbf{u}, \mathbf{v} \in \mathbf{W}^{1,p}$ the stress-tensor and help-function $\mathcal{F}$ relates to the quasi-norm, or so-called "natural distance", of [3] given by $|\mathbf{v}|^2_{(p,\mathbf{u})} := \int_\Omega (\epsilon + |\mathbf{Du}| + |\mathbf{Dv}|)^{p-2} |\mathbf{Dv}|^2$ as

$$(\mathbf{S}(\mathbf{Du}) - \mathbf{S}(\mathbf{Dv}), \mathbf{Du} - \mathbf{Dv})_\Omega \sim \|\mathbf{F}(\mathbf{Du}) - \mathbf{F}(\mathbf{Dv})\|_2^2$$

$$\sim \int_\Omega \varphi_{\epsilon+|\mathbf{Du}|}(|\mathbf{Du} - \mathbf{Dv}|)$$

$$\sim |\mathbf{u} - \mathbf{v}|^2_{(p,\mathbf{u})}.$$

(iii) For $p \in (1, 2]$ and for all $\mathbf{v}, \mathbf{u} \in \mathbf{W}^{1,p}$,

$$\|\mathbf{D}(\mathbf{v} - \mathbf{u})\|_p^2 \lesssim \|\mathcal{F}(\mathbf{Dv}) - \mathcal{F}(\mathbf{Du})\|_2^2 \|\epsilon + |\mathbf{Dv}| + |\mathbf{Du}|\|_p^{2-p},$$

$$\|\mathcal{F}(\mathbf{Dv}) - \mathcal{F}(\mathbf{Du})\|_2^2 \lesssim \|\mathbf{D}(\mathbf{v} - \mathbf{u})\|_p^p,$$

$$\|\mathbf{S}(\mathbf{Dv}) - \mathbf{S}(\mathbf{Du})\|_{p'} \lesssim \|\mathcal{F}(\mathbf{Dv}) - \mathcal{F}(\mathbf{Du})\|_2^{\frac{2}{p'}}. \tag{B.1}$$

The constants depend on p, σ_0 and σ_1. In particular, it does not depend on ϵ.

Proof. See [14] and [19].   □

Lemma 14 (Properties of φ). *The following statements hold:*
(i)   *Due to the convexity of φ we have uniformly in $t, s, a \in \mathbb{R}_0^+$ that*

$$\varphi_a(t + s) \le \varphi_a(s) + \varphi_a(t).$$

(ii) *Young-type inequality: For all $\delta > 0$ there exists $c(\delta) > 0$ that only depends on p and δ so that for all $s, t, a \ge 0$,*

$$s\varphi_a'(t) + \varphi_a'(s)t \le \delta\varphi_a(s) + c(\delta)\varphi_a(t).$$

(iii) *Change of shift: for all $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^{d\times d}$, $\epsilon \ge 0$ we have*

$$\varphi_{\epsilon+|\mathbf{P}|}(|\mathbf{P} - \mathbf{Q}|) \sim \varphi_{\epsilon+|\mathbf{Q}|}(|\mathbf{P} - \mathbf{Q}|),$$

where the constant depends only on p. The same holds if φ is replaced by φ'. Furthermore, it holds that for $t \ge 0$,

$$\varphi_{\epsilon+|\mathbf{P}|}(t) \sim c(\delta)\varphi_{\epsilon+|\mathbf{Q}|}(t) + \delta\varphi_{\epsilon+|\mathbf{Q}|}(|\mathbf{P} - \mathbf{Q}|),$$

where $\delta > 0$ and $c(\delta) > 0$ only depend on p and δ.

Proof. These properties are shown in [14] and [19].   □

Further relations between $\mathbf{S}$, φ, and $\mathcal{F}$ are described in [18].

C Approximation and Stability Properties of i_h

From [2, Theorem 3.3] we have that for $\omega \in W^{1,v}$,

$$\|\omega - i_h\omega\|_{0,v;M} \lesssim h_x\|\partial_x\omega\|_{0,v;\Lambda(M)} + h_y\|\partial_y\omega\|_{0,v;\Lambda(M)},$$

$$\|\omega - i_h\omega\|_{1,v;M} \lesssim \|\omega\|_{1,v;\Lambda(M)},$$

and for $\omega \in W^{2,v}$,

$$\|\omega - i_h\omega\|_{0,v;M} \lesssim h_x^2\|\partial_x^2\omega\|_{0,v;\Lambda(M)} + h_y^2\|\partial_y^2\omega\|_{0,v;\Lambda(M)} + h_xh_y\|\partial_{xy}\omega\|_{0,v;\Lambda(M)},$$

$$\|\omega - i_h\omega\|_{1,v;M} \lesssim h_x\|\partial_x\omega\|_{1,v;\Lambda(M)} + h_y\|\partial_y\omega\|_{1,v;\Lambda(M)}.$$

