IMA Journal of Numerical Analysis (2022) **00**, 1–29
https://doi.org/10.1093/imanum/drac009

Fully discrete best-approximation-type estimates in $L^\infty(I; L^2(\Omega)^d)$ for finite element discretizations of the transient Stokes equations

NIKLAS BEHRINGER AND BORIS VEXLER*
Center for Mathematical Sciences, Technical University of Munich, 85748 Garching by Munich, Germany
*Corresponding author: vexler@ma.tum.de

AND

DMITRIY LEYKEKHMAN
Department of Mathematics, University of Connecticut, Storrs, CT 06269, USA

[Received on 22 July 2021; revised on 29 December 2021]

In this article, we obtain an optimal best-approximation-type result for fully discrete approximations of the transient Stokes problem. For the time discretization, we use the discontinuous Galerkin method and for the spatial discretization we use standard finite elements for the Stokes problem satisfying the discrete inf-sup condition. The analysis uses the technique of discrete maximal parabolic regularity. The results require only natural assumptions on the data and do not assume any additional smoothness of the solutions.

Keywords: transient Stokes; discontinuous Galerkin method; finite elements; best approximation; pointwise error estimates; *a priori* estimates.

1. Introduction

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

$$\partial_t \vec{u} - \Delta \vec{u} + \nabla p = \vec{f} \quad \text{in } I \times \Omega, \tag{1.1a}$$

$$\nabla \cdot \vec{u} = 0 \quad \text{in } I \times \Omega, \tag{1.1b}$$

$$\vec{u} = \vec{0} \quad \text{on } I \times \partial\Omega, \tag{1.1c}$$

$$\vec{u}(0) = \vec{u}_0 \quad \text{in } \Omega. \tag{1.1d}$$

Throughout this work we assume that $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$ is a bounded polygonal/polyhedral Lipschitz domain, $\mathcal{T} > 0$ and $I = (0, \mathcal{T}]$. We will require some (weak) assumptions on the data, which essentially allow for a weak formulation including both velocity and pressure and for $\vec{u} \in C(\bar{I}; L^2(\Omega)^d)$. We consider fully discrete approximations of problem (1.1), 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. Our goal is to obtain best-approximation-type results that do not involve any additional regularity assumptions on the solution beyond the regularity that follows directly from the assumed data above. Such results are important in the analysis of PDE-constrained optimal control problems that we have in mind. We refer, e.g., to Meidner *et al.* (2011), where such estimates are required for numerical analysis of an optimal control problem constrained by the heat equation with state constraints pointwise in time.

Our main result is of the following form:

$$\|\vec{u} - \vec{u}_{\tau h}\|_{L^\infty(I;L^2(\Omega))} \leqslant C\ell_\tau \left(\|\vec{u} - \vec{v}_{\tau h}\|_{L^\infty(I;L^2(\Omega))} + \|\vec{u} - R_h^S(\vec{u}, p)\|_{L^\infty(I;L^2(\Omega))}\right), \tag{1.2}$$

where $\vec{u}_{\tau h}$ is the fully discrete finite element approximation of the velocity $\vec{u}$, $\vec{v}_{\tau h}$ is an arbitrary function from the finite element approximation of the velocity spaces $X_\tau^w(\vec{V}_h)$, R_h^S is the Ritz projection for the stationary Stokes problem and ℓ_τ is a logarithmic term, explicitly given in the statements of the results; see Corollary 6.4.

Result (1.2) links 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 $X_\tau^w(\vec{V}_h)$, as well as the approximation of the stationary Stokes problem in $\vec{V}_h$. Such results go hand in hand with only natural assumptions on the problem data and thus are desirable in applications. For this result, we do not require additional regularity of the domain allowing, e.g., for reentrant corners and edges. Moreover, we do not require the mesh to be quasi-uniform or 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). The application of (1.2) in such cases would require corresponding results for the stationary Stokes problem to estimate $\vec{u} - R_h^S(\vec{u}, p)$; see Remark 7.1.

Under the additional assumption of convexity of Ω and some approximation properties of the discrete spaces, we prove error estimates of the form

$$\|\vec{u} - \vec{u}_{\tau h}\|_{L^\infty(I;L^2(\Omega))} \leqslant C\ell_\tau \left(\tau + h^2\right) \left(\|\vec{f}\|_{L^\infty(I;L^2(\Omega))} + \|\vec{u}_0\|_{\vec{V}^2}\right),$$

where $\vec{V}^2$ is an appropriate space introduced in the next section. This estimate seems to be optimal (probably up to logarithmic terms) with respect to both the assumed regularity of the data and the order of convergence.

In the case of the heat equation, a similar estimate with respect to $L^\infty(I; L^2(\Omega))$ is derived in Meidner *et al.* (2011) and for a nonautonomous parabolic problem in Leykekhman & Vexler (2018, Theorem 4.5). For corresponding estimates in the maximum norm in the case of the heat equation we refer to Schatz *et al.* (1980), Eriksson & Johnson (1995), Meidner *et al.* (2011) and Leykekhman & Vexler (2016), and for the maximum norm of the gradient to Thomée *et al.* (1989), Leykekhman & Wahlbin (2008) and Leykekhman & Vexler (2017b). Further, results are also available in the case of discretization only in space. For an overview and respective references, we refer to Leykekhman & Vexler (2016, 2017b).

We are not aware of any best approximation max-norm estimates in time *and* space for the nonstationary Stokes problem (1.1) in the literature. A result for the fully discrete problem in the form of $L^\infty(I; L^2(\Omega)^d)$ estimates based on discontinuous Galerkin methods is provided in Chrysafinos & Walkington (2010), including an overview of related results for (semi)discrete problems based on other discretization approaches. Recently, the numerical behavior of a stabilized discontinuous Galerkin scheme for the Stokes problem has been analyzed in Ahmed *et al.* (2017). Furthermore, there are results for the fully discrete Navier–Stokes problem under moderate regularity assumptions in Heywood & Rannacher (1990). Here we focus on an approach via a discontinuous Galerkin time-stepping scheme similar to the approach in Chrysafinos & Walkington (2010) and Leykekhman & Vexler (2017a).

However, all the results mentioned above differ from ours in an essential way. We give a more detailed comparison of our result and the existing results from the literature in Section 7.

Our main technical tools are continuous and discrete maximal parabolic regularity results. On the continuous level, we use the estimate

$$\|\partial_t \vec{u}\|_{L^s(I;L^2(\Omega))} + \|A\vec{u}\|_{L^s(I;L^2(\Omega))} + \|p\|_{L^s(I;L^2(\Omega))} \leqslant \frac{Cs^2}{s-1}\|\vec{f}\|_{L^s(I;L^2(\Omega))}$$

for $\vec{f} \in L^s(I;L^2(\Omega)^d), \vec{u}_0 = 0$, $1 < s < \infty$ and A being the Stokes operator (2.1); see Proposition 2.6 and Theorem 2.10 for the details and also for the formulation in the case $\vec{u}_0 \neq 0$. This estimate holds on a general Lipschitz domain Ω. Assuming in addition the convexity of Ω, we have

$$\|\partial_t \vec{u}\|_{L^s(I;L^2(\Omega))} + \|\vec{u}\|_{L^s(I;H^2(\Omega))} + \|\nabla p\|_{L^s(I;L^2(\Omega))} \leqslant \frac{Cs^2}{s-1}\|\vec{f}\|_{L^s(I;L^2(\Omega))};$$

see Remark 2.7 and Corollary 2.11. On the discrete level, we provide the corresponding estimates that hold even in the limit cases $s = 1$ and $s = \infty$ at the expense of a logarithmic term. In a way, we extend the discrete maximal parabolic regularity results from Leykekhman & Vexler (2017a) to the Stokes problem. The resulting estimate is

$$\|\partial_t \vec{u}_{\tau h}\|_{L^s(I;L^2(\Omega))} + \|A_h \vec{u}_{\tau h}\|_{L^s(I;L^2(\Omega))} \leqslant C \ln \frac{\mathcal{T}}{\tau}\|\vec{f}\|_{L^s(I;L^2(\Omega))},$$

where A_h is the discrete Stokes operator; see Theorem 5.2 for details and the precise formulation. Under the convexity assumption for the domain Ω, similar to the continuous case we also obtain

$$\|\Delta_h \vec{u}_{\tau h}\|_{L^s(I;L^2(\Omega))} + \|\nabla p_{\tau h}\|_{L^s(I;L^2(\Omega))} \leqslant \ln \frac{\mathcal{T}}{\tau}\|\vec{f}\|_{L^s(I;L^2(\Omega))},$$

where Δ_h is the discrete Laplace operator; see Remark 5.4 and Theorem 8.2 for details.

In the next section we introduce a framework of function spaces for the treatment of the stationary and transient Stokes problem, the Stokes operator and the resolvent problem. Moreover, we discuss the weak formulation and regularity issues for (1.1). In Section 3 we discuss the spatial discretization, introduce respective discrete spaces, operators and prove a discrete resolvent estimate. In Section 4 we present a full discretization of (1.1) and show discrete smoothing and discrete maximal regularity results for the velocity in Section 5 based on the operator calculus discussed for the heat equation in Leykekhman & Vexler (2017a). This allows us to prove best approximation results for the velocity in Section 6. In Section 7 we apply the best-approximation-type results to prove error estimates and compare these results to the existing results in the literature. Finally, in Section 8 we explore an expansion of the discrete maximal parabolic estimates to the pressure.

2. Results on the continuous level

In this section we introduce the function spaces we require for the analysis of (1.1) and state some of the main properties of these spaces. In the later sections we adopt a technique based on discrete maximal parabolic regularity from Leykekhman & Vexler (2017a), where we used an operator calculus for $-\Delta$

and its finite element analog $-\Delta_h$. In order to modify the corresponding results, we will introduce the continuous and the discrete Stokes operator. Furthermore, we will require analysis for the resolvent of these operators. In our presentation we follow the notation and presentation of Guermond & Pasciak (2008, 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 $L^2_0(\Omega)$ 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:

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

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

$$\vec{V}^0 = \{\vec{v} \in L^2(\Omega)^d \mid \nabla \cdot \vec{v} = 0 \text{ and } \vec{u} \cdot \vec{n} = 0 \text{ on } \partial\Omega\},$$

where n is the outer unit normal vector to $\partial\Omega$, and

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

by Galdi (2011, Theorems III.2.3 and III.4.1).

We define the vector-valued Laplace operator

$$-\Delta\colon D(\Delta) \to L^2(\Omega)^d,$$

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. Temam, 1977, Chapter I and Theorem 1.4 and Galdi, 2011, Theorem III.1.1):

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

As usual, we define the Helmholtz projection $\mathbb{P}\colon L^2(\Omega)^d \to \vec{V}^0$ (often 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 $A\colon \vec{V}^2 \to \vec{V}^0$ as

$$A = -\mathbb{P}\Delta|_{\vec{V}^2}. \tag{2.1}$$

The operator A is a self-adjoint, densely defined and positive definite operator on $\vec{V}^0$. We note that $D(A) = \vec{V}^2$. There holds

$$(A\vec{v}, \vec{v}) = \|\nabla\vec{v}\|^2_{L^2(\Omega)} \geqslant \lambda_0\|\vec{v}\|^2_{L^2(\Omega)}, \quad \vec{v} \in \vec{V}^2,$$

where $\lambda_0 > 0$ is the smallest eigenvalue of the Laplace operator $-\Delta$ given by

$$\lambda_0 = \inf_{v\in H^1_0(\Omega)} \frac{\|\nabla v\|^2_{L^2(\Omega)}}{\|v\|^2_{L^2(\Omega)}}. \tag{2.2}$$

Similar to the Laplace operator, for convex polyhedral domains Ω we have the following H^2 regularity bound due to Kellogg & Osborn (1976) and Dauge (1989):

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

2.2 *Stokes resolvent problem*

The key to our analysis is the spectral representation of the semigroup generated by A. For that, we consider the Stokes resolvent problem for $\vec{f} \in L^2(\Omega)^d$,

$$z\vec{u} - \Delta\vec{u} + \nabla p = \vec{f} \quad \text{in } \Omega, \tag{2.3a}$$

$$\nabla\cdot\vec{u} = 0 \quad \text{in } \Omega, \tag{2.3b}$$

$$\vec{u} = \vec{0} \quad \text{on } \partial\Omega. \tag{2.3c}$$

Here, $z \in \Sigma_{\theta,\bar{\omega}}$, which is defined as

$$\Sigma_{\theta,\bar{\omega}} = \{c \in \mathbb{C} | c \neq \bar{\omega} \text{ and } |\arg(c - \bar{\omega})| < \theta\}.$$

The solution $(\vec{u}, p)$ to (2.3) is a complex-valued function in $H^1_0(\Omega)^d \times L^2_0(\Omega)$ as complex-valued function spaces with a hermitian inner product. In our situation, we are interested in the case of $\theta \in (\pi/2, \pi)$ and $\bar{\omega} \in [-\lambda_0, 0]$ with $\lambda_0 > 0$ from (2.2).

