

# 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<sup>†</sup>, BORIS VEXLER<sup>‡</sup>

*Chair of Optimal Control, Center for Mathematical Sciences,  
Technical University of Munich, 85748 Garching by Munich, Germany*

AND

DMITRIY LEYKEKHMAN<sup>§</sup>

*Department of Mathematics, University of Connecticut, Storrs, CT 06269*

□

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,

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

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

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

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

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 which follows directly from the assumed data above. Such results are important in the analysis of PDE constrained

<sup>†</sup>Email: niklas.behringer@tum.de

<sup>‡</sup>Corresponding author. Email: vexler@ma.tum.de

<sup>§</sup>Email: dmitriy.leykekhman@uconn.edu

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

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

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  $\ell_\tau$  is a logarithmic term, explicitly given in the statements of the results, see Corollary 6.2.

The 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 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 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). 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  $\Omega$  and some approximation properties of the discrete spaces we prove error estimates of the form

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

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 non-autonomous 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 Eriksson & Johnson (1995); Leykekhman & Vexler (2016); Schatz *et al.* (1980); Meidner *et al.* (2011) and for the maximum norm of the gradient to Leykekhman & Vexler (2017b); Leykekhman & Wahlbin (2008); Thomée *et al.* (1989). Further results are also available in 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 in-stationary Stokes problem (1.1) in the literature. A result for the fully discrete problem in form of  $L^\infty(I;L^2(\Omega)^d)$  estimates based on discontinuous Galerkin methods is provided in Chrysafinos & Walkington (2010), including an overview over 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); 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

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

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 2.3 and Theorem 2.4 for the details and also for the formulation in the case  $\vec{u}_0 \neq 0$ . This estimate holds on a general Lipschitz domain  $\Omega$ . Assuming in addition the convexity of  $\Omega$ , we have

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

see Remark 2.4 and Corollary 2.1. On the discrete level, we provide the corresponding estimates which hold even in the limit cases  $s = 1$  and  $s = \infty$  at the expense of an 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

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

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  $\Omega$ , similar to the continuous case, we also obtain

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

where  $\Delta_h$  is the discrete Laplace operator, see Remark 5.2 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 the 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 result 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 (Guermont & 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  $L_0^2(\Omega)$  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

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

where the notation in the last line denotes the completion of the space  $\mathcal{V}$  with respect to the  $L^2(\Omega)^d$  and  $H^1(\Omega)^d$  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 (Galdi, 2011, 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

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

If the domain  $\Omega$  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, Theorem 1.4) and (Galdi, 2011, Theorem III.1.1))

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

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

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

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

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

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

$$\lambda_0 = \inf_{v \in H_0^1(\Omega)} \frac{\|\nabla v\|_{L^2(\Omega)}^2}{\|v\|_{L^2(\Omega)}^2}. \quad (2.2)$$

Similar to the Laplace operator, for convex polyhedral domains  $\Omega$  we have the following  $H^2$  regularity bound due to Dauge (1989); Kellogg & Osborn (1976)

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

## 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  $\vec{f} \in L^2(\Omega)^d$

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

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

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

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

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

The solution  $(\vec{u}, p)$  to (2.3) is a complex valued function in  $H_0^1(\Omega)^d \times L_0^2(\Omega)$  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

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

*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  $A - \bar{\omega}\text{Id}$  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  $A$  (2.1) one can rewrite the resolvent estimate (2.4) as

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

**REMARK 2.1** 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 non-smooth 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  $1 \leq s \leq \infty$  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 2.3 below. Therefore, we will first introduce the (standard) weak formulation on the divergence free space, then we discuss regularity issues and introduce a velocity-pressure weak formulation based on a slightly stronger assumption on the data.

**PROPOSITION 2.2** 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

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

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

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

*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 the 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  $\vec{u}_0 = 0$ .  $\square$

**REMARK 2.2** 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 2.2.

**REMARK 2.3** Under the assumptions from Proposition 2.2 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  $p = \partial_t P$ . Especially one can not 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 a slightly more regularity in order to be able to introduce the pressure  $p$  as a function, cf. Remark 2.3. 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 2.2 the equation (2.5) can be understood as an abstract parabolic problem

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

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

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

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 (2.6), 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  $s \rightarrow \infty$ . 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 non-homogeneous initial conditions.