Using $\|i_h\omega\| = \|\omega - i_h - \omega\| \le \|\omega - i_h\omega\| + \|\omega\|$ and the approximation properties, we get the stability properties

$$\|i_h\omega\|_M \lesssim \|\omega\|_{\Lambda(M)} + h_x\|\partial_x\omega\|_{\Lambda(M)} + h_y\|\partial_y\omega\|_{\Lambda(M)}, \tag{C.1}$$

$$\|\partial_x i_h\omega\|_{v;M} \lesssim \|\omega\|_{1,v;\Lambda(M)}, \tag{C.2}$$

$$\|\partial_y i_h\omega\|_{v;M} \lesssim \|\omega\|_{1,v;\Lambda(M)}. \tag{C.3}$$

In [10] there is another useful stability estimate

$$\|\partial_x i_h\omega\|_{v;M} \lesssim \|\partial_x\omega\|_{v;\Lambda(M)} + \frac{h_y}{h_x}\|\partial_y\omega\|_{v;\Lambda(M)}, \tag{C.4}$$

$$\|\partial_y i_h\omega\|_{v;M} \lesssim \|\partial_y\omega\|_{v;\Lambda(M)}. \tag{C.5}$$

These properties also extend to vector-valued functions $\boldsymbol{\omega} \in \mathbf{W}_0^{1,v}$.

References

[1] J. Ahlkrona, N. Kirchner and P. Lötstedt, A numerical study of scaling relations for non-Newtonian thin-film flows with applications in ice sheet modelling, *Quart. J. Mech. Appl. Math.* **66** (2013), no. 4, 417–435.

[2] T. Apel, *Anisotropic Finite Elements: Local Estimates and Applications*, Adv. Numer. Math., B. G. Teubner, Stuttgart, 1999.

[3] J. W. Barrett and W. B. Liu, Quasi-norm error bounds for the finite element approximation of a non-Newtonian flow, *Numer. Math.* **68** (1994), no. 4, 437–456.

[4] R. Becker, *An adaptive finite element method for the incompressible Navier–Stokes equations on time-dependent domains*, PhD thesis, Heidelberg University, 1995.

[5] R. Becker and M. Braack, A finite element pressure gradient stabilization for the Stokes equations based on local projections, *Calcolo* **38** (2001), no. 4, 173–199.

[6] L. Belenki, L. C. Berselli, L. Diening and M. Růžička, On the finite element approximation of *p*-Stokes systems, *SIAM J. Numer. Anal.* **50** (2012), no. 2, 373–397.

[7] J. Blasco, An anisotropic GLS-stabilized finite element method for incompressible flow problems, *Comput. Methods Appl. Mech. Engrg.* **197** (2008), no. 45–48, 3712–3723.

[8] M. Braack, Anisotropic H^1-stable projections on quadrilateral meshes, in: *Numerical Mathematics and Advanced Applications*, Springer, Berlin (2006), 495–503.

[9] M. Braack, G. Lube and L. Röhe, Divergence preserving interpolation on anisotropic quadrilateral meshes, *Comput. Methods Appl. Math.* **12** (2012), no. 2, 123–138.

[10] M. Braack and T. Richter, Local projection stabilization for the Stokes system on anisotropic quadrilateral meshes, in: *Numerical Mathematics and Advanced Applications*, Springer, Berlin (2006), 770–778.

[11] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, 3rd ed., Texts Appl. Math. 15, Springer, New York, 2008.

[12] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer Ser. Comput. Math. 15, Springer, New York, 1991.

[13] L. Diening, C. Kreuzer and E. Süli, Finite element approximation of steady flows of incompressible fluids with implicit power-law-like rheology, *SIAM J. Numer. Anal.* **51** (2013), no. 2, 984–1015.

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

[15] L. Diening, M. Růžička and K. Schumacher, A decomposition technique for John domains, *Ann. Acad. Sci. Fenn. Math.* **35** (2010), no. 1, 87–114.

[16] V. Girault and P.-A. Raviart, *Finite Element Methods for Navier–Stokes Equations: Theory and Algorithms*, Springer Ser. Comput. Math. 5, Springer, Berlin, 1986.

[17] C. Helanow and J. Ahlkrona, Stabilized equal low-order finite elements in ice sheet modeling—accuracy and robustness, *Comput. Geosci.* **22** (2018), no. 4, 951–974.

[18] A. Hirn, *Finite element approximation of problems in non-Newtonian fluid mechanics*, PhD thesis, Heidelberg Univeryity and Charles University, 2011.

[19] A. Hirn, Approximation of the *p*-Stokes equations with equal-order finite elements, *J. Math. Fluid Mech.* **15** (2013), no. 1, 65–88.

[20] R. E. Johnson and R. M. McMeeking, Near-surface flow in glaciers obeying Glen's law, *Quart. J. Mech. Appl. Math.* **37** (1984), no. 2, 273–291.

[21] G. Matthies, P. Skrzypacz and L. Tobiska, A unified convergence analysis for local projection stabilisations applied to the Oseen problem, *M2AN Math. Model. Numer. Anal.* **41** (2007), no. 4, 713–742.

[22] S. Micheletti, S. Perotto and M. Picasso, Stabilized finite elements on anisotropic meshes: A priori error estimates for the advection-diffusion and the Stokes problems, *SIAM J. Numer. Anal.* **41** (2003), no. 3, 1131–1162.

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