Proposition 2.1 The operator A is sectorial. In particular, for every $\theta \in (\pi/2, \pi)$, there exists a constant $C = C_\theta$ such that for all $z \in \Sigma_{\theta,\bar{\omega}}$ with $\bar{\omega} \in [-\lambda_0, 0]$ and $(\vec{u}, p)$ being the solution of (2.3) with $\vec{f} \in L^2(\Omega)^d$ there holds the following resolvent estimate:

$$\|\vec{u}\|_{L^2(\Omega)} \leqslant \frac{C}{|z - \bar{\omega}|}\|\mathbb{P}\vec{f}\|_{L^2(\Omega)}. \tag{2.4}$$

Proof. It is straightforward to check that every self-adjoint positive operator, which is densely defined on a Hilbert space is sectorial. This applies to the operator $A - \bar{\omega}\,\mathrm{Id}$ for all $\bar{\omega} \in [-\lambda_0, 0]$, which results in the resolvent estimate (2.4). Below, we provide a direct proof for the discrete version of the operator A; see Lemma 3.2, which is also applicable here. $\square$

Using the Stokes operator A from (2.1) one can rewrite the resolvent estimate (2.4) as

$$\|(z+A)^{-1}\mathbb{P}\vec{f}\|_{L^2(\Omega)} \leqslant \frac{C}{|z-\bar{\omega}|}\|\mathbb{P}\vec{f}\|_{L^2(\Omega)}.$$

REMARK 2.2 The resolvent estimates in L^p norms for $p \neq 2$ are also known. For example, for $d = 3$ on Lipschitz domains, Shen (2012) has shown a resolvent estimate for some interval of p satisfying $|*|1/p - 1/2 < 1/6 + \varepsilon$ for $\varepsilon > 0$. On smooth C^3 domains, it is known to hold even for $p = \infty$ (cf. Abe *et al.*, 2015). However, the extension to nonsmooth convex domains is still an open problem and it even appears in a collection of open problems (cf. Maz'ya, 2018, Problem 66).

2.3 *Weak formulation and regularity*

In this section we discuss the weak formulation and the regularity of the transient Stokes problem (1.1). 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 $L^2(I;L^2(\Omega)^d)$ and in $L^2(I;L^2(\Omega)^d)$ is denoted by $(\cdot,\cdot)_{I\times\Omega}$. We will also use the notation $(f,g)_{I\times\Omega}$ for the corresponding integral for $f \in L^s(I;L^2(\Omega)^d)$ and $g \in L^{s'}(I;L^2(\Omega)^d)$ with $1 \le s \le \infty$ and the dual exponent s'.

For the application of Galerkin finite element methods in space and time, we will require a space-time weak formulation of the transient Stokes equations with respect to both variables, velocity $\vec{u}$ and pressure p. In a standard variational setting, e.g., with $f \in L^2(I;(\vec{V}^1)')$ or $f \in L^1(I;L^2(\Omega)^d)$, this is not possible, since only distributional pressure can be expected in general; see Remark 2.5 below. Therefore, we will first introduce the (standard) weak formulation on the divergence-free space, and then we discuss regularity issues and introduce a velocity–pressure weak formulation based on a slightly stronger assumption on the data.

PROPOSITION 2.3 Let $\vec{f} \in L^1(I;L^2(\Omega)^d)$ 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 $\partial_t\vec{u} \in L^1(I;\vec{V}^0) + L^2(I;(\vec{V}^1)')$ fulfilling $\vec{u}(0) = \vec{u}_0$ and

$$\langle\partial_t\vec{u},\vec{v}\rangle + (\nabla\vec{u},\nabla\vec{v})_{I\times\Omega} = (\vec{f},\vec{v})_{I\times\Omega} \quad \text{for all } \vec{v} \in L^2(I;\vec{V}^1) \cap L^\infty(I;\vec{V}^0). \tag{2.5}$$

Here, $\langle\partial_t\vec{u},\vec{v}\rangle$ for $\partial_t\vec{u} = \vec{w}_1 + \vec{w}_2 \in L^1(I;\vec{V}^0) + L^2(I;(\vec{V}^1)')$ and $\vec{v} \in L^2(I;\vec{V}^1) \cap L^\infty(I;\vec{V}^0)$ is understood as

$$\langle\partial_t\vec{u},\vec{v}\rangle = (\vec{w}_1,\vec{v})_{I\times\Omega} + \langle\vec{w}_2,\vec{v}\rangle_{L^2(I;(\vec{V}^1)')\times L^2(I;\vec{V}^1)}.$$

There holds the estimate

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

Proof. For the existence of the solution under the stated assumptions and with the corresponding regularity we refer to Temam (1977, Chapter III, Theorem 1.1) and its extension to the case $\vec{f} \in L^1(I;L^2(\Omega)^d)$ on page 264. The notion of the solution in Temam (1977) is formulated in the almost everywhere sense on I, from which formulation (2.5) follows directly by integration in time. The uniqueness of $\vec{u}$ solving (2.5) is also obtained in the standard manner choosing $\vec{v} = \vec{u}$ for $\vec{f} = 0$ and $\vec{u}_0 = 0$. □

REMARK 2.4 Another possibility to formulate the notion of the weak solution is to assume $\vec{f} \in L^2(I; (\vec{V}^1)')$. Since we require in the sequel the additional assumption $\vec{f} \in L^s(I; L^2(\Omega)^d)$ for some $s > 1$ we prefer to use the formulation from Proposition 2.3.

REMARK 2.5 Under the assumptions from Proposition 2.3 the existence of the corresponding pressure p can be shown only in the following distributional sense. There exists $P \in C(\bar{I}, L^2(\Omega))$ such that the Stokes system holds in the distributional sense for $\vec{u}$ solving (2.5) and $p = \partial_t P$. In particular, one cannot expect in general $p \in L^1(I \times \Omega)$; cf. Temam (1977, Chapter III, p. 267).

In the following we will discuss some additional regularity for the solution. On the one hand, we need slightly more regularity in order to be able to introduce the pressure p as a function; cf. Remark 2.5. Moreover, additional regularity beyond $\vec{u} \in C(\bar{I}; L^2(\Omega)^d)$ is required if we use the best approximation result 1.2 for providing (optimal) error estimates; see Section 7. It is well known, cf. again Temam (1977), that in the sense of Proposition 2.3, equation 2.5 can be understood as an abstract parabolic problem

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

with the Stokes operator A defined in (2.1).

Furthermore, note that by Proposition 2.1, the operator A is sectorial and thus a generator of an analytic semigroup (Lunardi, 1995, Definition 2.0.1, 2.0.2). For the Hilbert space setting, it has then been shown in de Simon (1964) that this is equivalent to having a maximal regularity estimate of the form

$$\|\partial_t \vec{u}\|_{L^s(I;L^2(\Omega))} + \|A\vec{u}\|_{L^s(I;L^2(\Omega))} \leqslant C_s \|\vec{f}\|_{L^s(I;L^2(\Omega))} \tag{2.6}$$

for problem (1.1) with $\vec{u}_0 = 0, 1 < s < \infty$ and $\vec{f} \in L^s(I; L^2(\Omega)^d)$. For more details, we refer to Sohr (2014, Chapter IV, Theorem 1.6.3). In Section 5 we derive a respective estimate for a fully discrete version of (2.12), a so-called discrete maximal parabolic regularity result based on ideas from Leykekhman & Vexler (2017a).

For proving (optimal) error estimates in Section 7 we will use the maximal parabolic estimate (2.6) for $s \to \infty$. To this end, we will need precise dependence of the constant C_s on s from Ashyralyev & Sobolevskii (1994, Chapter 1, eq. (3.9), Theorem 3.2). Moreover, we require this regularity result also for the case of nonhomogeneous initial conditions.

To state this result, we consider the space of initial data $\vec{V}^0_{1-\frac{1}{s}}$ for $1 < s < \infty$ as in Ashyralyev & Sobolevskii (1994, Chapter 1, Section 3.3). The Banach space $\vec{V}^0_{1-\frac{1}{s}}$ with the norm

$$\|\vec{v}_0\|_{\vec{V}^0_{1-\frac{1}{s}}} = \left(\int_0^1 \|A \exp(-tA)\vec{v}_0\|^s_{\vec{V}^0} \, \mathrm{d}t \right)^{1/s} + \|\vec{v}_0\|_{\vec{V}^0} \tag{2.7}$$

contains all functions $\vec{v}_0 \in \vec{V}^0$ such that for a solution $\vec{u}$ to the transient Stokes problem with right-hand side $\vec{f} = 0$ and initial data $\vec{u}_0 = \vec{v}_0$ it holds that $A\vec{u} \in L^s(I; \vec{V}^0)$.

Proposition 2.6 Let $1 < s < \infty$, $\vec{f} \in L^s(I; L^2(\Omega)^d)$ and $\vec{u}_0 \in \vec{V}^0_{1-\frac{1}{s}}$. Then the solution $\vec{u}$ to the problem

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

fulfills $\partial_t \vec{u}, A\vec{u} \in L^s(I; L^2(\Omega)^d)$. Moreover, there is a constant C independent on s, f and u_0 such that

$$\|\partial_t \vec{u}\|_{L^s(I;L^2(\Omega))} + \|A\vec{u}\|_{L^s(I;L^2(\Omega))} \leqslant \frac{Cs^2}{s-1}\Big(\|\vec{f}\|_{L^s(I;L^2(\Omega))} + \|\vec{u}_0\|_{\vec{V}^0_{1-\frac{1}{s}}}\Big).$$

Proof. This follows from Ashyralyev & Sobolevskii (1994, Chapter 1, Theorems 3.2, 3.7) since A is the generator of an analytic semigroup. □

Remark 2.7 If the domain is polyhedral/polygonal and convex then Proposition 2.6 provides $\vec{u} \in L^s(I; H^2(\Omega)^d)$ and the estimate

$$\|\partial_t \vec{u}\|_{L^s(I;L^2(\Omega))} + \|\vec{u}\|_{L^s(I;H^2(\Omega))} \leqslant \frac{Cs^2}{s-1}\Big(\|\vec{f}\|_{L^s(I;L^2(\Omega))} + \|\vec{u}_0\|_{\vec{V}^0_{1-\frac{1}{s}}}\Big).$$

Remark 2.8 There holds $\vec{V}^1 \hookrightarrow \vec{V}^0_{\frac{1}{2}}$. This follows from the fact that for the homogeneous problem ($\vec{f} = 0$) with $\vec{u}_0 \in \vec{V}^1$ the following estimate holds:

$$\|A\vec{u}\|_{L^2(I;L^2(\Omega))} \leqslant 2\|A^{\frac{1}{2}}\vec{u}_0\|_{L^2(\Omega)} = 2\|\vec{u}_0\|_{\vec{V}^1}.$$

The above inequality is stated, e.g., in Sohr (2014, Chapter IV, Theorem 1.5.2). For the representation of the norm on $\vec{V}^1$ by $A^{\frac{1}{2}}$, see, e.g., Sohr (2014, Chapter III, Lemma 2.2.1). Therefore, for the range $1 < s \leq 2$, it is sufficient to assume $\vec{u}_0 \in \vec{V}^1$ for the estimate in Proposition 2.6.

Remark 2.9 If $u_0 \in \vec{V}^2$, there holds for every $1 < s < \infty$,

$$\|\vec{u}_0\|_{\vec{V}^0_{1-\frac{1}{s}}} \leqslant \|A\vec{u}_0\|_{L^2(\Omega)}.$$

We can argue as follows. Since $\vec{u}_0 \in \vec{V}^2$, we have that A commutes with $\exp(-tA)$ (cf. Sohr, 2014, Chapter II, eq. (3.2.19)) and due to the boundedness of $\exp(-tA)$ in the operator norm (cf. Sohr, 2014, Chapter IV, eq. (1.5.8)), we can conclude using Definition 2.13,

$$\begin{aligned} \|\vec{u}_0\|_{\vec{V}^0_{1-\frac{1}{s}}} &= \left(\int_0^1 \|A\exp(-tA)\vec{u}_0\|^s_{\vec{V}^0}\,\mathrm{d}t\right)^{1/s} + \|\vec{u}_0\|_{\vec{V}^0} \\ &\leqslant \left(\int_0^1 \|\exp(-tA)\|^s_{\vec{V}^0\to\vec{V}^0}\|A\vec{u}_0\|^s_{\vec{V}^0}\,\mathrm{d}t\right)^{1/s} + \|\vec{u}_0\|_{\vec{V}^0} \leqslant C\|A\vec{u}_0\|_{L^2(\Omega)}. \end{aligned}$$

Therefore, we have the following version of the maximal parabolic regularity estimate:

$$\|\partial_t \vec{u}\|_{L^s(I;L^2(\Omega))} + \|A\vec{u}\|_{L^s(I;L^2(\Omega))} \leqslant \frac{Cs^2}{s-1}\Big(\|\vec{f}\|_{L^s(I;L^2(\Omega))} + \|A\vec{u}_0\|_{L^2(\Omega)}\Big),$$

which we will use in particular for $s \to \infty$.