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

$$\|\vec{v}_0\|_{\vec{V}_{1-\frac{1}{s}}^0} = \left( \int_0^1 \|A \exp(-tA) \vec{v}_0\|_{\vec{V}_0}^s dt \right)^{1/s} + \|\vec{v}_0\|_{\vec{V}_0} \quad (2.7)$$

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  $A\vec{u} \in L^s(I;\vec{V}^0)$ .

**PROPOSITION 2.3** Let  $1 < s < \infty$ ,  $\vec{f} \in L^s(I;L^2(\Omega)^d)$  and  $\vec{u}_0 \in \vec{V}_{1-\frac{1}{s}}^0$ . Then the solution  $\vec{u}$  to the problem

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

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

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

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

**REMARK 2.4** If the domain is polyhedral/polygonal and convex then Proposition 2.3 provides  $\vec{u} \in L^s(I;H^2(\Omega)^d)$  and the estimate

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

**REMARK 2.5** There holds  $\vec{V}^1 \hookrightarrow \vec{V}_{\frac{1}{2}}^0$ . 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

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

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 2.3.

REMARK 2.6 If  $u_0 \in \vec{V}^2$  there holds for every  $1 < s < \infty$

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

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 (2.7):

$$\begin{aligned} \|\vec{u}_0\|_{\vec{V}^0_{1-\frac{1}{s}}} &= \left( \int_0^1 \|A \exp(-tA) \vec{u}_0\|_{\vec{V}^0}^s dt \right)^{1/s} + \|\vec{u}_0\|_{\vec{V}^0} \\ &\leq \left( \int_0^1 \|\exp(-tA)\|_{\vec{V}^0 \rightarrow \vec{V}^0}^s \|A\vec{u}_0\|_{\vec{V}^0}^s dt \right)^{1/s} + \|\vec{u}_0\|_{\vec{V}^0} \leq C \|A\vec{u}_0\|_{L^2(\Omega)}. \end{aligned}$$

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

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

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

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  $f \in L^1(I; L^2(\Omega)^d)$  from Proposition 2.2 is only slightly strengthened to  $\vec{f} \in L^s(I; L^2(\Omega)^d)$  for some  $s > 1$ .

**THEOREM 2.4** 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

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

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

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

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

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

*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 2.3 and  $\vec{u} \in L^2(I; \vec{V}^1) \cap C(\bar{I}, \vec{V}^0)$  from Proposition 2.2. 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  $\vec{u}(t) \in \vec{V}^1 \hookrightarrow H_0^1(\Omega)$  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) we have  $p(t) \in L_0^2(\Omega)$  and

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

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

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

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 2.3. 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

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

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

**COROLLARY 2.1** Let the assumptions of Theorem 2.4 be fulfilled. Let in addition the domain  $\Omega$  be convex. Then we have  $p \in L^s(I,H^1(\Omega))$  and the corresponding estimate holds

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

*Proof.* By convexity of  $\Omega$  we obtain  $\vec{u} \in L^s(I;H^2(\Omega)^d)$  and the corresponding estimate, see Remark 2.4. 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 2.4. This leads to the desired regularity and to the estimate.  $\square$

**REMARK 2.7** For regularity beyond these estimates, we want to highlight (Heywood & Rannacher, 1982, Corollary 2.1), where the authors show that bounds for, e.g.,  $\nabla^3 \vec{u}$ ,  $\partial_t \vec{u}$ , go hand in hand with the need of the data  $\vec{u}_0$ ,  $\vec{f}$  and initial pressure  $p_0$  (defined as  $\lim_{t \rightarrow 0} p(t)$ ) satisfying a nonlocal compatibility condition for  $t \rightarrow 0$  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., them satisfying a uniform discrete inf-sup condition,

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

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

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

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

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

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.,

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

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

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

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

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

where  $\lambda_0$  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  $\lambda_0$ .

