

$L^2(I; H^1(\Omega)^d)$ and $L^2(I; L^2(\Omega)^d)$ best
approximation type error estimates for Galerkin
solutions of transient Stokes problems

Dmitriy Leykekhman^{1*} and Boris Vexler^{2†}

^{1*}Department of Mathematics, University of Connecticut, 341 Mansfield
Road U1009, Storrs, 06269-1009, CT, USA.

²Chair of Optimal Control, Center for Mathematical Science, Technical
University of Munich, Garching by Munich, 85748, Germany.

*Corresponding author(s). E-mail(s): dmitriy.leykekhman@uconn.edu;

Contributing authors: vexler@tum.de;

[†]These authors contributed equally to this work.

Abstract

In this paper we establish best approximation type estimates for the fully discrete Galerkin solutions of transient Stokes problem in $\mathbf{L}^2(I; \mathbf{L}^2(\Omega)^d)$ and $\mathbf{L}^2(I; \mathbf{H}^1(\Omega)^d)$ norms. These estimates fill the gap in the error analysis of the transient Stokes problems and have a number of applications. The analysis naturally extends to inhomogeneous parabolic problems. The best type $\mathbf{L}^2(I; \mathbf{H}^1(\Omega))$ error estimate are new even for scalar parabolic problems.

Keywords: Stokes problem, transient Stokes, parabolic problems, finite elements, discontinuous Galerkin, error estimates, best approximation, fully discrete

1 Introduction

In this paper we consider the following transient Stokes problem with no-slip boundary conditions,

$$\begin{aligned} \partial_t \vec{u} - \Delta \vec{u} + \nabla p &= \vec{f} && \text{in } I \times \Omega, \\ \nabla \cdot \vec{u} &= 0 && \text{in } I \times \Omega, \\ \vec{u} &= \vec{0} && \text{on } I \times \partial\Omega, \\ \vec{u}(0) &= \vec{u}_0 && \text{in } \Omega. \end{aligned} \quad (1)$$

We assume that $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, is a bounded polygonal/polyhedral Lipschitz domain, $T > 0$ and $I = (0, T]$. In the next section we make precise assumptions on the data, which allow for a weak formulation of the problem and provide regularity for the velocity $\vec{u} \in L^2(I; H^1(\Omega)^d)$. We consider fully discrete approximations of problem (1), where we use compatible finite elements (i.e. satisfying a uniform inf-sup condition) for the space discretization and the discontinuous Galerkin method for the temporal discretization. In our previous work [2], we have established the best type error estimate in $L^\infty(I; L^2(\Omega))$ norm,

$$\|\vec{u} - \vec{u}_{\tau h}\|_{L^\infty(I; L^2(\Omega))} \leq C \ell_\tau \left(\|\vec{u} - \vec{\chi}\|_{L^\infty(I; L^2(\Omega))} + \|\vec{u} - R_h^S(\vec{u}, p)\|_{L^\infty(I; L^2(\Omega))} \right), \quad (2)$$

where $\vec{u}_{\tau h}$ is the fully discrete finite element approximation of the velocity $\vec{u}$, R_h^S is the Ritz projection for the stationary Stokes problem, $\vec{\chi}$ is an arbitrary function from the finite element approximation of the velocity spaces $X_\tau^q(\vec{V}_h)$, and ℓ_τ is a logarithmic term. Such results are desirable for example in the analysis of PDE constrained optimal control problems since they do not require any additional regularity assumptions on the solution beyond the regularity which follows directly from the problem data. The estimate (2) was an improvement of the main results in [4, Thm. 4.9], where the error was estimated simultaneously in $L^\infty(I; L^2(\Omega))$ and $L^2(I; H^1(\Omega))$ norms, and the bounds there required the presence of the "mixed terms". A natural question, which actually was raised by one of the referees for [2], is it possible to obtain a best type approximation result just w.r.t. $L^2(I; H^1(\Omega))$ norm? Surprisingly, such results are not available even for scalar parabolic problems. In this paper, we give positive answer to this question and establish the following best type error estimates

$$\|\vec{u} - \vec{u}_{\tau h}\|_{L^2(I \times \Omega)} \leq C \left(\|\vec{u} - \vec{\chi}\|_{L^2(I \times \Omega)} + \|\vec{u} - \pi_\tau \vec{u}\|_{L^2(I \times \Omega)} + \|\vec{u} - R_h^S(\vec{u}, p)\|_{L^2(I \times \Omega)} \right) \quad (3)$$

and

$$\begin{aligned} \|\nabla(\vec{u} - \vec{u}_{\tau h})\|_{L^2(I \times \Omega)} &\leq C \left(\|\nabla(\vec{u} - \vec{\chi})\|_{L^2(I \times \Omega)} + \|\nabla(\vec{u} - \pi_\tau \vec{u})\|_{L^2(I \times \Omega)} \right. \\ &\quad \left. + \|\nabla(\vec{u} - R_h^S(\vec{u}, p))\|_{L^2(I \times \Omega)} \right), \end{aligned} \quad (4)$$

where as above $\vec{u}_{\tau h}$ is the fully discrete finite element approximation of the velocity $\vec{u}$, $\vec{\chi}$ is an arbitrary function from the finite element approximation of the velocity spaces $X_\tau^q(\vec{V}_h)$, $\pi_\tau \vec{u}$ is a certain time projection of the $\vec{u}$ on the time discrete space X_τ^q , and

R_h^S is the Ritz projection for the stationary Stokes problem. The results (3) and (4) link the approximation error for the fully discrete transient Stokes problem to the best possible approximation of a continuous solution $\vec{u}$ in the discrete space $X_\tau^q(\vec{V}_h)$ as well as the approximation of the stationary Stokes problem in $\vec{V}_h$ and a time projection. Such results go in hand with only natural assumptions on the problem data and are desirable in a number of applications. In contrast to (2), where some mild assumptions on time steps are needed, the estimates (3) and (4) do not require any restrictions and are valid for discontinuous time discretizations on arbitrary partitions. In addition, the estimate (3) does not require any additional regularity of the domain, thus allowing, e.g., for reentrant corners and edges. Moreover, (3) does not require the mesh to be quasi-uniform nor shape regular. Therefore, the result is also true for graded and even anisotropic meshes (provided the discrete inf-sup condition holds uniformly on such meshes). However, the second estimate (4), does require the stability of the discrete Leray projection in H^1 norm, which so far is established for the quasi-uniform meshes on convex domains. These results also naturally hold for the inhomogeneous heat equation, where the proofs can be simplified and extended to more general meshes, (see Section 7).

Under the additional assumption of convexity of Ω and some approximation properties of the discrete spaces, from (3) and (4) we easily derive optimal order (in terms of regularity) error estimates of the form

$$\|\vec{u} - \vec{u}_{\tau h}\|_{L^2(I; H^s(\Omega))} \leq C \left(\tau^{1-s/2} + h^{2-s} \right) \left(\|\vec{f}\|_{L^2(I; L^2(\Omega))} + \|\vec{u}_0\|_{\vec{V}^1} \right), \quad s = 0, 1,$$

where $\vec{V}^1$ is an appropriate space introduced in the next section. This estimate is optimal with respect to both the assumed regularity of the data and the order of convergence and has already been used by the authors in several applications [17, 20].

The above results naturally hold for simpler case of inhomogeneous heat equation with straightforward change of operators and function spaces. The $L^2(I; L^2(\Omega))$ best type approximation result is essentially shown in [18], instead of the best approximation the optimal error estimate in terms of data of the form (5) is given, however, the $L^2(I; H^1(\Omega))$ best approximation result seems to be new.

The rest of the paper is structured as follows. In Section 2, we introduce the function spaces, key operators and weak formulations of the problem with regularity results. In Section 3, we introduce fully discrete Galerkin approximation of the problem. Section 4 is devoted to stability results of the fully discrete velocity solutions. In Section 4 we review some stability and approximation results for the stationary Stokes problem, which we require for our main results Theorem 11 and Theorem 13 in the next Section 6. Finally, in Section 7, we briefly discuss the extension of the main results to scalar parabolic problems.

2 Continuous problem

In this section, we introduce function spaces we require for the analysis of (1) and state some of the main properties of these spaces. In our presentation we follow the notation and presentation of [14, Section 1 and Section 2].

2.1 Function spaces and Stokes operator

In the following, we will use the usual notation to denote the Lebesgue spaces L^p and Sobolev spaces H^k and $W^{k,p}$. The space $L_0^2(\Omega)$ will denote a subspace of $L^2(\Omega)$ with mean-zero functions. The inner product on $L^2(\Omega)$ as well as on $L^2(\Omega)^d$ is denoted by $(\cdot, \cdot)$. To improve readability, we omit the superscript d when having for example $L^2(\Omega)^d$ appear as subscript to norms. We also introduce the following function spaces

$$\mathcal{V} = \{\vec{v} \in C_0^\infty(\Omega)^d \mid \nabla \cdot \vec{v} = 0\}, \quad \vec{V}^0 = \overline{\mathcal{V}}^{L^2}, \quad \vec{V}^1 = \overline{\mathcal{V}}^{H^1}, \quad (5)$$

where the notation in the last line denotes the completion of the space $\mathcal{V}$ with respect to the $L^2(\Omega)^d$ and $H^1(\Omega)^d$ topology, respectively. Notice that functions in $\vec{V}^1$ have zero boundary conditions in the trace sense. Alternatively we have

$$\vec{V}^1 = \{\vec{v} \in H_0^1(\Omega)^d \mid \nabla \cdot \vec{v} = 0\}$$

by [8, Theorem III.4.1].