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 and the assumption $\vec{f} \in L^1(I;L^2(\Omega)^d)$ from Proposition 2.3 is only slightly strengthened to $\vec{f} \in L^s(I;L^2(\Omega)^d)$ for some $s > 1$.

THEOREM 2.10 Let $\vec{f} \in L^s(I;L^2(\Omega)^d)$ for some $1 < s < \infty$ and $\vec{u}_0 \in \vec{V}^0_{1-\frac{1}{s}}$. Then there exists a unique solution $(\vec{u}, p)$ with

$$\vec{u} \in L^2(I;\vec{V}^1) \cap C(\bar{I}, \vec{V}^0), \qquad \partial_t \vec{u}, A\vec{u} \in L^s(I;L^2(\Omega)^d) \quad \text{and} \quad p \in L^s(I;L^2_0(\Omega))$$

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

$$(\partial_t \vec{u}, \vec{v})_{I\times\Omega} + (\nabla\vec{u}, \nabla\vec{v})_{I\times\Omega} - (p, \nabla\cdot\vec{v})_{I\times\Omega} + (\nabla\cdot\vec{u}, \xi)_{I\times\Omega} = (\vec{f}, \vec{v})_{I\times\Omega} \tag{2.8}$$

for all

$$\vec{v} \in L^2(I;H^1_0(\Omega)^d) \cap L^\infty(I;L^2(\Omega)^d) \quad \text{and} \quad \xi \in L^2(I;L^2_0(\Omega)).$$

There holds the estimate

$$\|p\|_{L^s(I;L^2(\Omega))} \leqslant \frac{Cs^2}{s-1}\left(\|\vec{f}\|_{L^s(I;L^2(\Omega))} + \|\vec{u}_0\|_{\vec{V}^0_{1-\frac{1}{s}}}\right).$$

Proof. We take the unique solution $\vec{u}$ with $\partial_t \vec{u}, A\vec{u} \in L^s(I;L^2(\Omega)^d)$ from Proposition 2.6 and $\vec{u} \in L^2(I;\vec{V}^1) \cap C(\bar{I}, \vec{V}_0)$ from Proposition 2.3. To prove the existence of the corresponding pressure p we consider for almost every $t \in I$ the element $\vec{g}(t) \in H^{-1}(\Omega)^d$,

$$\langle \vec{g}(t), w\rangle = (\vec{f}(t), w) - (\partial_t \vec{u}(t), w) - (\nabla\vec{u}(t), \nabla w), \quad w \in H^1_0(\Omega)^d,$$

i.e.,

$$\vec{g}(t) = \vec{f}(t) - \partial_t \vec{u}(t) + \Delta\vec{u}(t) \in H^{-1}(\Omega)^d.$$

This element is well defined due to $\vec{f}(t), \partial_t \vec{u}(t) \in L^2(\Omega)^d$ and $\vec{u}(t) \in \vec{V}^1 \hookrightarrow H^1_0(\Omega)$ for almost every $t \in I$. Moreover, there holds by (2.5),

$$\langle \vec{g}(t), w\rangle = 0 \quad \forall\, w \in \vec{V}^1,$$

for almost all $t \in I$, since (2.5) holds also pointwise almost everywhere. Therefore, we can apply Temam (1977, Chapter I, Proposition 1.1), which ensures the existence of a distribution $p(t)$ with

$$\nabla p(t) = \vec{g}(t) \tag{2.9}$$

in the distributional sense for almost every $t \in I$. By Temam (1977, Chapter I, Proposition 1.2 and Remark 1.4), we have that the gradient operator is an isomorphism from $L^2(\Omega)/\mathbb{R}$ into $H^{-1}(\Omega)^d$ and as a result we have unique $p(t) \in L_0^2(\Omega)$ such that

$$\|p(t)\|_{L^2(\Omega)} \leqslant C\|\vec{g}(t)\|_{H^{-1}(\Omega)}.$$

Using the definition of $\vec{g}$ we obtain $p \in L^s(I; L_0^2(\Omega))$ and

$$\begin{aligned}\|p\|_{L^s(I;L^2(\Omega))} \leqslant C\|\vec{g}\|_{L^s(I;H^{-1}(\Omega))} &\leqslant C\left(\|\vec{f}\|_{L^s(I;L^2(\Omega))} + \|\partial_t\vec{u}\|_{L^s(I;L^2(\Omega))} + \|\nabla\vec{u}\|_{L^s(I;L^2(\Omega))}\right)\\ &\leqslant \frac{Cs^2}{s-1}\left(\|\vec{f}\|_{L^s(I;L^2(\Omega))} + \|\vec{u}_0\|_{\vec{V}^0_{1-\frac{1}{s}}}\right),\end{aligned}$$

where we have used $\|\nabla\vec{v}\|_{L^2(\Omega)} \leq C\|A\vec{v}\|_{L^2(\Omega)}$ for every $\vec{v} \in \vec{V}^1$ and Proposition 2.6. With this regularity we obtain from (2.9) and the definition of $\vec{g}$,

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

for all $\vec{v} \in L^2(I; H_0^1(\Omega)^d) \cap L^\infty(I; L^2(\Omega)^d)$. Furthermore, it holds that

$$(\nabla\cdot\vec{u}, \xi)_{I\times\Omega} = 0 \quad \forall\, \xi \in L^2(I; L_0^2(\Omega))$$

by $\vec{u} \in L^2(I; \vec{V}^1)$. This results in the stated weak formulation. □

COROLLARY 2.11 Let the assumptions of Theorem 2.10 be fulfilled. Let in addition the domain Ω be convex. Then we have $p \in L^s(I, H^1(\Omega))$ and the corresponding estimate holds:

$$\|p\|_{L^s(I;H^1(\Omega))} \leqslant \frac{Cs^2}{s-1}\left(\|\vec{f}\|_{L^s(I;L^2(\Omega))} + \|\vec{u}_0\|_{\vec{V}^0_{1-\frac{1}{s}}}\right).$$

Proof. By convexity of Ω, we obtain $\vec{u} \in L^s(I; H^2(\Omega)^d)$ and the corresponding estimate; see Remark 2.7. Then we have

$$\vec{g}(t) = \vec{f}(t) - \partial_t\vec{u}(t) + \Delta\vec{u}(t) \in L^2(\Omega)^d$$

for almost all $t \in I$ in the notation of the proof of Theorem 2.10. This leads to the desired regularity and to the estimate. □

REMARK 2.12 For regularity beyond these estimates, we want to highlight Heywood & Rannacher (1982, Corollary 2.1), where the authors show that bounds for, e.g., $\nabla^3\vec{u}$, $\partial_{tt}\vec{u}$, go hand in hand with the need for the data $\vec{u}_0$, $\vec{f}$ and initial pressure p_0 (defined as $\lim_{t\to 0} p(t)$) to satisfy a nonlocal compatibility condition for $t \to 0$ at the boundary, which is potentially difficult to verify.

3. Spatial discretization and discrete resolvent estimates

In this section we consider the discrete version of the operators presented in the previous section.

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., they satisfy a uniform discrete inf-sup condition,

$$\sup_{\vec{v}_h \in \vec{X}_h} \frac{(q_h, \nabla \cdot \vec{v}_h)}{\|\nabla \vec{v}_h\|_{L^2(\Omega)}} \geqslant \beta \|q_h\|_{L^2(\Omega)} \quad \forall\, q_h \in M_h, \tag{3.1}$$

with a constant $\beta > 0$ independent of h. We introduce the usual discrete Laplace operator $-\Delta_h \colon \vec{X}_h \to \vec{X}_h$ by

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

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

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

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

$$(\mathbb{P}_h \vec{u}, \vec{v}_h) = (\vec{u}, \vec{v}_h) \quad \forall\, \vec{v}_h \in \vec{V}_h. \tag{3.2}$$

Using $\mathbb{P}_h$, we define the discrete Stokes operator $A_h \colon \vec{V}_h \to \vec{V}_h$ 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

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

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

$$(A_h \vec{v}_h, \vec{v}_h) = (\nabla \vec{v}_h, \nabla \vec{v}_h) \geqslant \lambda_0 \|\vec{v}_h\|^2_{L^2(\Omega)}, \tag{3.3}$$

where λ_0 is the smallest eigenvalue of $-\Delta$; see (2.2). This implies that the eigenvalues of A_h are also positive and bounded from below by λ_0.

Moreover, we define the orthogonal space $\vec{V}_h^{\perp} \subset \vec{X}_h$ as

$$\vec{V}_h^{\perp} = \vec{w}_h \in X_h \mid (\vec{w}_h, \vec{v}_h) = 0 \quad \forall\, \vec{v}_h \in \vec{V}_h.$$

The following classical result, cf., e.g., Girault & Raviart (1986, Chapter II, Theorem 1.1), will be used to provide existence and uniqueness of the fully discrete pressure in the sequel.

LEMMA 3.1 For every $\vec{w}_h \in \vec{V}_h^{\perp}$, there exists a unique $p_h \in M_h$ such that

$$(\vec{w}_h, \vec{v}_h) = (p_h, \nabla \cdot \vec{v}_h) \quad \forall \vec{v}_h \in \vec{X}_h.$$

There holds

$$\|p_h\|_{L^2(\Omega)} \leqslant \frac{1}{\beta} \|\nabla(-\Delta_h)^{-1}\vec{w}_h\|_{L^2(\Omega)}.$$

Proof. Note that we can decompose $\vec{X}_h = \vec{V}_h \oplus \vec{V}_h^{\perp}$ and there holds $\dim \vec{V}_h^{\perp} = \dim M_h$. The equation for p_h can then be equivalently rewritten as

$$p_h \in M_h \quad : \quad (p_h, \nabla \cdot \vec{v}_h) = (\vec{w}_h, \vec{v}_h) \quad \forall \vec{v}_h \in \vec{V}_h^{\perp}.$$

The uniqueness of p_h (as well as the estimate) follows then directly from the inf-sup condition (3.1). The existence follows from uniqueness due to $\dim \vec{V}_h^{\perp} = \dim M_h$. □

3.2 *Discrete resolvent estimate*

For a given $\vec{f} \in L^2(\Omega)^d$ the discrete version of the resolvent problem (2.3) takes the form

$$\vec{u}_h \in \vec{V}_h \quad : \quad z(\vec{u}_h, \vec{v}_h) + (\nabla\vec{u}_h, \nabla\vec{v}_h) = (\vec{f}, \vec{v}_h) \quad \forall \vec{v}_h \in \vec{V}_h, \tag{3.4}$$

which we can also write compactly using the discrete operator as

$$(z + A_h)\vec{u}_h = \mathbb{P}_h\vec{f}. \tag{3.5}$$

Next we establish the discrete resolvent estimate in the $L^2(\Omega)^d$ norm, which is the discrete version of Proposition 2.1.

LEMMA 3.2 For any $\theta \in (\pi/2, \pi)$, there exists a constant $C = C_\theta$ such that for any $\nu \in [0, \lambda_0]$ with $\lambda_0 > 0$ being the smallest eigenvalue of $-\Delta$, see (2.2), it holds that

$$\|\vec{u}_h\|_{L^2(\Omega)} = \|(z + A_h)^{-1}\mathbb{P}_h\vec{f}\|_{L^2(\Omega)} \leqslant \frac{C_\theta}{|z + \nu|}\|\vec{f}\|_{L^2(\Omega)} \quad \forall z \in \Sigma_{\theta,-\nu},$$

where $\vec{u}_h \in \vec{V}_h$ is the solution to (3.5) with right-hand side $\vec{f} \in L^2(\Omega)^d$.