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

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

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

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

*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 be then 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  $\dim \vec{V}_h^\perp = \dim M_h$ .  $\square$

### 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

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

which we can also write compactly using the discrete operator as

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

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

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

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

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

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  $((-\Delta_h - \nu)\vec{u}_h, \vec{u}_h) \geq 0$ . Since  $z$  is restricted to the sector  $\Sigma_{\theta, -\nu}$  we can rewrite (3.6) as

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

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

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

which after an application of the Cauchy-Schwarz inequality completes the proof.  $\square$

## 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  $I = (0, \mathcal{T}]$  into subintervals  $I_m = (t_{m-1}, t_m]$  of length  $\tau_m = t_m - t_{m-1}$ , where  $0 = t_0 < t_1 < \dots < t_{M-1} < t_M = \mathcal{T}$ . 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  $\tau$  such that

$$\tau_{\min} \geq C\tau^\beta.$$

2. There is a constant  $\kappa > 0$  independent of  $\tau$  such that for all  $m = 1, 2, \dots, M-1$

$$\kappa^{-1} \leq \frac{\tau_m}{\tau_{m+1}} \leq \kappa.$$

3. It holds  $\tau \leq \frac{\mathcal{I}}{4}$ .

For a given Banach space  $\mathcal{B}$  and the order  $w \in \mathbb{N}$  we define the semi-discrete space  $X_\tau^w(\mathcal{B})$  of piecewise polynomial functions in time as

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

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

$$\mathcal{P}_{w, I_m}(\mathcal{B}) = \left\{ \vec{v}_\tau: I_m \rightarrow \mathcal{B} \mid \vec{v}_\tau(t) = \sum_{j=0}^w \vec{v}^j \phi_j(t), \vec{v}^j \in \mathcal{B}, j = 0, \dots, w \right\}.$$

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

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

For later use we introduce  $P_\tau: L^2(I; L^2(\Omega)^d) \rightarrow X_\tau^w(L^2(\Omega)^d)$  as the  $L^2$  projection in time by

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

We will use the following two standard properties

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

and

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

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  $X_\tau^w(\vec{V}_h)$ :

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

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 2.4, ensuring the Galerkin orthogonality relation, see also Proposition 4.2 below.

The above formulation is not a conforming discretization of the divergence free formulation (2.5) due to the fact that  $X_\tau^w(\vec{V}_h)$  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

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

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

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

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 (4.4) and (4.5).

PROPOSITION 4.1 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). Let now  $\vec{u}_{\tau h} \in X_\tau^w(\vec{V}_h)$  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

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

and one obtains  $\vec{w}_{\tau h}(t) \in \vec{V}_h^\perp$  for every  $t \in I_m$ ,  $m = 1, 2, \dots, M$ . 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

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

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 4.2 Let the assumptions of Theorem 2.4 be fulfilled, i.e.,  $\vec{f} \in L^s(I; L^2(\Omega)^d)$  for some  $s > 1$  and  $\vec{u}_0 \in \vec{V}_{1-\frac{1}{s}}^0$ . Then there holds for the solution  $(\vec{u}, p)$  of (2.8)

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

and consequently

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

*Proof.* In the setting of Theorem 2.4 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  $B$

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

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  $m = 1, 2, \dots, M$

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

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 form of the Dunford-Taylor integral, (cf. (Eriksson *et al.*, 1998, p. 1321–1322))

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

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 rational function with the numerator of degree  $w$  and denominator of degree  $w + 1$ . The contour  $\Gamma$  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

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

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).  $\square$

For the inhomogeneous problem (and  $\vec{u}_0 = 0$ ) we obtain the discrete analog of Proposition 2.3. 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 an logarithmic term.