We define the vector-valued Laplace operator

$$-\Delta: D(\Delta) \rightarrow L^2(\Omega)^d,$$

where the domain $D(\Delta)$ is understood with respect to $L^2(\Omega)^d$ and is given as

$$D(\Delta) = \{\vec{v} \in H_0^1(\Omega)^d \mid \Delta \vec{v} \in L^2(\Omega)\}.$$

If the domain Ω is convex, then the standard $H^2(\Omega)$ regularity implies $D(\Delta) = H_0^1(\Omega)^d \cap H^2(\Omega)^d$. In addition, we introduce the space $\vec{V}^2$ as

$$\vec{V}^2 = \vec{V}^1 \cap D(\Delta).$$

We will also use the following Helmholtz decomposition (cf. [19, Chapter I, Theorem 1.4] and [8, Theorem III.1.1])

$$L^2(\Omega)^d = \vec{V}^0 \oplus \nabla (H^1(\Omega) \cap L_0^2(\Omega)). \quad (6)$$

As usual we define the Helmholtz projection $\mathbb{P}: L^2(\Omega)^d \rightarrow \vec{V}^0$ (often also called the Leray projection) as the L^2 -projection from $L^2(\Omega)^d$ onto $\vec{V}^0$. Using $\mathbb{P}$ and $-\Delta$, we define the Stokes operator $A: \vec{V}^2 \rightarrow \vec{V}^0$ as

$$A = -\mathbb{P}\Delta|_{\vec{V}^2}. \quad (7)$$

The operator A is a self adjoint, densely defined and positive definite operator on $\vec{V}^0$. We note that $D(A) = \vec{V}^2$. Similar to the Laplace operator, for convex polyhedral domains Ω we have the following H^2 regularity bound due to [5, 15]

$$\|\vec{v}\|_{H^2(\Omega)} \leq C \|A\vec{v}\|_{L^2(\Omega)}, \quad \forall \vec{v} \in \vec{V}^2. \quad (8)$$

2.2 Weak formulation and regularity

In this section we discuss the weak formulation and the regularity of the transient Stokes problem (1). We will use the notation $L^s(I; X)$ for the corresponding Bochner space with a Banach space X . Moreover, we will use also the standard notation $H^1(I; X)$. The inner product in $L^2(I; L^2(\Omega)^d)$ is denoted by $(\cdot, \cdot)_{I \times \Omega}$.

Proposition 1. *Let $\vec{f} \in L^2(I; (\vec{V}^1)')$ and $\vec{u}_0 \in \vec{V}^0$. Then there exists a unique solution $\vec{u} \in L^2(I; \vec{V}^1) \cap C(\bar{I}, \vec{V}^0)$ with $\partial_t \vec{u} \in L^2(I; (\vec{V}^1)')$ fulfilling $\vec{u}(0) = \vec{u}_0$ and*

$$\langle \partial_t \vec{u}, \vec{v} \rangle + (\nabla \vec{u}, \nabla \vec{v})_{I \times \Omega} = (\vec{f}, \vec{v})_{I \times \Omega} \quad \text{for all } \vec{v} \in L^2(I; \vec{V}^1). \quad (9)$$

There holds

$$\|\nabla \vec{u}\|_{L^2(I; L^2(\Omega))} + \|\vec{u}'\|_{L^2(I; (\vec{V}^1)')} + \|\vec{u}\|_{C(\bar{I}; L^2(\Omega))} \leq C \left(\|\vec{f}\|_{L^2(I; (\vec{V}^1)')} + \|\vec{u}_0\|_{\vec{V}^0} \right).$$

Proof. For the proof of the above result we refer to [19, Chapter III, Theorem 1.1]. $\square$

It is well known, cf. again [19], that the equation (9) can be understood as an abstract parabolic problem

$$\begin{aligned} \partial_t \vec{u} + A\vec{u} &= \mathbb{P}\vec{f} \quad \text{for a.a. } t \in I, \\ \vec{u}(0) &= \vec{u}_0, \end{aligned} \quad (10)$$

with the Stokes operator A defined in (7).

The next theorem provides the space-time weak formulation in both variables, velocity and pressure. Please note, that no additional regularity of the domain is required.

Theorem 2. *Let $\vec{f} \in L^2(I; L^2(\Omega)^d)$ and $\vec{u}_0 \in \vec{V}^1$. Then there exists a unique solution $(\vec{u}, p)$ with*

$$\vec{u} \in L^2(I; \vec{V}^1), \quad \partial_t \vec{u}, A\vec{u} \in L^2(I; L^2(\Omega)^d) \quad \text{and} \quad p \in L^2(I; L_0^2(\Omega))$$

fulfilling $\vec{u}(0) = \vec{u}_0$ and

$$(\partial_t \vec{u}, \vec{v})_{I \times \Omega} + (\nabla \vec{u}, \nabla \vec{v})_{I \times \Omega} - (p, \nabla \cdot \vec{v})_{I \times \Omega} + (\nabla \cdot \vec{u}, \xi)_{I \times \Omega} = (\vec{f}, \vec{v})_{I \times \Omega} \quad (11)$$

for all

$$\vec{v} \in L^2(I; H_0^1(\Omega)^d) \quad \text{and} \quad \xi \in L^2(I; L_0^2(\Omega)).$$

There holds the estimate

$$\|\partial_t \vec{u}\|_{L^2(I \times \Omega)} + \|A\vec{u}\|_{L^2(I \times \Omega)} + \|p\|_{L^2(I \times \Omega)} \leq C \left(\|\vec{f}\|_{L^2(I \times \Omega)} + \|\vec{u}_0\|_{\vec{V}^1} \right).$$

Proof. The proof is given in [2], Theorem 2.10. $\square$

Corollary 3. *Let the assumptions of Theorem 2 hold and let in addition the domain Ω be convex. Then we have $\vec{u} \in L^2(I; H^2(\Omega)^d)$ and $p \in L^2(I, H^1(\Omega))$ with the corresponding estimates*

$$\|\vec{u}\|_{L^2(I; H^2(\Omega))} + \|p\|_{L^2(I; H^1(\Omega))} \leq C \left(\|\vec{f}\|_{L^2(I \times \Omega)} + \|\vec{u}_0\|_{\vec{V}^1} \right).$$

Proof. This result is shown in [19], Proposition 1.2 in Chapter 3, for C^2 domains, but the proof is valid for convex domains as well. See also Corollary 2.11 in [2]. $\square$

3 Fully discrete discretization

In this section we consider the discrete version of the operators presented in the previous section and introduce fully discrete Galerkin solution.

3.1 Spatial discretization

Let $\{\mathcal{T}_h\}$ be a family of triangulations of $\bar{\Omega}$, consisting of closed simplices, where we denote by h the maximum mesh-size. Let $\vec{X}_h \subset H_0^1(\Omega)^d$ and $M_h \subset L_0^2(\Omega)$ be a pair of compatible finite element spaces, i.e., them satisfying a uniform discrete inf-sup condition,

$$\sup_{\vec{v}_h \in \vec{X}_h} \frac{(q_h, \nabla \cdot \vec{v}_h)}{\|\nabla \vec{v}_h\|_{L^2(\Omega)}} \geq \beta \|q_h\|_{L^2(\Omega)} \quad \forall q_h \in M_h, \quad (12)$$

with a constant $\beta > 0$ independent of h .

We introduce the usual discrete Laplace operator $-\Delta_h: \vec{X}_h \rightarrow \vec{X}_h$ by

$$(-\Delta_h \vec{z}_h, \vec{v}_h) = (\nabla \vec{z}_h, \nabla \vec{v}_h), \quad \forall \vec{z}_h, \vec{v}_h \in \vec{X}_h. \quad (13)$$

To define a discrete version of the Stokes operator A , we first define the space of discretely divergence-free vectors $\vec{V}_h$ as

$$\vec{V}_h = \{\vec{v}_h \in \vec{X}_h \mid (\nabla \cdot \vec{v}_h, q_h) = 0 \quad \forall q_h \in M_h\}. \quad (14)$$

Using this space we can define the discrete Leray projection $\mathbb{P}_h: L^1(\Omega)^d \rightarrow \vec{V}_h$ to be the L^2 -projection onto $\vec{V}_h$, i.e.,

$$(\mathbb{P}_h \vec{u}, \vec{v}_h) = (\vec{u}, \vec{v}_h) \quad \forall \vec{v}_h \in \vec{V}_h. \quad (15)$$

Using $\mathbb{P}_h$, we define the discrete Stokes operator $A_h: \vec{V}_h \rightarrow \vec{V}_h$ as $A_h = -\mathbb{P}_h \Delta_h|_{\vec{V}_h}$. By this definition we have that for $\vec{u}_h \in \vec{V}_h$, $A_h \vec{u}_h \in \vec{V}_h$ fulfills

$$(A_h \vec{u}_h, \vec{v}_h) = (\nabla \vec{u}_h, \nabla \vec{v}_h), \quad \forall \vec{v}_h \in \vec{V}_h. \quad (16)$$

Notice, since $\vec{V}_h \subset \vec{X}_h$, for $\vec{v}_h \in \vec{V}_h$ we obtain

$$(A_h \vec{v}_h, \vec{v}_h) = (\nabla \vec{v}_h, \nabla \vec{v}_h) \geq \lambda_0 \|\vec{v}_h\|_{L^2(\Omega)}^2, \quad (17)$$

where λ_0 is the smallest eigenvalue of $-\Delta$. This implies that the eigenvalues of A_h are also positive and bounded from below by λ_0 .

For deriving error estimates, we make additional assumptions on the approximation properties of the finite element spaces. We assume, there exist approximation operators P_h and r_h that fulfill the following properties. For $P_h \in \mathcal{L}(H_0^1(\Omega)^d; \vec{X}_h)$ and $r_h \in \mathcal{L}(L^2(\Omega); \bar{M}_h)$ with $\bar{M}_h$ corresponding to M_h without the zero-mean value constraint, we assume the following assumptions hold.