Proof. Testing (3.4) with $\vec{u}_h$, we have

$$(z + \nu)\|\vec{u}_h\|^2_{L^2(\Omega)} + ((-\Delta_h - \nu)\vec{u}_h, \vec{u}_h) = (\vec{f}, \vec{u}_h) \tag{3.6}$$

for any $\nu > 0$. Since $-\Delta_h$ is positive definite with (3.3), we have that $-\Delta_h - \nu$ is still a non-negative operator for $\nu \in [0, \lambda_0]$ and thus $((-\Delta_h - \nu)\vec{u}_h, \vec{u}_h) \geqslant 0$. Since z is restricted to the sector $\Sigma_{\theta,-\nu}$, we can rewrite (3.6) as

$$|z + \nu|e^{i\phi}\|\vec{u}_h\|^2_{L^2(\Omega)} + \delta = (\vec{f}, \vec{u}_h), \tag{3.7}$$

where $\delta \geqslant 0$ and $|\phi| < \theta$. If we multiply (3.7) by $e^{-i\phi/2}$, take the real part and use that $\cos(\theta/2) > 0$, this results in

$$|z+\nu| \|\vec{u}_h\|^2_{L^2(\Omega)} \leqslant \cos(\theta/2)^{-1} |(\vec{f}, \vec{u}_h)| = C_\theta |(\vec{f}, \vec{u}_h)|,$$

which after an application of the Cauchy–Schwarz inequality completes the proof. □

4. 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 Chrysafinos & Walkington (2010). For that we partition $I = (0, \mathscr{T}]$ into subintervals $I_m = (t_{m-1}, t_m]$ of length $\tau_m = t_m - t_{m-1}$, where $0 = t_0 < t_1 < \cdots < t_{M-1} < t_M = \mathscr{T}$. The maximal and minimal time steps are denoted by $\tau = \max_m \tau_m$ and $\tau_{\min} = \min_m \tau_m$, respectively. The time partition fulfills the following assumptions.

1. There are constants $C, \beta > 0$ independent of τ such that
$$\tau_{\min} \geqslant C\tau^{\beta}.$$
2. There is a constant $\kappa > 0$ independent of τ such that for all $m = 1, 2, \ldots, M-1$,
$$\kappa^{-1} \leqslant \frac{\tau_m}{\tau_{m+1}} \leqslant \kappa.$$
3. It holds that $\tau \leqslant \frac{\mathscr{T}}{4}$.

For a given Banach space $\mathcal{B}$ and the order $w \in \mathbb{N}$, we define the semidiscrete space $X^w_\tau(\mathcal{B})$ of piecewise polynomial functions in time as

$$X^w_\tau(\mathcal{B}) = \left\{\vec{v}_\tau \in L^2(I; \mathcal{B}) | \vec{v}_\tau|_{I_m} \in \mathcal{P}_{w,I_m}(\mathcal{B}), m = 1, 2, \ldots, M\right\},$$

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

$$\mathcal{P}_{w,I_m}(\mathcal{B}) = \left\{\vec{v}_\tau : I_m \to \mathcal{B} \middle| \vec{v}_\tau(t) = \textstyle\sum_{j=0}^{w} \vec{v}^j \phi_j(t),\ \vec{v}^j \in \mathcal{B}, j = 0, \ldots, w\right\}.$$

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

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

For later use, we introduce $P_\tau : L^2(I; L^2(\Omega)^d) \to X^w_\tau(L^2(\Omega^d))$ as the L^2 projection in time by

$$(\vec{v} - P_\tau \vec{v}, \vec{w}_\tau)_{I \times \Omega} = 0 \quad \forall \vec{w}_\tau \in X^w_\tau(L^2(\Omega^d)). \tag{4.1}$$

We will use the following two standard properties:

$$\|P_\tau \vec{v}\|_{L^s(I,L^2(\Omega))} \leqslant C\|\vec{v}\|_{L^s(I,L^2(\Omega))} \quad \forall\, \vec{v} \in L^s(I, L^2(\Omega)^d),\ 1 \leqslant s \leqslant \infty \tag{4.2}$$

and

$$\|\vec{v} - P_\tau \vec{v}\|_{L^\infty(I,L^2(\Omega))} \leqslant C\tau^{1-\frac{1}{s}} \|\partial_t \vec{v}\|_{L^s(I,L^2(\Omega))} \quad \forall\, \vec{v} \in W^{1,s}(I, L^2(\Omega)^d),\ 1 \leqslant s \leqslant \infty. \tag{4.3}$$

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 $X^w_\tau(\vec{V}_h)$:

$$\vec{u}_{\tau h} \in X^w_\tau(\vec{V}_h) \ :\ \mathfrak{B}(\vec{u}_{\tau h}, \vec{v}_{\tau h}) = (\vec{f}, \vec{v}_{\tau h})_{I \times \Omega} + (\vec{u}_0, \vec{v}^+_{\tau h,0})_\Omega \quad \forall\, \vec{v}_{\tau h} \in X^w_\tau(\vec{V}_h). \tag{4.4}$$

By a standard argument one can see that this formulation possesses a unique solution (existence follows from uniqueness by the fact that (4.4) is equivalent to a quadratic system of linear equations).

REMARK 4.1 Note that the data $\vec{f}$ and $\vec{u}_0$ in (4.4) 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 (3.2)) without changing the solution. Therefore, this formulation makes sense for a general $\vec{f} \in L^1(I; L^1(\Omega)^d)$ and $\vec{u}_0 \in L^1(\Omega)^d$. However, for the error analysis later on we will require the assumptions from Theorem 2.10, ensuring the Galerkin orthogonality relation; see also Proposition 4.3 below.

The above formulation is not a conforming discretization of the divergence free-formulation (2.5) due to the fact that $X^w_\tau(\vec{V}_h)$ is not a subspace of $L^2(I; \vec{V}^1)$. In order to introduce a velocity pressure discrete formulation (as a discretization of (2.8)) we consider the following bilinear form:

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

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

$$B((\vec{u}_{\tau h}, p_{\tau h}), (\vec{v}_{\tau h}, q_{\tau h})) = (\vec{f}, \vec{v}_{\tau h})_{I \times \Omega} + (\vec{u}_0, \vec{v}^+_{\tau h,0})_\Omega \qquad \forall\, (\vec{v}_{\tau h}, q_{\tau h}) \in X^w_\tau(\vec{X}_h \times M_h). \tag{4.5}$$

We note that for the temporal discretization, we use polynomials of the same order for the velocity and the pressure. The next proposition states the equivalence of the formulations (4.4) and (4.5).

PROPOSITION 4.2 For a solution $(\vec{u}_{\tau h}, p_{\tau h})$ of (4.5), the discrete velocity $\vec{u}_{\tau h}$ fulfills (4.4). Moreover, for a solution $\vec{u}_{\tau h}$ of (4.4), there exists a unique $p_{\tau h} \in X_\tau^w(M_h)$ such that the pair $(\vec{u}_{\tau h}, p_{\tau h})$ fulfills (4.5). In particular, the solution of (4.5) is unique.

Proof. From (4.5) we have $\vec{u}_{\tau h} \in X_\tau^w(\vec{V}_h)$ and so it trivially fulfills (4.4). Now let $\vec{u}_{\tau h} \in X_\tau^w(\vec{V}_h)$ be the solution of (4.4). We define $\vec{w}_{\tau h} \in X_\tau^w(\vec{X}_h)$ by

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

It follows immediately that

$$(\vec{w}_{\tau h}, \vec{v}_{\tau h})_{I \times \Omega} = 0 \quad \forall \vec{v}_{\tau h} \in X_\tau^w(\vec{V}_h)$$

and one obtains $\vec{w}_{\tau h}(t) \in \vec{V}_h^\perp$ for every $t \in I_m, m = 1, 2, \ldots, M$. At the same time, we have globally $\vec{w}_{\tau h} \in X_\tau^w(\vec{V}_h^\perp)$. By Lemma 3.1 we get the existence and uniqueness of the pressure $p_{\tau h}(t) \in M_h$ with

$$(\vec{w}_{\tau h}(t), \vec{v}_h) = (p_{\tau h}(t), \nabla \cdot \vec{v}_h) \quad \forall \vec{v}_h \in \vec{X}_h,$$

for every $t \in I_m, m = 1, 2, \ldots, M$.

Therefore, $p_{\tau h} \in X_\tau^w(M_h)$ and there holds

$$(\vec{w}_{\tau h}, \vec{v}_{\tau h})_{I \times \Omega} = (p_{\tau h}, \nabla \cdot \vec{v}_{\tau h})_{I \times \Omega} \quad \forall \vec{v}_{\tau h} \in X_\tau^w(\vec{X}_h).$$

This completes the proof. $\square$

The next proposition provides the Galerkin orthogonality relation for the velocity pressure discretization (4.5), which is essential for our analysis. Please note that for the velocity formulation (4.4) the Galerkin orthogonality does not hold due to the fact that $X_\tau^w(\vec{V}_h)$ is not a subspace of $L^2(I, \vec{V}^1)$.

PROPOSITION 4.3 Let the assumptions of Theorem 2.10 be fulfilled, i.e., $\vec{f} \in L^s(I; L^2(\Omega)^d)$ for some $s > 1$ and $\vec{u}_0 \in \vec{V}^0_{1-\frac{1}{s}}$. Then there holds, for the solution $(\vec{u}, p)$ of (2.8),

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

and consequently,

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

Proof. In the setting of Theorem 2.10, we have $\vec{u} \in L^2(I, \vec{V}^1) \cap C(\bar{I}, \vec{V}^0)$, $\partial_t \vec{u} \in L^s(I, L^2(\Omega)^d)$ and $p \in L^s(I, L^2(\Omega))$. Therefore, all terms in the bilinear form are well defined. For the test space we have

$$X_\tau^w(\vec{X}_h) \subset L^2(I; H_0^1(\Omega)^d) \cap L^\infty(I; L^2(\Omega)^d) \quad \text{and} \quad X_\tau^w(M_h) \subset L^2(I; L_0^2(\Omega)).$$

Therefore, we can choose $(\vec{v}_{\tau h}, q_{\tau h}) \in X_\tau^w(\vec{X}_h \times M_h)$ as test functions in (2.8). Moreover, all jump terms vanish due to $\vec{u} \in C(\bar{I}, \vec{V}^0)$, and

$$(\vec{u}_0^+, \vec{v}^+_{\tau h,0}) = (\vec{u}_0, \vec{v}^+_{\tau h,0})$$

due to $\vec{u}(0) = \vec{u}_0$. This completes the proof. $\square$

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

$$\begin{aligned} B((\vec{u},p),(\vec{v},q)) = &-\sum_{m=1}^{M}\langle \vec{u}, \partial_t \vec{v}\rangle_{I_m\times\Omega} + (\nabla\vec{u},\nabla\vec{v})_{I\times\Omega} - (p,\nabla\cdot\vec{v})_{I\times\Omega} \\ &+ (\nabla\cdot\vec{u},q)_{I\times\Omega} - \sum_{m=1}^{M-1}(\vec{u}_m^-,[\vec{v}]_m)_\Omega + (\vec{u}_M^-,\vec{v}_M^-)_\Omega, \end{aligned} \tag{4.6}$$

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

5. Fully discrete smoothing and maximal regularity estimates

The goal of this section is to extend the results on the discrete maximal parabolic regularity for the discretization of the heat equation from Leykekhman & Vexler (2017a) to the transient Stokes equations. The results in Leykekhman & Vexler (2017a) rely solely on the resolvent estimates for $-\Delta_h$ and the Dunford–Taylor operator calculus. Since Lemma 3.2 establishes the resolvent estimate for A_h, all the results from Leykekhman & Vexler (2017a) continue to hold for A_h as well. We will state the results below.

The first result is a smoothing estimate for the homogeneous problem ($f = 0$).

THEOREM 5.1 Let $\vec{f} = \vec{0}$, and let $\vec{u}_0 \in L^2(\Omega)^d$. Let $\vec{u}_{\tau h} \in X_\tau^w(\vec{V}_h)$ be the solution to (4.4). Then there holds for $m = 1, 2, \ldots, M$,

$$\|\partial_t\vec{u}_{\tau h}\|_{L^\infty(I_m;L^2(\Omega))} + \|A_h\vec{u}_{\tau h}\|_{L^\infty(I_m;L^2(\Omega))} + \|\tau_m^{-1}[\vec{u}_{\tau h}]_{m-1}\|_{L^2(\Omega)} \leqslant \frac{C}{t_m}\|\mathbb{P}_h\vec{u}_0\|_{L^2(\Omega)}.$$

Here we have $[\vec{u}_{\tau h}]_0 = \vec{u}^+_{\tau h,0} - \mathbb{P}_h\vec{u}_0$.