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

$$\begin{aligned} & \left( \sum_{m=1}^M \|\partial_t \vec{u}_{\tau h}\|_{L^s(I_m; L^2(\Omega))}^s \right)^{1/s} + \|A_h \vec{u}_{\tau h}\|_{L^s(I; L^2(\Omega))} \\ & + \left( \sum_{m=1}^M \tau_m \|\tau_m^{-1} [\vec{u}_{\tau h}]_{m-1}\|_{L^2(\Omega)}^s \right)^{1/s} \leq C \ln \frac{\mathcal{J}}{\tau} \|\mathbb{P}_h \vec{f}\|_{L^s(I; L^2(\Omega))}. \end{aligned} \quad (5.1)$$

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

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

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).  $\square$

**REMARK 5.1** Due to the stability of the discrete Leray projection  $\mathbb{P}_h$  in  $L^2$ , we can drop it in the above estimates.

**REMARK 5.2** If we assume the domain  $\Omega$  to be convex and the family of meshes  $\{\mathcal{T}_h\}$  to be shape regular and quasi-uniform then there holds

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

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

For  $s = 2$  it is possible to get rid of the logarithmic term in (5.1) similar to (Meidner & Vexler, 2008, Theorem 4.6).

**THEOREM 5.3** For  $\vec{f} \in L^2(I, L^2(\Omega)^d)$  and  $\vec{u}_0 = 0$ . Let  $\vec{u}_{\tau h} \in X_\tau^w(\vec{V}_h)$  be the solution to (4.4). Then there

holds

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

*Proof.* The proof follows the lines of the corresponding proof in Meidner & Vexler (2008) replacing  $-\Delta_h$  by  $A_h$ .  $\square$

## 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 estimates for the velocity. In order to state the results, we need to 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

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

$$(\nabla \cdot (\vec{w} - R_h^S(\vec{w}, \varphi)), q_h) = 0, \quad \forall q_h \in M_h. \quad (6.1b)$$

**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  $\vec{w}$ . 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 2.4 in order to use the Galerkin orthogonality relation from Proposition 4.2.

**THEOREM 6.1** Let  $\vec{f} \in L^s(I; L^2(\Omega)^d)$  for some  $s > 1$  and  $\vec{u}_0 \in \vec{V}_{1-\frac{1}{s}}^0$ . 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

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

*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:

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



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  $\tilde{t} = \mathcal{I}$ . 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

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

By the Galerkin orthogonality from Proposition 4.2 we have

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

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's inequality and using  $\vec{u} \in C(\bar{I}; L^2(\Omega)^d)$ , we obtain

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

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

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

Combining the estimates, we conclude

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

and an application of Theorem 5.2 leads to

$$\|\vec{u}_{\tau h}(\tilde{t})\|_{L^2(\Omega)}^2 \leq C \ln \frac{\mathcal{J}}{\tau} \left( \|\vec{u}\|_{L^\infty(I;L^2(\Omega))} + \|\vec{u} - R_h^S(\vec{u}, p)\|_{L^\infty(I;L^2(\Omega))} \right) \times \|\vec{u}_{\tau h}(\tilde{t})\|_{L^2(\Omega)} \|\boldsymbol{\theta}\|_{L^1(I_M)}.$$

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

The following corollary provides a version of Theorem 6.1 involving the  $L^2$  projection in time  $P_\tau$  defined in (4.1).

**COROLLARY 6.1** Under the conditions of Theorem 6.1 there holds

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

*Proof.* We obtain this by arguing as in the proof of Theorem 6.1, only the term in (6.4) is estimated differently

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

where we have used the definition (4.1) of  $\mathcal{P}_\tau$ .  $\square$

As a corollary from Theorem 6.1 we obtain a best approximation type result.

**COROLLARY 6.2** Under the conditions of Theorem 6.1 there holds

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

*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 6.1. 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 6.1 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 6.1 we obtain then

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

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.  $\square$

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

**COROLLARY 6.3** Under the conditions of Theorem 6.1 there holds

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

*Proof.* Using Corollary 6.1 and arguing as in the proof of Corollary 6.2 we obtain

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

Choosing  $\vec{v}_{\tau h} = P_\tau R_h^S(\vec{u}, p) \in X_\tau^w(\vec{V}_h)$  we obtain the result.  $\square$

## 7. Error estimates and comparison to the literature

In this section we apply our result from Corollary 6.3 to derive a priori error estimates. Due to the nature of the result from Corollary 6.3 we obtain error estimates which 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  $\Omega$  to be polygonal/polyhedral and convex.