Assumption 1 (Stability of P_h in $H^1(\Omega)^d$). *There exists a constant C independent of h such that*

$$\|\nabla P_h(\vec{v})\|_{L^2(\Omega)} \leq C \|\nabla \vec{v}\|_{L^2(\Omega)}, \quad \forall \vec{v} \in H_0^1(\Omega)^d. \quad (18)$$

Assumption 2 (Preservation of discrete divergence for P_h). *It holds*

$$(\nabla \cdot (\vec{v} - P_h(\vec{v})), q_h) = 0, \quad \forall q_h \in \bar{M}_h, \quad \forall \vec{v} \in H_0^1(\Omega)^d. \quad (19)$$

Assumption 3 (Inverse Inequality). *There is a constant C independent of h such that*

$$\|\vec{v}_h\|_{H^1(\Omega)} \leq Ch^{-1} \|\vec{v}_h\|_{L^2(\Omega)} \quad \forall \vec{v}_h \in \vec{X}_h. \quad (20)$$

Assumption 4 (L^2 approximation). *For any $\vec{v} \in H^2(\Omega)^d$ and any $q \in H^1(\Omega)$ exists C independent of h , $\vec{v}$ and q such that*

$$\|P_h(\vec{v}) - \vec{v}\|_{L^2(\Omega)} + h \|\nabla(P_h(\vec{v}) - \vec{v})\|_{L^2(\Omega)} \leq Ch^2 \|\nabla^2 \vec{v}\|_{L^2(\Omega)}, \quad (21)$$

$$\|r_h(q) - q\|_{L^2(\Omega)} \leq Ch \|\nabla q\|_{L^2(\Omega)}. \quad (22)$$

Assumptions 1-4 hold for example for Taylor-Hood and Mini elements [10, Sec. 6] and the constructions of the operators P_h and r_h are also provided there.

3.2 Temporal discretization: the discontinuous Galerkin method

In this section we introduce the discontinuous Galerkin method for the time discretization of the transient Stokes equations, a similar method was considered, e.g., in [4] and [2]. For that, we partition $I = (0, T]$ into subintervals $I_m = (t_{m-1}, t_m]$ of length $\tau_m = t_m - t_{m-1}$, where $0 = t_0 < t_1 < \dots < t_{M-1} < t_M = T$. The maximal and minimal time steps are denoted by $\tau = \max_m \tau_m$ and $\tau_{\min} = \min_m \tau_m$, respectively. For a given Banach space $\mathcal{B}$ and the order $q \in \mathbb{N}$ we define the semi-discrete space $X_\tau^q(\mathcal{B})$ of piecewise polynomial functions in time as

$$X_\tau^q(\mathcal{B}) = \{\vec{v}_\tau \in L^2(I; \mathcal{B}) \mid \vec{v}_\tau|_{I_m} \in \mathcal{P}_q(\mathcal{B}), m = 1, 2, \dots, M\}, \quad (23)$$

where $\mathcal{P}_q(\mathcal{B})$ is the space of polynomial functions of degree less or equal q in time with values in $\mathcal{B}$, i.e.,

$$\mathcal{P}_q(\mathcal{B}) = \{\vec{v}_\tau: I_m \rightarrow \mathcal{B} | \vec{v}_\tau(t) = \sum_{j=0}^q \vec{v}^j \phi_j(t), \vec{v}^j \in \mathcal{B}, j = 0, \dots, q\}. \quad (24)$$

Here, $\{\phi_j(t)\}$ is a polynomial basis in t of the space $\mathcal{P}_q(I_m)$ of polynomials with degree less or equal q over the interval I_m . We use the following standard notation for a function $\vec{u} \in X_\tau^q(L^2(\Omega)^d)$

$$\vec{u}_m^+ = \lim_{\varepsilon \rightarrow 0^+} \vec{u}(t_m + \varepsilon), \quad \vec{u}_m^- = \lim_{\varepsilon \rightarrow 0^+} \vec{u}(t_m - \varepsilon), \quad [\vec{u}]_m = \vec{u}_m^+ - \vec{u}_m^-. \quad (25)$$

We define the bilinear form $\mathfrak{B}$ by

$$\mathfrak{B}(\vec{u}, \vec{v}) = \sum_{m=1}^M (\partial_t \vec{u}, \vec{v})_{I_m \times \Omega} + (\nabla \vec{u}, \nabla \vec{v})_{I \times \Omega} + \sum_{m=2}^M ([\vec{u}]_{m-1}, \vec{v}_{m-1}^+)_{\Omega} + (\vec{u}_0^+, \vec{v}_0^+)_{\Omega}.$$

With this bilinear form we define the fully discrete approximation for the transient Stokes problem on the discrete divergence free space $X_\tau^q(\vec{V}_h)$:

$$\vec{u}_{\tau h} \in X_\tau^q(\vec{V}_h) : \mathfrak{B}(\vec{u}_{\tau h}, \vec{v}_{\tau h}) = (\vec{f}, \vec{v}_{\tau h})_{I \times \Omega} + (\vec{u}_0, \vec{v}_{\tau h,0}^+)_{\Omega} \quad \forall \vec{v}_{\tau h} \in X_\tau^q(\vec{V}_h). \quad (26)$$

By a standard argument one can see that this formulation possesses a unique solution (existence follows from uniqueness by the fact that (26) is equivalent to a quadratic system of linear equations).

Remark 1. *Note, that the data $\vec{f}$ and $\vec{u}_0$ in (26) can be replaced by $\mathbb{P}_h \vec{f}$ and $\mathbb{P}_h \vec{u}_0$ respectively (with $\mathbb{P}_h$ being the discrete Leray projection (15)) without changing the solution.*

The above formulation is not a conforming discretization of the divergence free formulation (9) due to the fact that $X_\tau^q(\vec{V}_h)$ is not a subspace of $L^2(I; \vec{V}^1)$. In order to introduce a velocity-pressure discrete formulation (as a discretization of (11)) we consider the following bilinear form

$$\begin{aligned} B((\vec{u}, p), (\vec{v}, q)) &= \sum_{m=1}^M (\partial_t \vec{u}, \vec{v})_{I_m \times \Omega} + (\nabla \vec{u}, \nabla \vec{v})_{I \times \Omega} - (p, \nabla \cdot \vec{v})_{I \times \Omega} + (\nabla \cdot \vec{u}, q)_{I \times \Omega} \\ &\quad + \sum_{m=2}^M ([\vec{u}]_{m-1}, \vec{v}_{m-1}^+)_{\Omega} + (\vec{u}_0^+, \vec{v}_0^+)_{\Omega}. \end{aligned} \quad (27)$$

The corresponding fully discrete formulation reads: find $(\vec{u}_{\tau h}, p_{\tau h}) \in X_\tau^q(\vec{X}_h \times M_h)$ such that

$$B((\vec{u}_{\tau h}, p_{\tau h}), (\vec{v}_{\tau h}, q_{\tau h})) = (\vec{f}, \vec{v}_{\tau h})_{I \times \Omega} + (\vec{u}_0, \vec{v}_{\tau h,0}^+)_{\Omega} \quad \forall (\vec{v}_{\tau h}, q_{\tau h}) \in X_\tau^q(\vec{X}_h \times M_h). \quad (28)$$

We note, that for the temporal discretization we use polynomials of the same order for the velocity and pressure. The next proposition states the equivalence of the formulation (26) and (28).

Proposition 4. *For a solution $(\vec{u}_{\tau h}, p_{\tau h})$ of (28) the discrete velocity $\vec{u}_{\tau h}$ fulfills (26). Moreover, for a solution $\vec{u}_{\tau h}$ of (26) there exists a unique $p_{\tau h} \in X_\tau^q(M_h)$ such that the pair $(\vec{u}_{\tau h}, p_{\tau h})$ fulfills (28). In particular the solution of (28) is unique.*

Proof. The proof is given in [2], Proposition 4.2. $\square$

The next proposition provides the Galerkin orthogonality relation for the velocity pressure discretization (28), which is essential for our analysis. Please note, that for the velocity formulation (26) the Galerkin orthogonality does not hold due to the fact that $X_\tau^q(\vec{V}_h)$ is not a subspace of $L^2(I, \vec{V}^1)$.

Proposition 5. *Let the assumptions of (2) be fulfilled, i.e., $\vec{f} \in L^2(I; L^2(\Omega)^d)$ and $\vec{u}_0 \in \vec{V}^0$. Then there holds for the solution $(\vec{u}, p)$ of (11)*

$$B((\vec{u}, p), (\vec{v}_{\tau h}, q_{\tau h})) = (\vec{f}, \vec{v}_{\tau h})_{I \times \Omega} + (\vec{u}_0, \vec{v}_{\tau h, 0}^+)_{\Omega} \quad \forall (\vec{v}_{\tau h}, q_{\tau h}) \in X_\tau^q(\vec{X}_h \times M_h)$$

and consequently

$$B((\vec{u} - \vec{u}_{\tau h}, p - p_{\tau h}), (\vec{v}_{\tau h}, q_{\tau h})) = 0 \quad \forall (\vec{v}_{\tau h}, q_{\tau h}) \in X_\tau^q(\vec{X}_h \times M_h). \quad (29)$$

Proof. The proof is given in [2], Proposition 4.3. $\square$

In the following, we also consider a dual problem, where we use a dual representation of the bilinear form B

$$\begin{aligned} B((\vec{u}, p), (\vec{v}, q)) = & - \sum_{m=1}^M \langle \vec{u}, \partial_t \vec{v} \rangle_{I_m \times \Omega} + (\nabla \vec{u}, \nabla \vec{v})_{I \times \Omega} - (p, \nabla \cdot \vec{v})_{I \times \Omega} \\ & + (\nabla \cdot \vec{u}, q)_{I \times \Omega} - \sum_{m=1}^{M-1} (\vec{u}_m^-, [\vec{v}]_m)_{\Omega} + (\vec{u}_M^-, \vec{v}_M^-)_{\Omega}, \end{aligned} \quad (30)$$

which is obtained by integration by parts and rearranging the terms in the sum.

We will also need the following projection π_τ for $v \in C(I, L^2(\Omega))$ with $\pi_\tau v|_{I_m} \in \mathcal{P}_q(L^2(\Omega))$ for $m = 1, 2, \dots, M$ on each subinterval I_m by

$$(\pi_\tau v - v, \phi)_{I_m \times \Omega} = 0, \quad \forall \phi \in \mathcal{P}_{q-1}(I_m, L^2(\Omega)), \quad q > 0, \quad (31a)$$

$$\pi_\tau v(t_m^-) = v(t_m^-). \quad (31b)$$

In the case $q = 0$, $\pi_\tau v$ is defined solely by the second condition. The following approximation property also holds

$$\|\pi_\tau u - u\|_{L^2(I_m \times \Omega)} \leq C\tau \|\partial_t u\|_{L^2(I_m \times \Omega)} \quad \forall u \in H^1(I_m; L^2(\Omega)). \quad (32)$$

4 Fully discrete stability results

In this section, we establish stability results, which are essential for our main result.