Proof. The key step in proving this smoothing estimate is the representation of the solution on I_m in the form of the Dunford–Taylor integral (cf. Eriksson *et al.*, 1998, pp. 1321–1322),

$$A_h\vec{u}^-_{\tau,m} = \frac{1}{2\pi \mathrm{i}}\int_\Gamma \prod_{l=1}^{m} r(\tau_l z)A_h R(z,A_h)\,\mathrm{d}z\,\mathbb{P}_h\vec{u}_0 \qquad \text{for } m = 2,\ldots,M,$$

which is an operator equality on $\vec{V}_h$ since $\vec{u}_0$ is replaced by $\mathbb{P}_h\vec{u}_0 \in \vec{V}_h$; cf. Remark 4.1. Here, $r(z)$ is a subdiagonal Padè approximation, which is a rational function with numerator of degree w and denominator of degree $w+1$. The contour Γ is a curve contained in the resolvent set of A_h such that Lemma 3.2 can be applied for $-z \in \Gamma$ and $R(z,A_h)$ the resolvent operator, i.e., $R(z,A_h) = (z - A_h)^{-1}$. Then the proof of

$$\|A_h\vec{u}_{\tau h}\|_{L^\infty(I_m;L^2(\Omega))} \leqslant \frac{C}{t_m}\|\mathbb{P}_h\vec{u}_0\|_{L^2(\Omega)}$$

is the same as for Eriksson *et al.* (1998, Theorem 5.1). The estimates for the time derivative and for the jumps follow as in Leykekhman & Vexler (2017a, Theorem 4 and Theorem 5). □

For the inhomogeneous problem (and $\vec{u}_0 = 0$) we obtain the discrete analog of Proposition 2.6. On the continuous level, the corresponding estimate is true for $1 < s < \infty$. The following discrete maximal parabolic regularity result covers the limit cases $s = 1$ and $s = \infty$ at the expense of a logarithmic term.

THEOREM 5.2 Let $1 \leqslant s \leqslant \infty, \vec{f} \in L^s(I, L^2(\Omega)^d)$ and $\vec{u}_0 = 0$. Let $\vec{u}_{\tau h} \in X^w_\tau(\vec{V}_h)$ be the solution to (4.4). Then for $s < \infty$, there holds

$$\left(\sum_{m=1}^{M} \|\partial_t \vec{u}_{\tau h}\|^s_{L^s(I_m;L^2(\Omega))}\right)^{1/s} + \|A_h \vec{u}_{\tau h}\|_{L^s(I;L^2(\Omega))} + \left(\sum_{m=1}^{M} \tau_m \|\tau_m^{-1}[\vec{u}_{\tau h}]_{m-1}\|^s_{L^2(\Omega)}\right)^{1/s} \leqslant C \ln \frac{\mathcal{T}}{\tau} \|\mathbb{P}_h \vec{f}\|_{L^s(I;L^2(\Omega))}.$$

For $s = \infty$ the estimate takes the form

$$\max_{1 \leqslant m \leqslant M} \|\partial_t \vec{u}_{\tau h}\|_{L^\infty(I_m;L^2(\Omega))} + \|A_h \vec{u}_{\tau h}\|_{L^\infty(I;L^2(\Omega))} + \max_{1 \leqslant m \leqslant M} \|\tau_m^{-1}[\vec{u}_{\tau h}]_{m-1}\|_{L^2(\Omega)} \leqslant C \ln \frac{\mathcal{T}}{\tau} \|\mathbb{P}_h \vec{f}\|_{L^\infty(I;L^2(\Omega))}.$$

Here we have $[\vec{u}_{\tau h}]_0 = \vec{u}^+_{\tau h,0}$.

Proof. The result follows from the smoothing estimate in Theorem 5.1 as in the proof of Leykekhman & Vexler (2017a, Theorems 6–8). □

REMARK 5.3 Due to the stability of the discrete Leray projection $\mathbb{P}_h$ in L^2, we can drop it in the above estimates.

REMARK 5.4 If we assume the domain Ω to be convex and the family of meshes $\{\mathcal{T}_h\}$ to be shape regular and quasi-uniform then there holds

$$\|\Delta_h \vec{v}_h\|_{L^2(\Omega)} \leqslant c\|A_h \vec{v}_h\|_{L^2(\Omega)} \quad \forall \vec{v}_h \in \vec{V}_h$$

by Guermond & Pasciak (2008, Lemma 4.1) or Heywood & Rannacher (1982, Corollary 4.4) and therefore the corresponding estimates hold also for $\|\Delta_h \vec{u}_{\tau h}\|_{L^s(I;L^2(\Omega))}$.

6. Best-approximation-type estimates

The results in Section 5 allow us to show an $L^\infty(I; L^2(\Omega))$ best-approximation-type error estimate for the velocity. In order to state the results, we need to introduce an analog of the Ritz projection for the stationary Stokes problem $(R^S_h(\vec{w}, \varphi), R^{S,p}_h(\vec{w}, \varphi)) \in \vec{X}_h \times M_h$ of $(\vec{w}, \varphi) \in H^1_0(\Omega)^d \times L^2(\Omega)$ given by the relation

$$(\nabla(\vec{w} - R^S_h(\vec{w}, \varphi)), \nabla \vec{v}_h) - (\varphi - R^{S,p}_h(\vec{w}, \varphi), \nabla \cdot \vec{v}_h) = 0 \qquad \forall \vec{v}_h \in \vec{X}_h, \tag{6.1a}$$

$$(\nabla \cdot (\vec{w} - R^S_h(\vec{w}, \varphi)), q_h) = 0 \qquad \forall q_h \in M_h. \tag{6.1b}$$

REMARK 6.1 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 a $\vec{w}$. In this case, the same projection operator is defined, e.g., in Girault *et al.* (2015).

In the following we will make the same assumptions on the data $\vec{f}$ and $\vec{u}_0$ as in Theorem 2.10 in order to use the Galerkin orthogonality relation from Proposition 4.3.

THEOREM 6.2 Let $\vec{f} \in L^s(I; L^2(\Omega)^d)$ for some $s > 1$ and $\vec{u}_0 \in \vec{V}^0_{1-\frac{1}{s}}$. Let $(\vec{u}, p)$ be the solution of (2.8) and $(\vec{u}_{\tau h}, p_{\tau h})$ solve the respective finite element problem (4.5). Then, there holds

$$\|\vec{u}_{\tau h}\|_{L^\infty(I;L^2(\Omega))} \leqslant C \ln \frac{\mathcal{T}}{\tau} \Big(\|\vec{u}\|_{L^\infty(I;L^2(\Omega))} + \|\vec{u} - R_h^S \vec{u}\|_{L^\infty(I;L^2(\Omega))} \Big).$$

Proof. We proceed with a proof along the arguments of Leykekhman & Vexler (2016, Theorem 1). Let $\tilde{t} \in (0, \mathcal{T}]$ and without loss of generality assume $\tilde{t} \in (t_{M-1}, \mathcal{T}]$.

Consider the following dual problem:

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

Here, $\theta \in C^1(I)$ is a regularized delta function (cf. Schatz & Wahlbin, 1995, Appendix A.5) in time with the following properties:

$$\operatorname{supp}\theta \subset (t_{M-1}, \mathcal{T}), \qquad \|\theta\|_{L^1(I_M)} \leqslant C \qquad \text{and} \qquad (\theta, \vec{v}_\tau)_{I_M} = \vec{v}_\tau(\tilde{t}) \quad \forall \vec{v}_\tau \in \mathcal{P}_w(I_M). \tag{6.3}$$

Note that the authors in Schatz & Wahlbin (1995, Appendix A.5) assume $\tilde{t}$ to be an element of an open interval but the argument there can be extended to the case $\tilde{t} = \mathcal{T}$. The corresponding finite element approximation $(\vec{g}_{\tau h}, \lambda_{\tau h}) \in X_\tau^w(\vec{X}_h \times M_h)$ is given by

$$B((\vec{v}_{\tau h}, q_{\tau h}), (\vec{g}_{\tau h}, \lambda_{\tau h})) = \big(\vec{u}_{\tau h}(\tilde{t})\theta, \vec{v}_{\tau h}\big)_{I \times \Omega} \qquad \forall (\vec{v}_{\tau h}, q_{\tau h}) \in X_\tau^w(\vec{X}_h \times M_h).$$

By the Galerkin orthogonality from Proposition 4.3, we have

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

where we have used the dual representation of the bilinear form B from (4.6). 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 Hölder inequality and using $\vec{u} \in C(\bar{I}; L^2(\Omega)^d)$, we obtain

$$
\begin{aligned}
J_1 &\leqslant \sum_{m=1}^{M} \|\vec{u}\|_{L^\infty(I_m;L^2(\Omega))} \|\partial_t \vec{g}_{\tau h}\|_{L^1(I_m;L^2(\Omega))} \leqslant \|\vec{u}\|_{L^\infty(I;L^2(\Omega))} \sum_{m=1}^{M} \|\partial_t \vec{g}_{\tau h}\|_{L^1(I_m;L^2(\Omega))},\\
J_4 &\leqslant \sum_{m=1}^{M} \|\vec{u}_m^-\|_{L^2(\Omega)} \|[\vec{g}_{\tau h}]_{m-1}\|_{L^2(\Omega)} \leqslant \|\vec{u}\|_{L^\infty(I;L^2(\Omega))} \sum_{m=1}^{M} \|[\vec{g}_{\tau h}]_m\|_{L^2(\Omega)}.
\end{aligned}
$$

For $J_2 + J_3$, we can argue by using the projection R_h^S defined in (6.1). Then we have

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

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 $(\vec{u}, p)$ pointwise 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. Remark 6.1. With this we can use the definition of the discrete Stokes operator A_h resulting in

$$
\begin{aligned}
(\nabla R_h^S(\vec{u},p), \nabla \vec{g}_{\tau h})_{I\times\Omega} &= (R_h^S(\vec{u},p), A_h \vec{g}_{\tau h})_{I\times\Omega}\\
&\leqslant \Big(\|\vec{u}\|_{L^\infty(I;L^2(\Omega))} + \|\vec{u} - R_h^S(\vec{u},p)\|_{L^\infty(I;L^2(\Omega))}\Big) \|A_h \vec{g}_{\tau h}\|_{L^1(I;L^2(\Omega))}.
\end{aligned} \tag{6.4}
$$

Combining the estimates we conclude

$$
\begin{aligned}
\|\vec{u}_{\tau h}(\tilde{t})\|_{L^2(\Omega)}^2 &= -\sum_{m=1}^{M} (\vec{u}, \partial_t \vec{g}_{\tau h})_{I_m\times\Omega} + (\nabla R_h^S(\vec{u},p), \nabla \vec{g}_{\tau h})_{I\times\Omega} - \sum_{m=1}^{M} (\vec{u}_m^-, [\vec{g}_{\tau h}]_m)_\Omega\\
&\leqslant C\Big(\|\vec{u}\|_{L^\infty(I;L^2(\Omega))} + \|\vec{u} - R_h^S(\vec{u},p)\|_{L^\infty(I;L^2(\Omega))}\Big)\\
&\quad\times \left(\sum_{m=1}^{M} \|\partial_t \vec{g}_{\tau h}\|_{L^1(I_m;L^2(\Omega))} + \|A_h \vec{g}_{\tau h}\|_{L^1(I;L^2(\Omega))} + \sum_{m=1}^{M} \|[\vec{g}_{\tau h}]_m\|_{L^2(\Omega)}\right)
\end{aligned}
$$

and an application of Theorem 5.2 leads to

$$
\|\vec{u}_{\tau h}(\tilde{t})\|_{L^2(\Omega)}^2 \leqslant C \ln\frac{\mathscr{T}}{\tau}\Big(\|\vec{u}\|_{L^\infty(I;L^2(\Omega))} + \|\vec{u} - R_h^S(\vec{u},p)\|_{L^\infty(I;L^2(\Omega))}\Big) \times \|\vec{u}_{\tau h}(\tilde{t})\|_{L^2(\Omega)} \|\theta\|_{L^1(I_M)}.
$$

Canceling and using that $\|\theta\|_{L^1(I_M)} \leqslant C$, we complete the proof of the theorem. □

The following corollary provides a version of Theorem 6.2 involving the L^2 projection in time P_τ defined in (4.1).

COROLLARY 6.3 Under the conditions of Theorem 6.2, there holds

$$\|\vec{u}_{\tau h}\|_{L^\infty(I;L^2(\Omega))} \leqslant C \ln \frac{\mathcal{T}}{\tau} \Big(\|\vec{u}\|_{L^\infty(I;L^2(\Omega))} + \|\vec{u} - P_\tau R_h^S \vec{u}\|_{L^\infty(I;L^2(\Omega))} \Big).$$

Proof. We obtain this by arguing as in the proof of Theorem 6.2, only the term in (6.4) is estimated differently,

$$\begin{aligned}(\nabla R_h^S(\vec{u},p), \nabla \vec{g}_{\tau h})_{I\times\Omega} &= (R_h^S(\vec{u},p), A_h \vec{g}_{\tau h})_{I\times\Omega} = (P_\tau R_h^S(\vec{u},p), A_h \vec{g}_{\tau h})_{I\times\Omega} \\ &\leqslant \Big(\|\vec{u}\|_{L^\infty(I;L^2(\Omega))} + \|\vec{u} - P_\tau R_h^S(\vec{u},p)\|_{L^\infty(I;L^2(\Omega))} \Big) \|A_h \vec{g}_{\tau h}\|_{L^1(I;L^2(\Omega))},\end{aligned}$$

where we have used definition 4.6 of $\mathcal{P}_\tau$. □

As a corollary from Theorem 6.2, we obtain a best-approximation-type result.