**REMARK 7.1** Note that the results from Corollary 6.2 or Corollary 6.3 can be applied also to non-convex domains and to meshes, which are not necessarily quasi-uniform, including graded or even anisotropic refinement towards 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  $\vec{u} - R_h^S(\vec{u}, p)$  in (6.5).

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

**ASSUMPTION 7.1** There exists an interpolation operator  $i_h: H^2(\Omega)^d \cap H_0^1(\Omega)^d \rightarrow \vec{X}_h$  and  $r_h: L^2(\Omega) \rightarrow M_h$  such that

$$\|\nabla(\vec{v} - i_h \mathbf{v})\|_{L^2(\Omega)} \leq ch \|\nabla^2 \mathbf{v}\|_{L^2(\Omega)} \quad \forall \vec{v} \in H^2(\Omega)^d \cap H_0^1(\Omega)$$

and

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

This assumption is fulfilled for a variety of finite element pairs including, e.g., Taylor-Hood as well as mini element 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 7.2** Let  $\Omega$  be convex and Assumption 7.1 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  $p \in H^1(\Omega) \cap L_0^2(\Omega)$  the following estimate holds

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

*Proof.* We refer, e.g., to (Girault & Raviart, 1986, Theorem 1.9).  $\square$

In the following theorem we provide an error estimate of order  $\mathcal{O}(\tau + h^2)$  up to a logarithmic term under minimal assumption 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 7.3** Let  $\Omega$  be convex,  $\vec{f} \in L^\infty(I, L^2(\Omega)^d)$  and  $\vec{u}_0 \in \vec{V}^2$  and let Assumption 7.1 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

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

*Proof.* We start with the result from Corollary 6.3 and obtain

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

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

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

The spatial error is estimated by Proposition 7.2 resulting in

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

Using maximal parabolic regularity and the convexity of  $\Omega$ , see Remark 2.4, Remark 2.6 and Corollary 2.1, we obtain

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

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

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

Combining these terms we obtain the desired estimate.  $\square$

**REMARK 7.2** Under the additional assumption  $\bar{\mathbf{u}}_t, \Delta \bar{\mathbf{u}}, \nabla p \in L^\infty(I, L^2(\Omega)^d)$ , it is possible to remove one of the logarithmic terms in the result of Theorem 7.3.

To compare our error estimate from Theorem 7.3 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 7.3 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.

The authors of Chrysafinos & Walkington (2010) operate in a similar setting as 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)

$$\begin{aligned}
\|\bar{\mathbf{u}} - \bar{\mathbf{u}}_{\tau h}\|_{L^\infty(I;L^2(\Omega))} &\leq C \left( h \left( \|\bar{\mathbf{u}}\|_{L^2(I;H^2(\Omega))} + h \|\bar{\mathbf{u}}\|_{L^\infty(I;H^2(\Omega))} \right) + \tau \left( \|\bar{\mathbf{u}}\|_{H^1(I;H^1(\Omega))} + \|\bar{\mathbf{u}}\|_{W^{1,\infty}(I;L^2(\Omega))} \right) \right) \\
&\quad + h \|\bar{\mathbf{u}}_0\|_{H^1(\Omega)} + \|\bar{\mathbf{u}}\|_{C(I;H^2(\Omega))} \min \left( h^{3/2}/\tau, \sqrt{h/\tau} \right) h^{3/2} + h \|p\|_{L^2(I;H^1(\Omega))} \quad (7.1)
\end{aligned}$$

with corresponding regularity assumptions on the solution  $(\vec{u}, p)$ . Comparing our result in Theorem 7.3 with (7.1), we want to emphasize that we require much less regularity, provide a better convergence order with respect to  $h$  and do not have any “mixed terms” containing both  $h$  and  $\tau$  (apart from the logarithmic term).