Theorem 6. For $\vec{f} \in L^2(I, L^2(\Omega)^d)$ and $\vec{u}_0 \in \vec{V}^1$. Let $\vec{u}_{\tau h} \in X_\tau^q(\vec{V}_h)$ be the solution to (26). Then there holds

$$\begin{aligned} & \left(\sum_{m=1}^M \|\partial_t \vec{u}_{\tau h}\|_{L^2(I_m \times \Omega)}^2 \right)^{1/2} + \|A_h \vec{u}_{\tau h}\|_{L^2(I \times \Omega)} + \left(\sum_{m=1}^M \tau_m \|\tau_m^{-1} [\vec{u}_{\tau h}]_{m-1}\|_{L^2(\Omega)}^2 \right)^{1/2} \\ & \leq C \left(\|\mathbb{P}_h \vec{f}\|_{L^2(I \times \Omega)} + \|\nabla \mathbb{P}_h \vec{u}_0\|_{L^2(\Omega)} \right). \end{aligned} \quad (33)$$

Proof. The proof follows the lines of the corresponding proof in [18] replacing $-\Delta_h$ by A_h . $\square$

For the $L^2(I; H^1(\Omega)^d)$ norm estimates we also need the following result.

Theorem 7. For $\vec{f} \in L^2(I, L^2(\Omega)^d)$ and $\vec{u}_0 \in \vec{V}^0$. Let $\vec{u}_{\tau h} \in X_\tau^q(\vec{V}_h)$ be the solution to (26). Then there holds

$$\begin{aligned} & \left(\sum_{m=1}^M \|\partial_t \nabla A_h^{-1} \vec{u}_{\tau h}\|_{L^2(I_m \times \Omega)}^2 \right)^{1/2} + \|\nabla \vec{u}_{\tau h}\|_{L^2(I \times \Omega)} \\ & + \left(\sum_{m=1}^M \tau_m \|\tau_m^{-1} [\nabla A_h^{-1} \vec{u}_{\tau h}]_{m-1}\|_{L^2(\Omega)}^2 \right)^{1/2} \leq C \left(\|\nabla A_h^{-1} \mathbb{P}_h \vec{f}\|_{L^2(I \times \Omega)} + \|\mathbb{P}_h \vec{u}_0\|_{L^2(\Omega)} \right). \end{aligned} \quad (34)$$

Proof. The proof goes along the lines of the proof of Theorem 4.1 [18]. We will provide the proof for completeness. We will provide to the proof for $q \geq 1$, the case of $q = 0$, i.e. the lowest order piecewise constant case, we leave to the reader.

Writing (26) in an alternative formulation of the dG time stepping scheme at each I_m and testing with $\vec{v}_{\tau h} = \vec{u}_{\tau h}$, we obtain

$$(\partial_t \vec{u}_{\tau h}, \vec{u}_{\tau h})_{I_m \times \Omega} + \|\nabla \vec{u}_{\tau h}\|_{L^2(I_m \times \Omega)}^2 + ([\vec{u}_{\tau h}]_{m-1}, \vec{u}_{\tau h, m-1}^+)_{\Omega} = (\vec{f}, \vec{u}_{\tau h})_{I_m \times \Omega}.$$

Using the identities

$$(\partial_t \vec{u}_{\tau h}, \vec{u}_{\tau h})_{I_m \times \Omega} = \frac{1}{2} \|\vec{u}_{\tau h, m}^-\|_{L^2(\Omega)}^2 - \frac{1}{2} \|\vec{u}_{\tau h, m-1}^+\|_{L^2(\Omega)}^2$$

and

$$([\vec{u}_{\tau h}]_{m-1}, \vec{u}_{\tau h, m-1}^+)_{\Omega} = \frac{1}{2} \|\vec{u}_{\tau h, m-1}^+\|_{L^2(\Omega)}^2 + \frac{1}{2} \|[\vec{u}_{\tau h}]_{m-1}\|_{L^2(\Omega)}^2 - \frac{1}{2} \|\vec{u}_{\tau h, m-1}^-\|_{L^2(\Omega)}^2$$

gives us

$$\frac{1}{2} \|\vec{u}_{\tau h, m}^-\|_{L^2(\Omega)}^2 - \frac{1}{2} \|\vec{u}_{\tau h, m-1}^-\|_{L^2(\Omega)}^2 + \frac{1}{2} \|[\vec{u}_{\tau h}]_{m-1}\|_{L^2(\Omega)}^2 + \|\nabla \vec{u}_{\tau h}\|_{L^2(I_m \times \Omega)}^2 = (\vec{f}, \vec{u}_{\tau h})_{I_m \times \Omega}.$$

Summing over m , we obtain

$$\frac{1}{2}\|\vec{u}_{\tau h,M}^-\|_{L^2(\Omega)}^2 + \frac{1}{2}\sum_{m=1}^M \|[\vec{u}_{\tau h}]_{m-1}\|_{L^2(\Omega)}^2 + \|\nabla \vec{u}_{\tau h}\|_{L^2(I \times \Omega)}^2 = (\vec{f}, \vec{u}_{\tau h})_{I \times \Omega} + \frac{1}{2}\|\mathbb{P}_h \vec{u}_0\|_{L^2(\Omega)}^2.$$

To treat the term involving $\vec{f}$, we write it as

$$(\vec{f}, \vec{u}_{\tau h})_{I \times \Omega} = (\mathbb{P}_h \vec{f}, \vec{u}_{\tau h})_{I \times \Omega} = (A_h A_h^{-1} \mathbb{P}_h \vec{f}, \vec{u}_{\tau h})_{I \times \Omega} = (\nabla A_h^{-1} \mathbb{P}_h \vec{f}, \nabla \vec{u}_{\tau h})_{I \times \Omega},$$

where in the last step we used that $\vec{u}_{\tau h}$ is discrete divergence free. Using the Cauchy-Schwarz and geometric-arithmetic mean inequalities

$$(\vec{f}, \vec{u}_{\tau h})_{I \times \Omega} \leq \frac{1}{2}\|\nabla \vec{u}_{\tau h}\|_{L^2(I \times \Omega)}^2 + \frac{1}{2}\|\nabla A_h^{-1} \mathbb{P}_h \vec{f}\|_{L^2(I \times \Omega)}^2.$$

Thus, we obtain

$$\|\nabla \vec{u}_{\tau h}\|_{L^2(I \times \Omega)}^2 \leq \|\nabla A_h^{-1} \mathbb{P}_h \vec{f}\|_{L^2(I \times \Omega)}^2 + \|\mathbb{P}_h \vec{u}_0\|_{L^2(\Omega)}^2. \quad (35)$$

Next, we test (26) with $\vec{v}_{\tau h} |_{I_m} = (t - t_{m-1})A_h^{-1}\partial_t \vec{u}_{\tau h}$. Noticing that the jump terms disappear and using the identity,

$$(\partial_t \vec{u}_{\tau h}, A_h^{-1} \partial_t \vec{u}_{\tau h})_{\Omega} = (A_h A_h^{-1} \partial_t \vec{u}_{\tau h}, A_h^{-1} \partial_t \vec{u}_{\tau h})_{\Omega} = \|\nabla A_h^{-1} \partial_t \vec{u}_{\tau h}\|_{L^2(\Omega)}^2,$$

on each time interval I_m , which follows from $\partial_t \vec{u}_{\tau h}$ being discretely divergence free for each $t \in I_m$, we obtain

$$\begin{aligned} \int_{I_m} (t - t_{m-1}) \|\nabla A_h^{-1} \partial_t \vec{u}_{\tau h}\|_{L^2(\Omega)}^2 dt &= - \int_{I_m} (t - t_{m-1}) (\nabla \vec{u}_{\tau h}, \nabla A_h^{-1} \partial_t \vec{u}_{\tau h})_{\Omega} dt \\ &\quad + \int_{I_m} (t - t_{m-1}) (\vec{f}, A_h^{-1} \partial_t \vec{u}_{\tau h})_{\Omega} dt \\ &= \int_{I_m} (t - t_{m-1}) (-\nabla \vec{u}_{\tau h} + \nabla A_h^{-1} \mathbb{P}_h \vec{f}, \nabla A_h^{-1} \partial_t \vec{u}_{\tau h})_{\Omega} dt. \end{aligned}$$

Using the Cauchy-Schwarz inequality

$$\begin{aligned} \int_{I_m} (t - t_{m-1}) (-\nabla \vec{u}_{\tau h} + \nabla A_h^{-1} \mathbb{P}_h \vec{f}, \nabla A_h^{-1} \partial_t \vec{u}_{\tau h})_{\Omega} dt &\leq \\ \left(\int_{I_m} (t - t_{m-1}) \|\nabla \vec{u}_{\tau h} - \nabla A_h^{-1} \mathbb{P}_h \vec{f}\|_{L^2(\Omega)}^2 dt \right)^{1/2} &\left(\int_{I_m} (t - t_{m-1}) \|\nabla A_h^{-1} \partial_t \vec{u}_{\tau h}\|_{L^2(\Omega)}^2 dt \right)^{1/2} \end{aligned}$$

and canceling, we have

$$\int_{I_m} (t - t_{m-1}) \|\nabla A_h^{-1} \partial_t \vec{u}_{\tau h}\|_{L^2(\Omega)}^2 dt \leq \int_{I_m} (t - t_{m-1}) \|\nabla \vec{u}_{\tau h} - \nabla A_h^{-1} \mathbb{P}_h \vec{f}\|_{L^2(\Omega)}^2 dt.$$