COROLLARY 6.4 Under the conditions of Theorem 6.2, there holds

$$\|\vec{u} - \vec{u}_{\tau h}\|_{L^\infty(I;L^2(\Omega))} \leqslant C \ln \frac{\mathcal{T}}{\tau} \Bigg(\inf_{\vec{v}_{\tau h} \in X_\tau^w(\vec{V}_h)} \|\vec{u} - \vec{v}_{\tau h}\|_{L^\infty(I;L^2(\Omega))} + \|\vec{u} - R_h^S(\vec{u},p)\|_{L^\infty(I;L^2(\Omega))} \Bigg). \quad (6.5)$$

Proof. The desired result follows by considering $(\vec{u} - \vec{v}_{\tau h}, p - q_{\tau h})$ instead of $(\vec{u}, p)$ with arbitrary $(\vec{v}_{\tau h}, q_{\tau h}) \in X_\tau^w(\vec{V}_h \times M_h)$ in the proof of Theorem 6.2. This allows us to replace $(\vec{u}_{\tau h}, p_{\tau h})$ by $(\vec{u}_{\tau h} - \vec{v}_{\tau h}, p_{\tau h} - q_{\tau h})$.

Note that $\vec{u} - \vec{v}_{\tau h}$ is discrete divergence-free and so $R_h^S(\vec{u} - \vec{v}_{\tau h}, p - q_{\tau h})(t) \in \vec{V}_h$ for almost all $t \in I$; see Remark 6.1. Therefore, the argument in the proof of Theorem 6.2 involving the discrete Stokes operator A_h is still valid. Moreover, by the definition of the projection R_h^S, we have

$$R_h^S(\vec{u} - \vec{v}_{\tau h}, p - q_{\tau h}) = R_h^S(\vec{u}, p) - \vec{v}_{\tau h}.$$

As in the proof of Theorem 6.2, then we obtain

$$\|\vec{u}_{\tau h} - \vec{v}_{\tau h}\|_{L^\infty(I;L^2(\Omega))} \leqslant C \ln \frac{\mathcal{T}}{\tau} \left(\|\vec{u} - \vec{v}_{\tau h}\|_{L^\infty(I;L^2(\Omega))} + \|\vec{u} - R_h^S(\vec{u},p)\|_{L^\infty(I;L^2(\Omega))} \right).$$

We complete the proof using

$$\vec{u} - \vec{u}_{\tau h} = \vec{u} - \vec{v}_{\tau h} + \vec{v}_{\tau h} - \vec{u}_{\tau h}$$

and the triangle inequality. □

The error estimate in the next section is based on the following variant of our result.

COROLLARY 6.5 Under the conditions of Theorem 6.2, there holds

$$\|\vec{u} - \vec{u}_{\tau h}\|_{L^\infty(I;L^2(\Omega))} \leqslant C \ln \frac{\mathcal{T}}{\tau} \|\vec{u} - P_\tau R_h^S(\vec{u},p)\|_{L^\infty(I;L^2(\Omega))}.$$

Proof. Using Corollary 6.3 and arguing as in the proof of Corollary 6.4, we obtain

$$\|\vec{u}-\vec{u}_{\tau h}\|_{L^\infty(I;L^2(\Omega))} \leqslant C\ln\frac{\mathcal{T}}{\tau}\left(\inf_{\vec{v}_{\tau h}\in X^w_\tau(\vec{V}_h)}\|\vec{u}-\vec{v}_{\tau h}\|_{L^\infty(I;L^2(\Omega))}+\|\vec{u}-P_\tau R^S_h(\vec{u},p)\|_{L^\infty(I;L^2(\Omega))}\right).$$

Choosing $\vec{v}_{\tau h}=P_\tau R^S_h(\vec{u},p)\in X^w_\tau(\vec{V}_h)$ we obtain the result. □

7. Error estimates and comparison to the literature

In this section we apply our result from Corollary 6.5 to derive *a priori* error estimates. Due to the nature of the result from Corollary 6.5, we obtain error estimates that are optimal (probably up to a logarithmic term) with respect to both the orders of approximation and the assumed regularity. Moreover, we compare these estimates with the results from the literature. For this section we assume the domain Ω to be polygonal/polyhedral and convex.

Remark 7.1 Note that the results from Corollary 6.4 or 6.5 can be applied also to nonconvex domains and to meshes, which are not necessarily quasi-uniform, including graded or even anisotropic refinement toward reentrant corners or edges. In this case, one can use error estimates for the solution of the stationary Stokes equations for such cases; see Apel & Kempf (2021), and the references therein, in order to estimate $\vec{u}-R^S_h(\vec{u},p)$ in (6.5).

For this section we assume the following standard approximation properties for the spaces $\vec{X}_h$ and M_h.

Assumption 7.2 There exists an interpolation operator $i_h\colon H^2(\Omega)^d\cap H^1_0(\Omega)^d \to \vec{X}_h$ and $r_h\colon L^2(\Omega)\to M_h$ such that

$$\|\nabla(\vec{v}-i_h v)\|_{L^2(\Omega)} \leqslant ch\|\nabla^2 v\|_{L^2(\Omega)} \quad \forall\,\vec{v}\in H^2(\Omega)^d\cap H^1_0(\Omega)$$

and

$$\|q-r_h q\|_{L^2(\Omega)} \leqslant ch\|\nabla q\|_{L^2(\Omega)} \quad \forall\, q\in H^1(\Omega).$$

This assumption is fulfilled for a variety of finite element pairs including, e.g., Taylor-Hood as well as Mini finite elements on a family of shape-regular meshes. Under this assumption, the following standard estimate holds for the Ritz projection for the Stokes problem (6.1).

Proposition 7.3 Let Ω be convex and Assumption 7.2 be fulfilled. There is a constant $C>0$ such that for all $(\vec{u},p)$ with $\vec{u}\in H^2(\Omega)^d\cap\vec{V}^1$ and $p\in H^1(\Omega)\cap L^2_0(\Omega)$ the following estimate holds:

$$\|\vec{u}-R^S_h(\vec{u},p)\|_{L^2(\Omega)} \leqslant Ch^2\left(\|\nabla^2\vec{u}\|_{L^2(\Omega)}+\|\nabla p\|_{L^2(\Omega)}\right).$$

Proof. We refer, e.g., to Girault & Raviart (1986, Theorem 1.9). □

In the following theorem, we provide an error estimate of order $\mathcal{O}(\tau+h^2)$ up to a logarithmic term under minimal assumptions on the data. The estimate holds for every choice of degree w in the temporal discretization but especially for $w=0$, i.e., for the dG(0) discretization, which is known to be a variant of the implicit Euler scheme.

THEOREM 7.4 Let Ω be convex, $\vec{f} \in L^\infty(I, L^2(\Omega)^d)$ and $\vec{u}_0 \in \vec{V}^2$, and let Assumption 7.2 be fulfilled. Let $(\vec{u}, p)$ be the solution of (2.8) and $(\vec{u}_{\tau h}, p_{\tau h})$ solve the respective finite element problem (4.5). Then there holds

$$\|\vec{u} - \vec{u}_{\tau h}\|_{L^\infty(I;L^2(\Omega))} \leqslant C \left(\ln \frac{\mathcal{T}}{\tau}\right)^2 \left(\tau + h^2\right) \left(\|\vec{f}\|_{L^\infty(I;L^2(\Omega))} + \|\vec{u}_0\|_{\vec{V}^2}\right).$$

Proof. We start with the result from Corollary 6.5 and obtain

$$\begin{aligned}
\|\vec{u} - \vec{u}_{\tau h}\|_{L^\infty(I;L^2(\Omega))} &\leqslant C \ln \frac{\mathcal{T}}{\tau} \|\vec{u} - P_\tau R_h^S(\vec{u}, p)\|_{L^\infty(I;L^2(\Omega))} \\
&\leqslant C \ln \frac{\mathcal{T}}{\tau} \left(\|\vec{u} - P_\tau \vec{u}\|_{L^\infty(I;L^2(\Omega))} + \|P_\tau(\vec{u} - R_h^S(\vec{u}, p))\|_{L^\infty(I;L^2(\Omega))}\right) \\
&\leqslant C \ln \frac{\mathcal{T}}{\tau} \left(\|\vec{u} - P_\tau \vec{u}\|_{L^\infty(I;L^2(\Omega))} + \tau^{-\frac{1}{s}} \|P_\tau(\vec{u} - R_h^S(\vec{u}, p))\|_{L^s(I;L^2(\Omega))}\right) \\
&\leqslant C \ln \frac{\mathcal{T}}{\tau} \left(\|\vec{u} - P_\tau \vec{u}\|_{L^\infty(I;L^2(\Omega))} + \tau^{-\frac{1}{s}} \|\vec{u} - R_h^S(\vec{u}, p)\|_{L^s(I;L^2(\Omega))}\right),
\end{aligned}$$

where we have used an inverse inequality for some $1 < s < \infty$ and the stability of P_τ in L^s from (4.2). The temporal projection error is estimated by (4.3) resulting in

$$\|\vec{u} - P_\tau \vec{u}\|_{L^\infty(I;L^2(\Omega))} \leqslant C\tau^{1-\frac{1}{s}} \|\partial_t \vec{u}\|_{L^s(I;L^2(\Omega))}.$$

The spatial error is estimated by Proposition 7.3 resulting in

$$\|\vec{u} - R_h^S(\vec{u}, p)\|_{L^s(I;L^2(\Omega))} \leqslant Ch^2 \left(\|\nabla^2 \vec{u}\|_{L^s(I;L^2(\Omega))} + \|\nabla p\|_{L^s(I;L^2(\Omega))}\right).$$

Using maximal parabolic regularity and the convexity of Ω (see Remarks 2.7 and 2.9 and Corollary 2.11) we obtain

$$\|\partial_t \vec{u}\|_{L^s(I;L^2(\Omega))} + \|\nabla^2 \vec{u}\|_{L^s(I;L^2(\Omega))} + \|\nabla p\|_{L^s(I;L^2(\Omega))} \leqslant \frac{Cs^2}{s-1} \left(\|\vec{f}\|_{L^s(I;L^2(\Omega))} + \|\vec{u}_0\|_{\vec{V}^2}\right).$$

For $s \geq 2$, we have $\frac{s^2}{s-1} \leqslant 2s$. We choose $s = 2\ln \frac{\mathcal{T}}{\tau} \geq 2$ and get

$$\tau^{-\frac{1}{s}} = \mathcal{T}^{-\frac{1}{s}} \left(\frac{\mathcal{T}}{\tau}\right)^{\frac{1}{s}} \leqslant C(\mathcal{T}) e^{\frac{1}{2}}.$$

Combining these terms, we obtain the desired estimate. □

REMARK 7.5 Under the additional assumption $\vec{u}_t, \Delta\vec{u}, \nabla p \in L^\infty(I, L^2(\Omega)^d)$, it is possible to remove one of the logarithmic terms in the result of Theorem 7.4.

To compare our error estimate from Theorem 7.4 with the results from the literature we first remark that our result especially holds for $w = 0$, i.e., the dG(0) discretization in time, which is known to be a variant of the implicit Euler scheme. In Heywood & Rannacher (1986), the authors discuss the discretization of the transient Navier–Stokes equation by the implicit Euler scheme in time and finite elements in space. They prove an estimate of order $\mathcal{O}(\tau + h^2)$ (which corresponds to Theorem 7.4 up to a logarithmic term) for the velocity error in the $L^\infty(I, L^2(\Omega)^d)$ norm; see Heywood & Rannacher (1986, p. 765). However, they require stronger regularity assumptions, in particular $\partial_t \vec{f} \in L^\infty(I; L^2(\Omega)^d)$. Note that our setting and the setting from Heywood & Rannacher (1986) are not fully comparable.