### 8. Discrete regularity estimate for the pressure

The above results were so far solely focused on the velocity estimates. To provide estimates in tune with Corollary 2.1 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  $\Omega$  to be convex and will use the estimate discussed in Remark 5.2, i.e.,

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

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

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

Note, that  $\nabla l_h$  is well-defined for these finite element spaces. Finite element spaces that fulfill these assumptions are among others the space of Taylor-Hood finite elements or the space of the mini element.

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

**THEOREM 8.2** Let  $1 \leq s \leq \infty$ ,  $\Omega$  be convex,  $\vec{f} \in L^s(I, L^2(\Omega)^d)$  and  $\vec{u}_0 = 0$ . Let moreover the assumptions of Proposition 8.1 and Remark 5.2 be fulfilled. Let  $(\vec{u}_{\tau h}, p_{\tau h}) \in X_\tau^w(\vec{X}_h \times M_h)$  be the solution to (4.5). Then there holds

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

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

Let  $\tilde{t} \in I_{\tilde{m}}$  for  $1 \leq \tilde{m} \leq M$  and let  $\theta(t)$  from the proof of Theorem 6.1 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,

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

Since  $p_{\tau h}(\vec{x})$  is in  $X_\tau^w(L^2(\Omega))$ , using the orthogonal projection  $P_\tau$  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 being constant in space and  $\vec{v}_h \in \vec{X}_h$  being 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

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

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

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

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

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

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

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

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

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

By (8.1) we obtain

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

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

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

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 as above. We can expand the norm as follows

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

Similarly to the case  $s = \infty$ , using regularized Dirac functions  $\theta^m(t)$  from the proof of Theorem 6.1 for  $\tilde{t}_m \in I_m$ , we have

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

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

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

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

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

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

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

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

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

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

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

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

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

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

**COROLLARY 8.1** Let  $p_{\tau h}$  be the pressure solution to (4.5) with  $\vec{f} = \vec{0}$ . Let moreover the assumptions of Proposition 8.1 and Remark 5.2 be fulfilled. Then, there holds for  $m = 1, 2, \dots, M$

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

*Proof.* The result follows by the same arguments which 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.  $\square$

## Funding

Deutsche Forschungsgemeinschaft (German Research Foundation) - Project number 188264188/GRK1754 to N.B.; NSF grant DMS-1913133 to D.L.

## REFERENCES

- ABE, K., GIGA, Y. & HIEBER, M. (2015) Stokes resolvent estimates in spaces of bounded functions. *Ann. Sci. Éc. Norm. Supér. (4)*, **48**, 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**, Paper No. 15, 20.
- ASHYRALYEV, A. & SOBOLEVSKII, P. E. (1994) *Well-posedness of parabolic difference equations*. Operator Theory: Advances and Applications, vol. 69. Basel: Birkhäuser Verlag, pp. xiv+349.
- 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. 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. & LARSSON, S. (1998) Adaptive finite element methods for parabolic problems. VI. Analytic semigroups. *SIAM J. Numer. Anal.*, **35**, 1315–1325.
- 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.
- GALDI, G. P. (2011) *An introduction to the mathematical theory of the Navier-Stokes equations*. Springer Monographs in Mathematics, second edn. New York: Springer, pp. xiv+1018. Steady-state problems.
- 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-Verlag, pp. x+374.



- 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 Ladyzenskaja-Babuska-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. 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. 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. Functional Analysis*, **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 Basel AG, pp. xviii+424.
- MAZ'YA, V. (2018) Seventy five (thousand) unsolved problems in analysis and partial differential equations. *Integral Equations Operator Theory*, **90**, Art. 25, 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.
- MEIDNER, D. & VEXLER, B. (2008) 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.
- 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. Ration. Mech. Anal.*, **205**, 395–424.
- SOHR, H. (2014) *The Navier-Stokes Equations: An Elementary Functional Analytic Approach*. Birkhäuser Advanced Texts Basler Lehrbücher. Basel: Birkhäuser.
- TEMAM, R. (1977) *Navier-Stokes equations. Theory and numerical analysis*. Amsterdam-New York-Oxford: North-Holland Publishing Co., pp. x+500. Studies in Mathematics and its Applications, Vol. 2.
- 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.