Using the following inequality

$$\int_{I_m} \|v_k\|^2 dt \leq C\tau_m^{-1} \int_{I_m} (t - t_{m-1}) \|v_k\|^2 dt \quad \forall v_k \in \mathcal{P}_q(I_m; \vec{V}_h), \quad (36)$$

which can be easily verified by a usual scaling argument, we obtain

$$\begin{aligned} \int_{I_m} \|\nabla A_h^{-1} \partial_t \vec{u}_{\tau h}\|_{L^2(\Omega)}^2 dt &\leq C\tau_m^{-1} \int_{I_m} (t - t_{m-1}) \|\nabla A_h^{-1} \partial_t \vec{u}_{\tau h}\|_{L^2(\Omega)}^2 dt \\ &\leq C\tau_m^{-1} \int_{I_m} (t - t_{m-1}) \| -\nabla \vec{u}_{\tau h} + \nabla A_h^{-1} \mathbb{P}_h \vec{f} \|_{L^2(\Omega)}^2 dt \\ &\leq C \int_{I_m} \| -\nabla \vec{u}_{\tau h} + \nabla A_h^{-1} \mathbb{P}_h \vec{f} \|_{L^2(\Omega)}^2 dt. \end{aligned}$$

Summing over m and using the estimate for the gradient (35), we obtain

$$\sum_{m=1}^M \|\nabla A_h^{-1} \partial_t \vec{u}_{\tau h}\|_{L^2(I_m \times \Omega)}^2 \leq C \left(\|\nabla A_h^{-1} \mathbb{P}_h \vec{f}\|_{L^2(I \times \Omega)}^2 + \|\mathbb{P}_h \vec{u}_0\|_{L^2(\Omega)}^2 \right). \quad (37)$$

It remains to estimate the jump terms. We test (26) with $\vec{v}_{\tau h} |_{I_m} = [A_h^{-1} \vec{u}_{kh}]_{m-1}$. On each time interval I_m this time we obtain,

$$\|\nabla [A_h^{-1} \vec{u}_{\tau h}]_{m-1}\|_{L^2(\Omega)}^2 = (\vec{f} + A_h \vec{u}_{\tau h} + \partial_t \vec{u}_{\tau h}, [A_h^{-1} \vec{u}_{\tau h}]_{m-1})_{I_m \times \Omega}.$$

Integrating by parts and using the Cauchy-Schwarz and geometric-arithmetic mean inequalities

$$\begin{aligned} &(\vec{f} + A_h \vec{u}_{\tau h} + \partial_t \vec{u}_{\tau h}, [A_h^{-1} \vec{u}_{\tau h}]_{m-1})_{I_m \times \Omega} \\ &= (\nabla A_h^{-1} \mathbb{P}_h \vec{f} + \nabla \vec{u}_{\tau h} + \nabla A_h^{-1} \partial_t \vec{u}_{\tau h}, [\nabla A_h^{-1} \vec{u}_{\tau h}]_{m-1})_{I_m \times \Omega} \\ &\leq \|\nabla A_h^{-1} \mathbb{P}_h \vec{f} + \nabla \vec{u}_{\tau h} + \nabla A_h^{-1} \partial_t \vec{u}_{\tau h}\|_{L^2(I_m \times \Omega)} \|[\nabla A_h^{-1} \vec{u}_{\tau h}]_{m-1}\|_{L^2(I_m \times \Omega)} \\ &\leq \frac{\tau_m}{2} \|\nabla A_h^{-1} \mathbb{P}_h \vec{f} + \nabla \vec{u}_{\tau h} + \nabla A_h^{-1} \partial_t \vec{u}_{\tau h}\|_{L^2(I_m \times \Omega)}^2 + \frac{1}{2\tau_m} \|[\nabla A_h^{-1} \vec{u}_{\tau h}]_{m-1}\|_{L^2(I_m \times \Omega)}^2 \\ &= \frac{\tau_m}{2} \|\nabla A_h^{-1} \mathbb{P}_h \vec{f} + \nabla \vec{u}_{\tau h} + \nabla A_h^{-1} \partial_t \vec{u}_{\tau h}\|_{L^2(I_m \times \Omega)}^2 + \frac{1}{2} \|[\nabla A_h^{-1} \vec{u}_{\tau h}]_{m-1}\|_{L^2(\Omega)}^2, \end{aligned}$$

where we used that $[\nabla A_h^{-1} \vec{u}_{\tau h}]_{m-1}$ is constant in time on I_m . Thus, we obtain

$$\tau_m^{-1} \|[\nabla A_h^{-1} \vec{u}_{\tau h}]_{m-1}\|_{L^2(\Omega)}^2 \leq \|\nabla A_h^{-1} \mathbb{P}_h \vec{f} + \nabla \vec{u}_{\tau h} + \nabla A_h^{-1} \partial_t \vec{u}_{\tau h}\|_{L^2(I_m \times \Omega)}^2.$$

Summing over m and using (35) and (37)

$$\sum_{m=1}^M \tau_m^{-1} \|[\nabla A_h^{-1} \vec{u}_{\tau h}]_{m-1}\|_{L^2(\Omega)}^2 \leq C \left(\|\nabla A_h^{-1} \mathbb{P}_h \vec{f}\|_{L^2(I \times \Omega)}^2 + \|\mathbb{P}_h \vec{u}_0\|_{L^2(\Omega)}^2 \right). \quad (38)$$

Combining (35), (37), and (38), we obtain the result. $\square$

5 Stationary Stokes results

For the best approximation type error estimates in the next section, we need to review basic results on the stationary Stokes projection and on the discrete Leray projection.

5.1 Discrete Stokes projection

We introduce an analogue of the Ritz projection for the stationary Stokes problem $(R_h^S(\vec{w}, \varphi), R_h^{S,p}(\vec{w}, \varphi)) \in \vec{X}_h \times M_h$ of $(\vec{w}, \varphi) \in H_0^1(\Omega)^d \times L^2(\Omega)$ given by the relation

$$(\nabla(\vec{w} - R_h^S(\vec{w}, \varphi)), \nabla \vec{v}_h) - (\varphi - R_h^{S,p}(\vec{w}, \varphi), \nabla \cdot \vec{v}_h) = 0, \quad \forall \vec{v}_h \in \vec{X}_h, \quad (39a)$$

$$(\nabla \cdot (\vec{w} - R_h^S(\vec{w}, \varphi)), q_h) = 0, \quad \forall q_h \in M_h. \quad (39b)$$

Remark 2. If $\vec{w}$ is discrete divergence free, i.e., $(\nabla \cdot \vec{w}, q_h) = 0$ for all $q_h \in M_h$, then we have $R_h^S(\vec{w}, \varphi) \in \vec{V}_h$. We will use this projection operator only for such $\vec{w}$. In this case the same projection operator is defined, e.g., in [12].

Proposition 8. The Stokes projection is stable in H^1 norm, i.e.

$$\|\nabla R_h^S(\vec{w}, \varphi)\|_{L^2(\Omega)} \leq C (\|\nabla \vec{w}\|_{L^2(\Omega)} + \|\varphi\|_{L^2(\Omega)})$$

and if $\vec{w} \in H^2(\Omega)^d$ and $\varphi \in H^1(\Omega)$, the following error estimate holds

$$\|\nabla(\vec{w} - R_h^S(\vec{w}, \varphi))\|_{L^2(\Omega)} \leq Ch (\|\nabla^2 \vec{w}\|_{L^2(\Omega)} + \|\nabla \varphi\|_{L^2(\Omega)}).$$

Proposition 9. If Ω is convex, then

$$\|\vec{w} - R_h^S(\vec{w}, \varphi)\|_{L^2(\Omega)} \leq Ch (\|\nabla \vec{w}\|_{L^2(\Omega)} + \|\varphi\|_{L^2(\Omega)})$$

and

$$\|\vec{w} - R_h^S(\vec{w}, \varphi)\|_{L^2(\Omega)} \leq Ch^2 (\|\nabla^2 \vec{w}\|_{L^2(\Omega)} + \|\nabla \varphi\|_{L^2(\Omega)}).$$

For the proofs of the above results we refer to [7], Propositions 4.14 and 4.18.

5.2 Discrete Leray projection

We need the restriction on the mesh and H^2 regularity such that the Leray projection $\mathbb{P}_h$ is stable in $H^1(\Omega)$ norm. Such result is shown for quasi-uniform meshes in [4].

Lemma 10 (Stability of the Leray projection in H^1 norm). *Let Ω be convex and the triangulations $\{\mathcal{T}_h\}$ be quasi-uniform. Then, there exist a constant C such that*

$$\|\nabla \mathbb{P}_h \vec{v}\|_{L^2(\Omega)} \leq C \|\nabla \vec{v}\|_{L^2(\Omega)} \quad \forall \vec{v} \in \vec{V}^1.$$

Proof. The proof of this results is given in the Appendix of [4]. For the completeness, we repeat the argument here. Let $\vec{v} \in \vec{V}^1$ and consider $\vec{v}_h = R_h^S(\vec{v}, 0) \in \vec{V}_h$. Then by Proposition 8

$$\|\nabla \vec{v}_h\|_{L^2(\Omega)} \leq C\|\nabla \vec{v}\|_{L^2(\Omega)} + \|0\|_{L^2(\Omega)} = C\|\nabla \vec{v}\|_{L^2(\Omega)}.$$

On the other hand by Proposition 9,

$$\|\vec{v} - \vec{v}_h\|_{L^2(\Omega)} \leq Ch\|\nabla \vec{v}\|_{L^2(\Omega)}.$$

As a result, using the inverse inequality and the above two estimates, we obtain

$$\begin{aligned} \|\nabla \mathbb{P}_h \vec{v}\|_{L^2(\Omega)} &\leq \|\nabla (\mathbb{P}_h \vec{v} - \vec{v}_h)\|_{L^2(\Omega)} + \|\nabla \vec{v}_h\|_{L^2(\Omega)} \\ &\leq Ch^{-1}\|\mathbb{P}_h \vec{v} - \vec{v}_h\|_{L^2(\Omega)} + C\|\nabla \vec{v}\|_{L^2(\Omega)} \\ &\leq Ch^{-1}\|\mathbb{P}_h(\vec{v} - \vec{v}_h)\|_{L^2(\Omega)} + C\|\nabla \vec{v}\|_{L^2(\Omega)} \\ &\leq Ch^{-1}\|\vec{v} - \vec{v}_h\|_{L^2(\Omega)} + C\|\nabla \vec{v}\|_{L^2(\Omega)} \\ &\leq C\|\nabla \vec{v}\|_{L^2(\Omega)}. \end{aligned}$$

□

Remark 3. *The stability of the discrete Leray projection is a delicate matter. As operator, $\mathbb{P}_h$ is well-defined for $L^2(\Omega)^d$ and even $L^1(\Omega)^d$ functions. In such a case, the following stability result can be obtained in fractional norms (cf. [14, Lemma 3.1])*

$$\|\mathbb{P}_h \vec{v}\|_{H^s(\Omega)} \leq C\|\vec{v}\|_{H^s(\Omega)} \quad \forall s \in [0, \frac{1}{2}), \quad \forall \vec{v} \in H^s(\Omega)^d, \quad (40)$$

and can not be extended to any $s \geq 1/2$ (cf. [14, Remark 3.1]).