Chrysafinos & Walkington (2010) operate in a similar setting to here, discussing the discontinuous Galerkin method for the temporal discretization of the Stokes problem. For the full discretization they derive the following estimate (see Chrysafinos & Walkington, 2010, Theorem 4.9):

$$\|\vec{u}-\vec{u}_{\tau h}\|_{L^\infty(I;L^2(\Omega))} \leqslant C\Big(h\Big(\|\vec{u}\|_{L^2(I;H^2(\Omega))}+h\|\vec{u}\|_{L^\infty(I;H^2(\Omega))}\Big)+\tau\left(\|\vec{u}\|_{H^1(I;H^1(\Omega))}+\|\vec{u}\|_{W^{1,\infty}(I;L^2(\Omega))}\right)$$
$$+ h\|\vec{u}_0\|_{H^1(\Omega)} + \|\vec{u}\|_{C(I;H^2(\Omega))} \min\left(h^{3/2}/\tau, \sqrt{h/\tau}\right)h^{3/2} + h\|p\|_{L^2(I;H^1(\Omega))}\Big)$$

with corresponding regularity assumptions on the solution $(\vec{u}, p)$. This estimate holds true even if the finite element meshes change from time step to time step. In the setting of a fixed spatial mesh as considered here, the result simplifies to

$$\|\vec{u} - \vec{u}_{\tau h}\|_{L^\infty(I;L^2(\Omega))} \leqslant C\Big(h\Big(\|\vec{u}\|_{L^2(I;H^2(\Omega))} + h\|\vec{u}\|_{L^\infty(I;H^2(\Omega))}\Big) + \tau\left(\|\vec{u}\|_{H^1(I;H^1(\Omega))} + \|\vec{u}\|_{W^{1,\infty}(I;L^2(\Omega))}\right)$$
$$+ h\|\vec{u}_0\|_{H^1(\Omega)} + h\|p\|_{L^2(I;H^1(\Omega))}\Big) \tag{7.1}$$

providing an $\mathcal{O}(\tau + h)$ order of convergence. The only first order in h is due to the fact that in Chrysafinos & Walkington (2010) the $L^\infty(I; L^2(\Omega))$ norm is estimated simultaneously with the $L^2(I; H^1(\Omega))$ norm of the error. Comparing our result in Theorem 7.4 with (7.1), we want to emphasize that we require much less regularity and provide a better convergence order with respect to h.

8. Discrete regularity estimate for the pressure

The above results have so far been solely focused on the velocity estimates. To provide estimates in tune with Corollary 2.11 also in the discrete setting, we extend the results from Section 5 to the gradient of the pressure for certain finite element discretizations. These pressure estimates do not immediately lead to best approximation results as in the velocity case above, but to our knowledge discrete regularity estimates for the pressure have not yet been reported. In this section we assume the domain Ω to be convex and will use the estimate discussed in Remark 5.4, i.e.,

$$\max_{1\leqslant m\leqslant M} \|\partial_t \vec{u}_{\tau h}\|_{L^\infty(I_m;L^2(\Omega))} + \|\Delta_h \vec{u}_{\tau h}\|_{L^\infty(I;L^2(\Omega))}$$
$$+ \max_{1\leqslant m\leqslant M} \|\tau_m^{-1}[\vec{u}_{\tau h}]_{m-1}\|_{L^2(\Omega)} \leqslant C \ln \frac{\mathcal{T}}{\tau} \|\mathbb{P}_h \vec{f}\|_{L^\infty(I;L^2(\Omega))}, \tag{8.1}$$

in the setting of Theorem 5.2.

In the sequel, we require the following proposition, which is given in Guzman *et al.* (2013, Theorems 3.6, 4.1).

PROPOSITION 8.1 Let $\vec{X}_h \times M_h$ fulfill the inf-sup condition in (3.1) and the discrete pressure space fulfill the assumption $M_h \subset L_0^2(\Omega) \cap H^1(\Omega)$. Then there holds

$$\sup_{\vec{v}_h \in \vec{X}_h, \vec{v}_h \neq \vec{0}} \frac{(\nabla l_h, \vec{v}_h)}{\|\vec{v}_h\|_{L^2(\Omega)}} \geqslant C \|\nabla l_h\|_{L^2(\Omega)} \qquad \forall\, l_h \in M_h.$$

Note, that ∇l_h is well defined for these finite element spaces. Finite element spaces that fulfill these assumptions are, among others, Taylor-Hood or Mini finite element spaces

The next theorem provides a discrete regularity estimate for the pressure.

THEOREM 8.2 Let $1 \leqslant s \leqslant \infty, \Omega$ be convex, $\vec{f} \in L^s(I, L^2(\Omega)^d)$ and $\vec{u}_0 = 0$. Moreover, let the assumptions of Proposition 8.1 and Remark 5.4 be fulfilled. Let $(\vec{u}_{\tau h}, p_{\tau h}) \in X_\tau^w(\vec{X}_h \times M_h)$ be the solution to (4.11). Then there holds

$$\|\nabla p_{\tau h}\|_{L^s(I;L^2(\Omega))} \leqslant C \ln \frac{\mathscr{T}}{\tau} \|\vec{f}\|_{L^s(I;L^2(\Omega))}.$$

Proof. We first consider the case $s = \infty$.

Let $\tilde{t} \in I_{\tilde{m}}$ for $1 \leqslant \tilde{m} \leqslant M$, and let $\theta(t)$ from the proof of Theorem 6.2 be the regularized Dirac function supported in the interior of the time interval $I_{\tilde{m}}$ (cf. (6.3)) such that for $\tilde{t} \in I_{\tilde{m}}$ it holds by Proposition 8.1 and integration by parts that

$$\begin{aligned}\|\nabla p_{\tau h}(\tilde{t})\|_{L^2(\Omega)} \leqslant C \sup_{\vec{v}_h \in \vec{X}_h, \vec{v}_h \neq \vec{0}} \frac{(\nabla p_{\tau h}(\tilde{t}), \vec{v}_h)_\Omega}{\|\vec{v}_h\|_{L^2(\Omega)}} &= C \sup_{\vec{v}_h \in \vec{X}_h, \vec{v}_h \neq \vec{0}} \frac{(\nabla p_{\tau h}, \theta \vec{v}_h)_{I\times\Omega}}{\|\vec{v}_h\|_{L^2(\Omega)}} \\ &= C \sup_{\vec{v}_h \in \vec{X}_h, \vec{v}_h \neq \vec{0}} \frac{(-p_{\tau h}, \theta \nabla \cdot \vec{v}_h)_{I\times\Omega}}{\|\vec{v}_h\|_{L^2(\Omega)}}.\end{aligned}$$

Since $p_{\tau h}(\vec{x})$ is in $X_\tau^w(L^2(\Omega))$, using the orthogonal projection P_τ from (4.1) we have

$$(p_{\tau h}, \theta \nabla \cdot \vec{v}_h)_{I\times\Omega} = (p_{\tau h}, P_\tau(\theta) \nabla \cdot \vec{v}_h)_{I\times\Omega} = (p_{\tau h}, \nabla \cdot (P_\tau(\theta)\vec{v}_h))_{I\times\Omega}.$$

Notice that since here $P_\tau(\theta(t)) \in X_\tau^w(L^2(\Omega))$ and it is constant in space and $\vec{v}_h \in \vec{X}_h$ is constant in time, we have $P_\tau(\theta)\vec{v}_h \in X_\tau^w(\vec{X}_h)$. Thus, testing the weak formulation of the fully discrete Stokes problem in (4.5) with $(P_\tau(\theta)\vec{v}_h, 0)$, we have

$$\begin{aligned}(p_{\tau h}, \nabla \cdot (P_\tau(\theta)\vec{v}_h))_{I\times\Omega} = \sum_{m=1}^{M} \langle \partial_t \vec{u}_{\tau h}, P_\tau(\theta)\vec{v}_h \rangle_{I_m\times\Omega} - (\Delta_h \vec{u}_{\tau h}, P_\tau(\theta)\vec{v}_h)_{I\times\Omega} \\ + \sum_{m=2}^{M} ([\vec{u}_{\tau h}]_{m-1}, (P_\tau(\theta))_{m-1}^{+} \vec{v}_h)_\Omega - (\vec{f}, P_\tau(\theta)\vec{v}_h)_{I\times\Omega},\end{aligned}$$

using that $\vec{u}_0 = \vec{0}$ and the definition of Δ_h for the second term. Using the Hölder inequality, we obtain the following estimate:

$$|(p_{\tau h}, \nabla \cdot P_\tau(\theta)\vec{v}_h)_{I\times\Omega}| \leqslant \left[\max_{1\leqslant m\leqslant M} \|\partial_t \vec{u}_{\tau h}\|_{L^\infty(I_m;L^2(\Omega))} + \|\Delta_h \vec{u}_{\tau h}\|_{L^\infty(I;L^2(\Omega))} + \|\vec{f}\|_{L^\infty(I;L^2(\Omega))}\right]$$
$$\times \|P_\tau(\theta)\vec{v}_h\|_{L^1(I;L^2(\Omega))}$$
$$+ \max_{2\leqslant m\leqslant M} \|\tau_m^{-1}[\vec{u}_{\tau h}]_{m-1}\|_{L^2(\Omega)} \sum_{m=1}^{M} \tau_m \|(P_\tau(\theta))_{m-1}^{+}\vec{v}_h\|_{L^2(\Omega)}.$$

Applying the stability of P_τ (4.2) and the bound of $\theta(t)$ in the L^1 norm, we have

$$\|P_\tau(\theta)\vec{v}_h\|_{L^1(I;L^2(\Omega))} \leqslant C\|\theta\|_{L^1(I)}\|\vec{v}_h\|_{L^2(\Omega)} \leqslant C\|\vec{v}_h\|_{L^2(\Omega)}.$$

Since θ is supported in the interior of $I_{\tilde{m}}$ and $P_\tau(\theta)_m = 0$ for all $m \neq \tilde{m}$, we obtain by

$$|P_\tau(\theta)_{\tilde{m}}^{+}| \leqslant C\|\theta\|_{L^\infty(I_{\tilde{m}})} \leqslant C\tau_{\tilde{m}}^{-1}$$

(cf. Schatz & Wahlbin, 1995, (eq. A.2)) and the second assumption on the time mesh that

$$\tau_{\tilde{m}+1}\|P_\tau(\theta)_{\tilde{m}}^{+}\vec{v}_h\|_{L^2(\Omega)} \leqslant C\|\vec{v}_h\|_{L^2(\Omega)}.$$

By (8.1) we obtain

$$\max_{1\leqslant m\leqslant M} \|\partial_t \vec{u}_{\tau h}\|_{L^\infty(I_m;L^2(\Omega))} + \|\Delta_h \vec{u}_{\tau h}\|_{L^\infty(I;L^2(\Omega))}$$
$$+ \max_{2\leqslant m\leqslant M} \|\tau_m^{-1}[\vec{u}_{\tau h}]_{m-1}\|_{L^2(\Omega)} \leqslant C \ln\frac{\mathcal{T}}{\tau}\|\vec{f}\|_{L^\infty(I;L^2(\Omega))}.$$

Collecting the estimates above, we establish that for any $\tilde{t} \in I$,

$$\|\nabla p_{\tau h}(\tilde{t})\|_{L^2(\Omega)} \leqslant C \sup_{\vec{v}_h\in\vec{X}_h, \vec{v}_h\neq\vec{0}} \frac{(-p_{\tau h}, \theta\nabla\cdot\vec{v}_h)_{I\times\Omega}}{\|\vec{v}_h\|_{L^2(\Omega)}} \leqslant C\ln\frac{\mathcal{T}}{\tau}\|\vec{f}\|_{L^\infty(I;L^2(\Omega))}.$$

Next we discuss the case $s = 1$. Here, direct application of Proposition 8.1 leads to a $\vec{v}_h$ that is dependent on time and thus cannot be separated from the time integral, which leads to technical difficulties. Thus, we will pursue a similar approach to above. We can expand the norm as follows:

$$\|\nabla p_{\tau h}\|_{L^1(I;L^2(\Omega))} = \sum_{m=1}^{M} \int_{I_m} \|\nabla p_{\tau h}\|_{L^2(\Omega)}\,\mathrm{d}t \leqslant \sum_{m=1}^{M} \tau_m \|\nabla p_{\tau h}\|_{L^\infty(I_m;L^2(\Omega))}.$$