Remark 4. *The stability of the Leray projection (40) is the only result that requires restriction on a mesh.*

6 Main results

Now we state our main best type approximation results.

6.1 L^2 error estimates

The first result establishes the following error estimate in $L^2(I \times \Omega)$ norm.

Theorem 11. *Let $\vec{f} \in L^2(I; L^2(\Omega)^d)$ and $\vec{u}_0 \in \vec{V}^1$. Let $(\vec{u}, p)$ be the solution of (11) and $(\vec{u}_{\tau h}, p_{\tau h})$ solve the respective finite element problem (28). Then, for any $\vec{\chi} \in X_\tau^q(\vec{V}_h)$, there holds*

$$\|\vec{u} - \vec{u}_{\tau h}\|_{L^2(I \times \Omega)} \leq C \left(\|\vec{u} - \vec{\chi}\|_{L^2(I \times \Omega)} + \|\vec{u} - \pi_\tau \vec{u}\|_{L^2(I \times \Omega)} + \|\vec{u} - R_h^S(\vec{u}, p)\|_{L^2(I \times \Omega)} \right). \quad (41)$$

Proof. The proof essentially follows by a duality argument and Theorem 6.

Consider the following dual problem

$$\begin{aligned} -\partial_t \vec{g}(t, \vec{x}) - \Delta \vec{g}(t, \vec{x}) + \nabla \lambda(t, \vec{x}) &= \vec{u}_{\tau h}, & (t, \vec{x}) &\in I \times \Omega, \\ \nabla \cdot \vec{g}(t, \vec{x}) &= 0, & (t, \vec{x}) &\in I \times \Omega, \\ \vec{g}(t, \vec{x}) &= 0, & (t, \vec{x}) &\in I \times \partial\Omega, \\ \vec{g}(T, x) &= 0, & \vec{x} &\in \Omega. \end{aligned} \quad (42)$$

The corresponding finite element approximation $(\vec{g}_{\tau h}, \lambda_{\tau h}) \in X_\tau^q(\vec{X}_h \times M_h)$ is given by

$$B((\vec{v}_{\tau h}, q_{\tau h}), (\vec{g}_{\tau h}, \lambda_{\tau h})) = (\vec{u}_{\tau h}, \vec{v}_{\tau h})_{I \times \Omega} \quad \forall (\vec{v}_{\tau h}, q_{\tau h}) \in X_\tau^q(\vec{X}_h \times M_h). \quad (43)$$

By the Galerkin orthogonality from Proposition 5, we have

$$\begin{aligned} \|\vec{u}_{\tau h}\|_{L^2(I \times \Omega)}^2 &= (\vec{u}_{\tau h}, \vec{u}_{\tau h})_{I \times \Omega} = B((\vec{u}_{\tau h}, p_{\tau h}), (\vec{g}_{\tau h}, \lambda_{\tau h})) = B((\vec{u}, p), (\vec{g}_{\tau h}, \lambda_{\tau h})) \\ &= - \sum_{m=1}^M (\vec{u}, \partial_t \vec{g}_{\tau h})_{I_m \times \Omega} + (\nabla \vec{u}, \nabla \vec{g}_{\tau h})_{I \times \Omega} - (p, \nabla \cdot \vec{g}_{\tau h})_{I \times \Omega} - \sum_{m=1}^M (\vec{u}_m^-, [\vec{g}_{\tau h}]_m)_\Omega \\ &= J_1 + J_2 + J_3 + J_4, \end{aligned}$$

where we have used the dual representation of the bilinear form B from (30). In the last sum we set $\vec{g}_{\tau h, M+1} = 0$ so that $[\vec{g}_{\tau h}]_M = -\vec{g}_{\tau h, M}$. Applying the Cauchy-Schwarz inequality and Theorem 6, we obtain

$$\begin{aligned} J_1 &\leq \sum_{m=1}^M \|\vec{u}\|_{L^2(I_m \times \Omega)} \|\partial_t \vec{g}_{\tau h}\|_{L^2(I_m \times \Omega)} \\ &\leq \left(\sum_{m=1}^M \|\vec{u}\|_{L^2(I_m \times \Omega)}^2 \right)^{1/2} \left(\sum_{m=1}^M \|\partial_t \vec{g}_{\tau h}\|_{L^2(I_m \times \Omega)}^2 \right)^{1/2} \leq C \|\vec{u}\|_{L^2(I \times \Omega)} \|\vec{u}_{\tau h}\|_{L^2(I \times \Omega)}. \end{aligned}$$

To treat J_4 , we use the projection defined in (31), the Cauchy-Schwarz inequality and the inverse inequality for time discrete functions and Theorem 6, we obtain

$$\begin{aligned}
J_4 &= - \sum_{m=1}^M (\vec{u}_m^-, [\vec{g}_{\tau h}]_m)_\Omega = - \sum_{m=1}^M ((\pi_\tau \vec{u})_m, [\vec{g}_{\tau h}]_m)_\Omega \\
&\leq \left(\sum_{m=1}^M \tau_m \|(\pi_\tau \vec{u})_m\|_{L^2(\Omega)}^2 \right)^{1/2} \left(\sum_{m=1}^M \tau_m^{-1} \|[\vec{g}_{\tau h}]_{m-1}\|_{L^2(\Omega)}^2 \right)^{1/2} \\
&\leq \left(\sum_{m=1}^M \tau_m \|\pi_\tau \vec{u}\|_{L^\infty(I_m; L^2(\Omega))}^2 \right)^{1/2} \left(\sum_{m=1}^M \tau_m \|\tau_m^{-1} [\vec{g}_{\tau h}]_{m-1}\|_{L^2(\Omega)}^2 \right)^{1/2} \\
&\leq C \left(\sum_{m=1}^M \|\pi_\tau \vec{u}\|_{L^2(I_m \times \Omega)}^2 \right)^{1/2} \left(\sum_{m=1}^M \tau_m \|\tau_m^{-1} [\vec{g}_{\tau h}]_{m-1}\|_{L^2(\Omega)}^2 \right)^{1/2} \\
&\leq C \|\pi_\tau \vec{u}\|_{L^2(I \times \Omega)} \|\vec{u}_{\tau h}\|_{L^2(I \times \Omega)}.
\end{aligned}$$

For $J_2 + J_3$ we can argue by using the projection R_h^S defined in (39). Then, we have

$$\begin{aligned}
J_2 + J_3 &= (\nabla \vec{u}, \nabla \vec{g}_{\tau h})_{I \times \Omega} - (p, \nabla \cdot \vec{g}_{\tau h})_{I \times \Omega} \\
&= (\nabla R_h^S(\vec{u}, p), \nabla \vec{g}_{\tau h})_{I \times \Omega} - (R_h^{S,p}(\vec{u}, p), \nabla \cdot \vec{g}_{\tau h})_{I \times \Omega} = (\nabla R_h^S(\vec{u}, p), \nabla \vec{g}_{\tau h})_{I \times \Omega},
\end{aligned} \tag{44}$$

where the last term vanishes, since $\vec{g}_{\tau h}$ is discretely divergence-free. Here and in what follows, the projection $(R_h^S, R_h^{S,p})$ is applied to time dependent functions $(\vec{u}, p)$ point-wise in time. Since $\nabla \cdot \vec{u}(t) = 0$ for almost all $t \in I$ we have $R_h^S(\vec{u}(t), p(t)) \in \vec{V}_h$, cf. (2). With this we can use the definition of the discrete Stokes operator A_h resulting in

$$\begin{aligned}
(\nabla R_h^S(\vec{u}, p), \nabla \vec{g}_{\tau h})_{I \times \Omega} &= (R_h^S(\vec{u}, p), A_h \vec{g}_{\tau h})_{I \times \Omega} \\
&\leq \|R_h^S(\vec{u}, p)\|_{L^2(I \times \Omega)} \|A_h \vec{g}_{\tau h}\|_{L^2(I \times \Omega)} \\
&\leq (\|\vec{u}\|_{L^2(I \times \Omega)} + \|\vec{u} - R_h^S(\vec{u}, p)\|_{L^2(I \times \Omega)}) \|\vec{u}_{\tau h}\|_{L^2(I \times \Omega)}.
\end{aligned} \tag{45}$$

Combining the estimates for J_1 , J_2 , J_3 and J_4 , we conclude

$$\|\vec{u}_{\tau h}\|_{L^2(I \times \Omega)} \leq C \left(\|\vec{u}\|_{L^2(I \times \Omega)} + \|\pi_\tau \vec{u}\|_{L^2(I \times \Omega)} + \|\vec{u} - R_h^S(\vec{u}, p)\|_{L^2(I \times \Omega)} \right).$$

Using that the Galerkin method is invariant on $X_\tau^q(\vec{V}_h \times M_h)$, by replacing $\vec{u}$ and $\vec{u}_{\tau h}$ with $\vec{u} - \vec{\chi}$ and $\vec{u}_{\tau h} - \vec{\chi}$ for any $\vec{\chi} \in X_\tau^q(\vec{V}_h)$, and using the triangle inequality we complete the proof of the theorem. $\square$

If the exact solution is sufficiently smooth then the above result easily leads to optimal convergence rates.

Corollary 12. *Let Ω be convex, $\vec{f} \in L^2(I, L^2(\Omega)^d)$ and $\vec{u}_0 \in \vec{V}^1$. Let $(\vec{u}, p)$ be the solution of (11) and $(\vec{u}_{\tau h}, p_{\tau h})$ solve the respective finite element problem (28). Assume*

in addition the approximation Assumptions 1-4 hold. Then,

$$\|\vec{u} - \vec{u}_{\tau h}\|_{L^2(I \times \Omega)} \leq C (\tau + h^2) \left(\|\vec{f}\|_{L^2(I \times \Omega)} + \|\vec{u}_0\|_{\vec{V}^1} \right).$$

Proof. From Theorem 11, we need to estimate three terms. The temporal error is estimated by (32) resulting in

$$\|\vec{u} - \pi_\tau \vec{u}\|_{L^2(I \times \Omega)} \leq C\tau \|\partial_t \vec{u}\|_{L^2(I; L^2(\Omega))}.$$

The spatial error using Proposition 9 is estimated by

$$\|\vec{u} - R_h^S(\vec{u}, p)\|_{L^2(I \times \Omega)} \leq Ch^2 (\|\nabla^2 \vec{u}\|_{L^2(I \times \Omega)} + \|\nabla p\|_{L^2(I \times \Omega)}).$$

Similarly, choosing $\vec{\chi} = P_\tau R_h^S(\vec{u}, p)$, where P_τ is the orthogonal L^2 projection onto X_τ^q , by the triangle inequality we obtain

$$\begin{aligned} \|\vec{u} - \vec{\chi}\|_{L^2(I \times \Omega)} &\leq \|\vec{u} - P_\tau \vec{u}\|_{L^2(I \times \Omega)} + \|P_\tau(\vec{u} - R_h^S(\vec{u}, p))\|_{L^2(I \times \Omega)} \\ &\leq C\tau \|\partial_t \vec{u}\|_{L^2(I \times \Omega)} + Ch^2 \|\nabla^2 \vec{u}\|_{L^2(I \times \Omega)}. \end{aligned}$$

Using Corollary 3, we obtain the result. $\square$

6.2 H^1 error estimates

The second result establishes the optimal error estimates $L^2(I; H^1(\Omega)^d)$ norm on quasi-uniform meshes.

Theorem 13. Let $\vec{f} \in L^2(I; L^2(\Omega)^d)$ and $\vec{u}_0 \in \vec{V}^1$. Let $(\vec{u}, p)$ be the solution of (11) and $(\vec{u}_{\tau h}, p_{\tau h})$ solve the respective finite element problem (28) on a family of quasi-uniform triangulations $\{\mathcal{T}_h\}$. Then, for any $\vec{\chi} \in X_\tau^q(\vec{V}_h)$, there holds

$$\begin{aligned} \|\nabla(\vec{u} - \vec{u}_{\tau h})\|_{L^2(I \times \Omega)} &\leq C \left(\|\nabla(\vec{u} - \vec{\chi})\|_{L^2(I \times \Omega)} + \|\nabla(\vec{u} - \pi_\tau \vec{u})\|_{L^2(I \times \Omega)} \right. \\ &\quad \left. + \|\nabla(\vec{u} - R_h^S(\vec{u}, p))\|_{L^2(I \times \Omega)} \right). \end{aligned} \quad (47)$$

Proof. The proof is similar to the proof for the $L^2(I; L^2(\Omega)^d)$ norm. This time, we consider the following dual problem

$$\begin{aligned} -\partial_t \vec{g}(t, \vec{x}) - \Delta \vec{g}(t, \vec{x}) + \nabla \lambda(t, \vec{x}) &= A_h \vec{u}_{\tau h}, & (t, \vec{x}) &\in I \times \Omega, \\ \nabla \cdot \vec{g}(t, \vec{x}) &= 0, & (t, \vec{x}) &\in I \times \Omega, \\ \vec{g}(t, \vec{x}) &= 0, & (t, \vec{x}) &\in I \times \partial\Omega, \\ \vec{g}(T, x) &= 0, & \vec{x} &\in \Omega. \end{aligned} \quad (48)$$

The corresponding finite element approximation $(\vec{g}_{\tau h}, \lambda_{\tau h}) \in X_\tau^q(\vec{X}_h \times M_h)$ is given by

$$B((\vec{v}_{\tau h}, q_{\tau h}), (\vec{g}_{\tau h}, \lambda_{\tau h})) = (A_h \vec{u}_{\tau h}, \vec{v}_{\tau h})_{I \times \Omega} \quad \forall (\vec{v}_{\tau h}, q_{\tau h}) \in X_\tau^q(\vec{X}_h \times M_h). \quad (49)$$

By the Galerkin orthogonality from Proposition 5, we have

$$\begin{aligned}
\|\nabla \vec{u}_{\tau h}\|_{L^2(I \times \Omega)}^2 &= (A_h \vec{u}_{\tau h}, \vec{u}_{\tau h})_{I \times \Omega} = B((\vec{u}_{\tau h}, p_{\tau h}), (\vec{g}_{\tau h}, \lambda_{\tau h})) = B((\vec{u}, p), (\vec{g}_{\tau h}, \lambda_{\tau h})) \\
&= - \sum_{m=1}^M (\vec{u}, \partial_t \vec{g}_{\tau h})_{I_m \times \Omega} + (\nabla \vec{u}, \nabla \vec{g}_{\tau h})_{I \times \Omega} - (p, \nabla \cdot \vec{g}_{\tau h}) - \sum_{m=1}^M (\vec{u}_m^-, [\vec{g}_{\tau h}]_m)_\Omega \\
&= J_1 + J_2 + J_3 + J_4,
\end{aligned}$$

where we have used the dual representation of the bilinear form B from (30). In the last sum we set $\vec{g}_{\tau h, M+1} = 0$ so that $[\vec{g}_{\tau h}]_M = -\vec{g}_{\tau h, M}$. Applying the Cauchy-Schwarz inequality and Theorem 7, we obtain

$$\begin{aligned}
J_1 &= - \sum_{m=1}^M (\mathbb{P}_h \vec{u}, \partial_t \vec{g}_{\tau h})_{I_m \times \Omega} \\
&= - \sum_{m=1}^M (\nabla \mathbb{P}_h \vec{u}, \nabla A_h^{-1} \partial_t \vec{g}_{\tau h})_{I_m \times \Omega} \\
&\leq \sum_{m=1}^M \|\nabla \mathbb{P}_h \vec{u}\|_{L^2(I_m \times \Omega)} \|\nabla A_h^{-1} \partial_t \vec{g}_{\tau h}\|_{L^2(I_m \times \Omega)} \\
&\leq \left(\sum_{m=1}^M \|\nabla \mathbb{P}_h \vec{u}\|_{L^2(I_m \times \Omega)}^2 \right)^{1/2} \left(\sum_{m=1}^M \|\nabla A_h^{-1} \partial_t \vec{g}_{\tau h}\|_{L^2(I_m \times \Omega)}^2 \right)^{1/2} \\
&\leq C \|\nabla \vec{u}\|_{L^2(I \times \Omega)} \|\nabla \vec{u}_{\tau h}\|_{L^2(I \times \Omega)},
\end{aligned}$$

where in the last step we used the stability of $\mathbb{P}_h$ in $H^1(\Omega)^d$ for divergence free functions from $H_0^1(\Omega)^d$.

To treat J_4 , we use the projection defined in (31), the Cauchy-Schwarz inequality and the inverse inequality for time discrete function and to obtain

$$\begin{aligned}
J_4 &= - \sum_{m=1}^M (\vec{u}_m^-, [\vec{g}_{\tau h}]_m)_\Omega = - \sum_{m=1}^M (\pi_\tau \mathbb{P}_h \vec{u}_m, [\vec{g}_{\tau h}]_m)_\Omega \\
&= - \sum_{m=1}^M (\pi_\tau \nabla \mathbb{P}_h \vec{u}_m, [\nabla A_h^{-1} \vec{g}_{\tau h}]_m)_\Omega \\
&\leq \left(\sum_{m=1}^M \tau_m \|\pi_\tau \nabla \mathbb{P}_h \vec{u}_m\|_{L^2(\Omega)}^2 \right)^{1/2} \left(\sum_{m=1}^M \tau_m^{-1} \|[\nabla A_h^{-1} \vec{g}_{\tau h}]_{m-1}\|_{L^2(\Omega)}^2 \right)^{1/2} \\
&\leq \left(\sum_{m=1}^M \tau_m \|\pi_\tau \nabla \mathbb{P}_h \vec{u}\|_{L^\infty(I_m; L^2(\Omega))}^2 \right)^{1/2} \left(\sum_{m=1}^M \tau_m \|\tau_m^{-1} [\nabla A_h^{-1} \vec{g}_{\tau h}]_{m-1}\|_{L^2(\Omega)}^2 \right)^{1/2} \\
&\leq C \left(\sum_{m=1}^M \|\pi_\tau \nabla \mathbb{P}_h \vec{u}\|_{L^2(I_m \times \Omega)}^2 \right)^{1/2} \left(\sum_{m=1}^M \tau_m \|\tau_m^{-1} [\nabla A_h^{-1} \vec{g}_{\tau h}]_{m-1}\|_{L^2(\Omega)}^2 \right)^{1/2} \\
&\leq C \|\pi_\tau \nabla \mathbb{P}_h \vec{u}\|_{L^2(I \times \Omega)} \|\nabla \vec{u}_{\tau h}\|_{L^2(I \times \Omega)} \\
&\leq C \|\pi_\tau \nabla \vec{u}\|_{L^2(I \times \Omega)} \|\nabla \vec{u}_{\tau h}\|_{L^2(I \times \Omega)},
\end{aligned}$$

where again in the last step we used the stability of $\mathbb{P}_h$ in $H^1(\Omega)^d$ for divergence free functions from $H_0^1(\Omega)^d$.