Similarly to the case $s=\infty$, using regularized Dirac functions $\theta^m(t)$ from the proof of Theorem 6.2 for $\tilde{t}_m \in I_m$, we have

$$\begin{aligned}\sum_{m=1}^{M} \tau_m \|\nabla p_{\tau h}(\tilde{t}_m)\|_{L^2(\Omega)} &\leqslant C \sum_{m=1}^{M} \tau_m \sup_{\vec{v}_h^m \in \vec{X}_h, \vec{v}_h^m \neq \vec{0}} \frac{(\nabla p_{\tau h}(\tilde{t}_m), \vec{v}_h^m)_\Omega}{\|\vec{v}_h^m\|_{L^2(\Omega)}} \\ &= C \sum_{m=1}^{M} \tau_m \sup_{\vec{v}_h^m \in \vec{X}_h, \vec{v}_h^m \neq \vec{0}} \frac{(\nabla p_{\tau h}, \theta^m \vec{v}_h^m)_{I_m \times \Omega}}{\|\vec{v}_h^m\|_{L^2(\Omega)}},\end{aligned}$$

where in the last step we used that θ^m is supported in I_m. Since $\vec{v}_h^m$ and τ_m are constants on each I_m, we can pull the supremum out of the sum, to obtain

$$\begin{aligned}\sum_{m=1}^{M} \tau_m \|\nabla p_{\tau h}(\tilde{t}_m)\|_{L^2(\Omega)} &\leqslant C \sum_{m=1}^{M} \sup_{\vec{v}_h^m \in \vec{X}_h, \vec{v}_h^m \neq \vec{0}} \left(\nabla p_{\tau h}, \frac{\theta^m \tau_m \vec{v}_h^m}{\|\vec{v}_h^m\|_{L^2(\Omega)}} \right)_{I_m \times \Omega} \\ &= C \sum_{m=1}^{M} \sup_{\vec{v}_h^m \in \vec{X}_h, \vec{v}_h^m \neq \vec{0}} \left(\nabla p_{\tau h}, \frac{P_\tau(\theta^m) \tau_m \vec{v}_h^m}{\|\vec{v}_h^m\|_{L^2(\Omega)}} \right)_{I_m \times \Omega} \\ &= C \sup_{\substack{\vec{v}_h^1 \in \vec{X}_h, \dots, \vec{v}_h^M \in \vec{X}_h \\ \vec{v}_h^1 \neq \vec{0}, \dots, \vec{v}_h^M \neq \vec{0}}} \left(\nabla p_{\tau h}, \sum_{m=1}^{M} \frac{P_\tau(\theta^m) \tau_m \vec{v}_h^m}{\|\vec{v}_h^m\|_{L^2(\Omega)}} \right)_{I \times \Omega} \\ &= C \sup_{\substack{\vec{v}_h^1 \in \vec{X}_h, \dots, \vec{v}_h^M \in \vec{X}_h \\ \vec{v}_h^1 \neq \vec{0}, \dots, \vec{v}_h^M \neq \vec{0}}} (\nabla p_{\tau h}, \tilde{\vec{v}}_{\tau h})_{I \times \Omega},\end{aligned}$$

where we defined $\tilde{\vec{v}}_{\tau h} \in X_\tau^w(\vec{X}_h)$ by

$$\tilde{\vec{v}}_{\tau h} := \sum_{m=1}^{M} \frac{P_\tau(\theta^m) \tau_m \vec{v}_h^m}{\|\vec{v}_h^m\|_{L^2(\Omega)}}.$$

Using the weak formulation in (4.5) and the Hölder estimate as before, we see

$$\begin{aligned}|(\nabla p_{\tau h}, \tilde{\vec{v}}_{\tau h})_{I \times \Omega}| \leqslant C \Bigg[&\sum_{m=1}^{M} \|\partial_t \vec{u}_{\tau h}\|_{L^1(I_m; L^2(\Omega))} + \|\Delta_h \vec{u}_{\tau h}\|_{L^1(I; L^2(\Omega))} + \sum_{m=2}^{M} \tau_m \|\tau_m^{-1} [\vec{u}_{\tau h}]_{m-1}\|_{L^2(\Omega)} \\ &+ \|\vec{f}\|_{L^1(I; L^2(\Omega))} \Bigg] \times \left[\max_{1 \leqslant m \leqslant M} \|\tilde{\vec{v}}_{\tau h, m-1}^{+}\|_{L^2(\Omega)} + \|\tilde{\vec{v}}_{\tau h}\|_{L^\infty(I; L^2(\Omega))} \right].\end{aligned}$$

By Remark 5.4 and Theorem 5.2, for $s = 1$ we have, similar to (8.1),

$$\sum_{m=1}^{M} \|\partial_t \vec{u}_{\tau h}\|_{L^1(I_m;L^2(\Omega))} + \|\Delta_h \vec{u}_{\tau h}\|_{L^1(I;L^2(\Omega))} + \sum_{m=2}^{M} \tau_m \|\tau_m^{-1}[\vec{u}_{\tau h}]_{m-1}\|_{L^2(\Omega)} \leqslant C \ln \frac{\mathscr{T}}{\tau} \|\vec{f}\|_{L^1(I;L^2(\Omega))}.$$

Using the stability of P_τ in $L^\infty(I)$ (4.2), we obtain

$$\begin{aligned}\|\tilde{\vec{v}}_{\tau h}\|_{L^\infty(I;L^2(\Omega))} &= \max_{1\leqslant m\leqslant M} \|\tilde{\vec{v}}_{\tau h,m}\|_{L^\infty(I_m;L^2(\Omega))} = \max_{1\leqslant m\leqslant M} \frac{\|P_\tau(\theta^m)\tau_m \vec{v}_h^m\|_{L^\infty(I_m;L^2(\Omega))}}{\|\vec{v}_h^m\|_{L^2(\Omega)}} \\ &\leqslant \max_{1\leqslant m\leqslant M} \frac{C\|\theta^m\|_{L^\infty(I_m)}\tau_m \|\vec{v}_h^m\|_{L^2(\Omega)}}{\|\vec{v}_h^m\|_{L^2(\Omega)}} \leqslant \max_{1\leqslant m\leqslant M} C,\end{aligned}$$

where we used that $\|\theta^m\|_{L^\infty(I_m)} \leqslant C\tau_m{}^{-1}$. Similarly, $\max_{1\leqslant m\leqslant M}\|\tilde{\vec{v}}^{+}_{\tau h,m-1}\|_{L^2(\Omega)} \leqslant C$. Combining the steps above, we arrive at the following estimate:

$$\|\nabla p_{\tau h}\|_{L^1(I;L^2(\Omega))} \leqslant \sum_{m=1}^{M} \tau_m \|\nabla p_{\tau h}\|_{L^\infty(I_m;L^2(\Omega))} \leqslant C \ln \frac{\mathscr{T}}{\tau} \|\vec{f}\|_{L^1(I;L^2(\Omega))}.$$

This shows the estimate in $L^1(I;L^2(\Omega))$. By interpolation we obtain the result for $1 \leqslant s \leqslant \infty$. □

Corollary 8.3 Let $p_{\tau h}$ be the pressure solution to (4.5) with $\vec{f} = \vec{0}$. Moreover, let the assumptions of Proposition 8.1 and Remark 5.4 be fulfilled. Then, there holds for $m = 1, 2, \ldots, M$,

$$\|\nabla p_{\tau h}\|_{L^\infty(I_m;L^2(\Omega))} \leqslant \frac{C}{t_m} \|\vec{u}_0\|_{L^2(\Omega)}.$$

Proof. The result follows by the same arguments that we used to show the theorem above. A notable difference is that we only consider I_m here, not the whole domain, and use Theorem 5.1 instead of Theorem 5.2. □

Funding

D. Leykekhman was partially supported by NSF grant DMS-1913133. N. Behringer was supported by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project number 188264188/GRK1754.

References

Abe, K., Giga, Y. & Hieber, M. (2015) Stokes resolvent estimates in spaces of bounded functions. *Ann. Sci. Éc. Norm. Supér. (4)*, **4**, 537–559.

Ahmed, N., Becher, S. & Matthies, G. (2017) Higher-order discontinuous Galerkin time stepping and local projection stabilization techniques for the transient Stokes problem. *Comput. Methods Appl. Mech. Engrg.*, **313**, 28–52.

Apel, T. & Kempf, V. (2021) Pressure-robust error estimate of optimal order for the Stokes equations: domains with re-entrant edges and anisotropic mesh grading. *Calcolo*, **58**, 20.

Ashyralyev, A. & Sobolevskii, P. E. (1994) *Well-Posedness of Parabolic Difference Equations*. Operator Theory: Advances and Applications, vol. 69. Basel: Birkhäuser.

Chrysafinos, K. & Walkington, N. J. (2010) Discontinuous Galerkin approximations of the Stokes and Navier–Stokes equations. *Math. Comp.*, **79**, 2135–2167.

Dauge, M. (1989) Stationary Stokes and Navier–Stokes systems on two- or three-dimensional domains with corners. Part I. Linearized equations. *SIAM J. Math. Anal.*, **20**, 74–97.

De Simon, L. (1964) Un'applicazione della teoria degli integrali singolari allo studio delle equazioni differenziali lineari astratte del primo ordine. *Rend. Sem. Mat. Univ. Padova*, **34**, 205–223.

Eriksson, K. & Johnson, C. (1995) Adaptive finite element methods for parabolic problems. II. Optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$. *SIAM J. Numer. Anal.*, **32**, 706–740.

Eriksson, K., Johnson, C. & Larsson, S. (1998) Adaptive finite element methods for parabolic problems VI: analytic semigroups. *SIAM J. Numer. Anal.*, **35**, 1315–1325.

Galdi, G. P. (2011) *An Introduction to the Mathematical Theory of the Navier–Stokes Equations: Steady-State Problems*, 2nd edn. New York: Springer.

Girault, V., Nochetto, R. H. & Scott, L. R. (2015) Max-norm estimates for Stokes and Navier–Stokes approximations in convex polyhedra. *Numer. Math.*, **131**, 771–822.

Girault, V. & Raviart, P. A. (1986) *Finite Element Methods for Navier–Stokes Equations*. Springer Series in Computational Mathematics, vol. 5. Berlin: Springer.

Guermond, J.-L. & Pasciak, J. E. (2008) Stability of discrete Stokes operators in fractional Sobolev spaces. *J. Math. Fluid Mech.*, **10**, 588–610.

Guzman, J., Salgado, A. J. & Sayas, F.-J. (2013) A note on the Ladyženskaja–Babuška–Brezzi condition. *J. Sci. Comput.*, **56**, 219–229.

Heywood, J. G. & Rannacher, R. (1982) Finite element approximation of the nonstationary Navier–Stokes problem. I. Regularity of solutions and second-order error estimates for spatial discretization. *SIAM J. Numer. Anal.*, **19**, 275–311.

Heywood, J. G. & Rannacher, R. (1986) Finite element approximation of the nonstationary Navier–Stokes problem, part II: stability of solutions and error estimates uniform in time. *SIAM J. Numer. Anal.*, **23**, 750–777.

Heywood, J. G. & Rannacher, R. (1990) Finite-element approximation of the nonstationary Navier–Stokes problem. Part IV: error analysis for second-order time discretization. *SIAM J. Numer. Anal.*, **27**, 353–384.

Kellogg, R. B. & Osborn, J. E. (1976) A regularity result for the Stokes problem in a convex polygon. *J. Funct. Anal.*, **21**, 397–431.

Leykekhman, D. & Vexler, B. (2016) Pointwise best approximation results for Galerkin finite element solutions of parabolic problems. *SIAM J. Numer. Anal.*, **54**, 1365–1384.

Leykekhman, D. & Vexler, B. (2017a) Discrete maximal parabolic regularity for Galerkin finite element methods. *Numer. Math.*, **135**, 923–952.

Leykekhman, D. & Vexler, B. (2017b) Global and interior pointwise best approximation results for the gradient of Galerkin solutions for parabolic problems. *SIAM J. Numer. Anal.*, **55**, 2025–2049.

Leykekhman, D. & Vexler, B. (2018) Discrete maximal parabolic regularity for Galerkin finite element methods for nonautonomous parabolic problems. *SIAM J. Numer. Anal.*, **56**, 2178–2202.

Leykekhman, D. & Wahlbin, L. B. (2008) A posteriori error estimates by recovered gradients in parabolic finite element equations. *BIT*, **48**, 585–605.

Lunardi, A. (1995) *Analytic Semigroups and Optimal Regularity in Parabolic Problems*. Modern Birkhäuser Classics. Basel: Birkhäuser/Springer.

Maz'ya, V. (2018) Seventy five (thousand) unsolved problems in analysis and partial differential equations. *Integral Equations Operator Theory*, **90**, 44.

Meidner, D., Rannacher, R. & Vexler, B. (2011) A priori error estimates for finite element discretizations of parabolic optimization problems with pointwise state constraints in time. *SIAM J. Control Optim.*, **49**, 1961–1997.

Schatz, A. H., Thomée, V. C. & Wahlbin, L. B. (1980) Maximum norm stability and error estimates in parabolic finite element equations. *Comm. Pure Appl. Math.*, **33**, 265–304.

Schatz, A. H. & Wahlbin, L. B. (1995) Interior maximum-norm estimates for finite element methods II. *Math. Comp.*, **64**, 907–928.

Shen, Z. (2012) Resolvent estimates in L^p for the Stokes operator in Lipschitz domains. *Arch. Rational Mech. Anal.*, **205**, 395–424.

Sohr, H. (2014) *The Navier–Stokes Equations: An Elementary Functional Analytic Approach*. Birkhäuser Advanced Texts. Basel: Birkhäuser.

Temam, R. (1977) *Navier–Stokes Equations: Theory and Numerical Analysis*. Studies in Mathematics and its Applications, vol. 2. Amsterdam–New York–Oxford: North-Holland.

Thomée, V., Xu, J. & Zhang, N. Y. (1989) Superconvergence of the gradient in piecewise linear finite-element approximation to a parabolic problem. *SIAM J. Numer. Anal.*, **26**, 553–573.