For $J_2 + J_3$ we can argue by using the projection R_h^S defined in (39). Then since $\vec{g}_{\tau h}$ is discretely divergence-free, we have

$$\begin{aligned}
J_2 + J_3 &= (\nabla \vec{u}, \nabla \vec{g}_{\tau h})_{I \times \Omega} - (p, \nabla \cdot \vec{g}_{\tau h})_{I \times \Omega} \\
&= (\nabla R_h^S(\vec{u}, p), \nabla \vec{g}_{\tau h})_{I \times \Omega} - (R_h^{S,p}(\vec{u}, p), \nabla \cdot \vec{g}_{\tau h})_{I \times \Omega} = (\nabla R_h^S(\vec{u}, p), \nabla \vec{g}_{\tau h})_{I \times \Omega} \\
&\leq \|\nabla R_h^S(\vec{u}, p)\|_{L^2(I \times \Omega)} \|\nabla \vec{g}_{\tau h}\|_{L^2(I \times \Omega)} \\
&\leq C (\|\nabla \vec{u}\|_{L^2(I \times \Omega)} + \|\nabla(\vec{u} - R_h^S(\vec{u}, p))\|_{L^2(I \times \Omega)}) \|\nabla \vec{u}_{\tau h}\|_{L^2(I \times \Omega)}.
\end{aligned}$$

Combining the estimates for J_1 , J_2 , J_3 and J_4 , we conclude

$$\|\nabla \vec{u}_{\tau h}\|_{L^2(I \times \Omega)} \leq C \left(\|\nabla \vec{u}\|_{L^2(I \times \Omega)} + \|\pi_\tau \nabla \vec{u}\|_{L^2(I \times \Omega)} + \|\nabla(\vec{u} - R_h^S(\vec{u}, p))\|_{L^2(I \times \Omega)} \right).$$

Using that the Galerkin method is invariant on $X_\tau^q(\vec{V}_h \times M_h)$, by replacing $\vec{u}$ and $\vec{u}_{\tau h}$ with $\vec{u} - \vec{\chi}$ and $\vec{u}_{\tau h} - \vec{\chi}$ for any $\vec{\chi} \in X_\tau^q(\vec{V}_h)$, and using the triangle inequality we complete the proof of the theorem. $\square$

Corollary 14. *Let Ω be convex, $\vec{f} \in L^2(I, L^2(\Omega)^d)$ and $\vec{u}_0 \in \vec{V}^1$. Let $(\vec{u}, p)$ be the solution of (11) and $(\vec{u}_{\tau h}, p_{\tau h})$ solve the respective finite element problem (28). Assume in addition the approximation Assumptions 1-4 hold. Then,*

$$\|\nabla(\vec{u} - \vec{u}_{\tau h})\|_{L^2(I \times \Omega)} \leq C \left(\tau^{1/2} + h \right) \left(\|\vec{f}\|_{L^2(I \times \Omega)} + \|\vec{u}_0\|_{\vec{V}^1} \right).$$

Proof. The proof is analogous to the proof of Corollary 12. The main difference is to use the estimate

$$\|\nabla(\vec{u} - \pi_\tau \vec{u})\|_{L^2(I \times \Omega)} \leq C\tau^{1/2} \left(\|\vec{f}\|_{L^2(I \times \Omega)} + \|\vec{u}_0\|_{\vec{V}^1} \right)$$

from Lemma 3.13 from [16]. $\square$

7 Inhomogeneous heat equation

The above results naturally hold for the scalar parabolic equation

$$\begin{aligned} u_t(t, x) - \Delta u(t, x) &= f(t, x), & (t, x) \in I \times \Omega, \\ u(t, x) &= 0, & (t, x) \in I \times \partial\Omega, \\ u(0, x) &= u_0(x), & x \in \Omega. \end{aligned} \tag{50}$$

All arguments stay almost unchanged. We only need to replace A and A_h with $-\Delta$ and $-\Delta_h$, the discrete Leray projection $\mathbb{P}_h$ with the L^2 -projection P_h , and the Stokes projection $R_h^S(\vec{u}, p)$ with the Ritz (elliptic) projection R_h . The mesh restrictions in the case of H^1 norm estimates can be relaxed. The only technical requirement is the stability of the L^2 -projection in H^1 norm, i.e.

$$\|\nabla P_h u\|_{L^2(\Omega)} \leq C \|\nabla u\|_{L^2(\Omega)}.$$

Such result is shown for locally quasi-uniform meshes [3], on adaptive meshes obtained by bisection method in 2D [1, 9], and also on more general meshes in any space dimensions [6].

In this situation the best type approximation result take the form

$$\|u - u_{\tau h}\|_{L^2(I \times \Omega)} \leq C \left(\|u - \chi\|_{L^2(I \times \Omega)} + \|R_h u - u\|_{L^2(I \times \Omega)} + \|\pi_\tau u - u\|_{L^2(I \times \Omega)} \right), \tag{51}$$

and since the Ritz projection R_h is stable in H^1 norm,

$$\|\nabla(u - u_{\tau h})\|_{L^2(I \times \Omega)} \leq C \left(\|\nabla(u - \chi)\|_{L^2(I \times \Omega)} + \|\nabla(\pi_\tau u - u)\|_{L^2(I \times \Omega)} \right), \tag{52}$$

for any space-time fully discrete function χ . Similarly, to the Stokes problem, under the additional assumption of convexity of Ω and some approximation properties of the discrete spaces, we easily derive optimal error estimates of the form

$$\|u - u_{\tau h}\|_{L^2(I; H^s(\Omega))} \leq C \left(\tau^{1-s/2} + h^{2-s} \right) \left(\|f\|_{L^2(I \times \Omega)} + \|u_0\|_{H^1(\Omega)} \right), \quad s = 0, 1.$$

References

- [1] BANK, R. & YSERENTANT, H., *On the H^1 -stability of the L_2 -projection onto finite element spaces*, Numer. Math. **126**, 361-381 (2014), <https://doi.org/10.1007/s00211-013-0562-4>

- [2] BEHRINGER, N., VEXLER, B. & LEYKEKHMAN, D., *Fully discrete best-approximation-type estimates in $L^\infty(I; L^2(\Omega)^d)$ for finite element discretizations of the transient Stokes equations*, IMA Journal Of Numerical Analysis, **43**, 852–880 (2023), <https://doi.org/10.1093/imanum/drac009>,
- [3] BRAMBLE, J., PASCIAK, J. & STEINBACH, O., *On the stability of the L^2 projection in $H^1(\Omega)$* , Math. Comp. **71**, 147–156 (electronic) (2002), <http://dx.doi.org/10.1090/S0025-5718-01-01314-X>
- [4] CHRYSAFINOS, K. & WALKINGTON, N., *Discontinuous Galerkin approximations of the Stokes and Navier-Stokes equations*, Math. Comp. **79**, 2135–2167 (2010)
- [5] DAUGE, M., *Stationary Stokes and Navier-Stokes systems on two- or three-dimensional domains with corners. I. Linearized equations.*, SIAM J. Math. Anal. **20**, 74–97 (1989)
- [6] DIENING, L., STORN, J. & AND TSCHERPEL, T., *On the Sobolev and L^p -stability of the L^2 -projection.*, SIAM J. Math. Anal. **59**, 2571–2607 (2021)
- [7] ERN, A. & GUERMOND, J., *Theory and practice of finite elements.*, Springer-Verlag, New York, 2004, <https://doi-org.eaccess.ub.tum.de/10.1007/978-1-4757-4355-5>
- [8] GALDI, G., *An introduction to the mathematical theory of the Navier-Stokes equations*, Springer, 2011, Steady-state problems
- [9] GASPOZ, F., HEINE, C. & SIEBERT, K., *Optimal grading of the newest vertex bisection and H^1 -stability of the L_2 -projection*, IMA J. Numer. Anal. **36**, 1217–1241 (2016), <https://doi.org/10.1093/imanum/drv044>
- [10] GIRAULT, V., NOCHETTO, R. & SCOTT, L., *Maximum-norm stability of the finite element Stokes projection*, J. Math. Pures Appl. (9), **84**, 279–330 (2005), <https://doi-org.eaccess.ub.tum.de/10.1016/j.matpur.2004.09.017>
- [11] GIRAULT, V. & SCOTT, L., *A quasi-local interpolation operator preserving the discrete divergence*, Calcolo, **40**, 1–19 (2003)
- [12] GIRAULT, V., NOCHETTO, R. & SCOTT, L., *Max-norm estimates for Stokes and Navier-Stokes approximations in convex polyhedra*, Numer. Math. **131**, 771–822 (2015)
- [13] GIRAULT, V. & RAVIART, P., *Finite element methods for Navier-Stokes equations.*, Springer-Verlag, (1986)
- [14] GUERMOND, J. & PASCIAK, J., *Stability of discrete Stokes operators in fractional Sobolev spaces*, J. Math. Fluid Mech. **10**, 588–610 (2008)
- [15] KELLOGG, R. & OSBORN, J., *A regularity result for the Stokes problem in a*

- convex polygon*, J. Functional Analysis, **21**, 397-431 (1976)
- [16] KUNISCH, K., PIEPER, K. & VEXLER, B., *Measure valued directional sparsity for parabolic optimal control problems*, SIAM J. Control Optim. **52**, 3078-3108 (2014), <https://doi.org/10.1137/140959055>
 - [17] LEYKEKHMAN, D. & VEXLER, B., *A priori error estimates for the fully discrete Galerkin approximations of state constrained optimal control problems governed by transient Stokes problem*, in preparation
 - [18] MEIDNER, D. & VEXLER, B., *A priori error estimates for space-time finite element discretization of parabolic optimal control problems. I. Problems without control constraints*, SIAM J. Control Optim. **47**, 1150-1177 (2008), <https://doi-org.eaccess.ub.tum.de/10.1137/070694016>
 - [19] TEMAM, R., *Navier-Stokes equations. Theory and numerical analysis*, North-Holland Publishing Co., 1977, Studies in Mathematics and its Applications, Vol. 2
 - [20] VEXLER, B. & WAGNER, J., $L^2(I; L^2(\Omega))$ Error estimates for dG-cG finite element discretizations of the 2D instationary Navier-Stokes equations, submitted