SIAM J. NUMER. ANAL.
Vol. 53, No. 2, pp. 1005–1031

ANALYSIS OF THE PRESSURE STABILIZED PETROV–GALERKIN METHOD FOR THE EVOLUTIONARY STOKES EQUATIONS AVOIDING TIME STEP RESTRICTIONS*

VOLKER JOHN† AND JULIA NOVO‡

Abstract. Optimal error estimates for the pressure stabilized Petrov–Galerkin method for the evolutionary Stokes equations are proved, in the case of regular solutions, without restriction on the length of the time step. These results clarify the "instability of the discrete pressure for small time steps" reported in the literature. First, the limit situation of the continuous-in-time discretization is considered and second the numerical analysis for the backward Euler scheme is carried out. The main results are applicable to higher order finite elements. The analytical results are strongly based on an appropriate approach for computing the initial velocity, which is suggested in this paper. Numerical studies confirm the theoretical results, showing in particular that this instability does not occur for the proposed initial condition.

Key words. evolutionary Stokes equations, PSPG stabilization, equal order pairs of finite element spaces, backward Euler scheme, optimal error estimates, small time step instability

AMS subject classifications. 65M12, 65M60

DOI. 10.1137/130944941

1. Introduction.

It is well known that stable mixed finite element approximations to the Stokes and Navier–Stokes equations are required to satisfy a discrete inf-sup condition. This condition prevents the use of many attractive mixed finite elements, such as equal-order continuous elements. If these kinds of elements are chosen, then one has to introduce some stabilization term to circumvent the inf-sup condition. One of the most popular stabilization schemes is the pressure stabilized Petrov–Galerkin (PSPG) method, which was first introduced in [7].

In this paper, the PSPG stabilization of the evolutionary Stokes problem

$$
\begin{aligned}
\partial_t \tilde{\boldsymbol{u}} - \nu \Delta \tilde{\boldsymbol{u}} + \nabla \tilde{p} &= \tilde{\boldsymbol{f}} && \text{in} \quad (0,T] \times \Omega, \\
\nabla \cdot \tilde{\boldsymbol{u}} &= 0 && \text{in} \quad [0,T] \times \Omega, \\
\tilde{\boldsymbol{u}} &= \boldsymbol{0} && \text{on} \quad [0,T] \times \partial\Omega, \\
\tilde{\boldsymbol{u}}(0,\boldsymbol{x}) &= \tilde{\boldsymbol{u}}_0(\boldsymbol{x}) && \text{in} \quad \Omega,
\end{aligned}
\tag{1.1}
$$

will be considered. In (1.1), $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, is a bounded domain, $(0,T)$ with $T < \infty$ the time interval, $\tilde{\boldsymbol{u}}$ is the velocity field, $\boldsymbol{u}_0$ the initial velocity, $\tilde{p}$ the pressure, and ν is the kinematic viscosity.

Finite element error analysis of the PSPG method can be found already in the literature. In [12], the PSPG method applied to a steady-state problem was analyzed. The stability and convergence of the PSPG method applied to the evolutionary Stokes equations have been already considered in [1, 3]. In [1], it is stated that if the time

---

*Received by the editors November 12, 2013; accepted for publication (in revised form) February 5, 2015; published electronically April 9, 2015.
http://www.siam.org/journals/sinum/53-2/94494.html

†Weierstrass Institute for Applied Analysis and Stochastics, Leibniz Institute in Forschungsverbund Berlin e. V. (WIAS), 10117 Berlin, Germany, and Department of Mathematics and Computer Science, Free University of Berlin, 14195 Berlin, Germany (john@wias-berlin.de). This work was performed while visiting the Universidad Autónoma de Madrid.

‡Departamento de Matemáticas, Universidad Autónoma de Madrid, Madrid, Spain (julia.novo@uam.es). The research of this author was supported by Spanish MEC under grants MTM2010-14919 and MTM2013-42538-P.

step is sufficiently small, then the fully discrete problem necessarily leads to unstable pressure approximations. Similar conclusions were drawn in [3]. In this paper, stability and optimal convergence were proved for the velocity for piecewise affine approximations and the backward Euler method. For higher order polynomials, the results in [3] hold under the assumption $\Delta t \geq h^2/\nu$. Concerning the pressure, the condition $\Delta t \geq h^2/\nu$ is also required for piecewise affine approximations. The case of higher order polynomials or the use of the Crank–Nicolson method remain open.

Our interest in the numerical analysis of the PSPG method arose from the reported instabilities for small time steps. In this paper, first the limit case of small time steps will be studied, which is the continuous-in-time problem. Any instability for small time steps will be reflected in the analysis of the continuous-in-time discretization. Second, an analysis of the PSPG method in combination with the backward Euler scheme will be presented. In both cases, the derivation of the error bounds relies on an appropriately chosen initial condition for the discrete velocity.

The main result for the continuous-in-time situation consists in the derivation of optimal error bounds for both the velocity and pressure approximations, Theorem 4.5. This result holds also for higher order finite element discretizations. Its derivation assumes sufficient regularity of the solution and that for the finite element pressure space it holds $Q_h \subset H^1(\Omega)$. The latter property is satisfied for commonly used equal order discretizations. To obtain the estimates for the evolutionary problem, in section 3 error estimates for the PSPG method applied to a steady-state Stokes problem will be proved. The error bounds for the pressure do not deteriorate for small values of the viscosity parameter ν. In addition to the main result, two error estimates will be proved that apply for the P_1/P_1 and affine Q_1/Q_1 pairs of finite element spaces. The first estimate, Theorem 4.2, is similar to the main result but it needs fewer regularity assumptions. In the second estimate, Theorem 4.7, the error of the pressure in $L^2(0,T;L^2)$ is considered. Finally, Theorems 4.10 and 4.11 are analogous to Theorems 4.5 and 4.7 but in the particular case in which an appropriate initial condition for the discrete velocity is chosen.

In the case of piecewise affine approximations, the error bound for the velocity depends on the velocity approximation error at the initial time $t = 0$, e.g., as in [3]. However, the error bound in the case of using higher order polynomials depends both on the velocity and pressure approximation error at $t = 0$. From our point of view, this issue is a key point for the better understanding of the higher order polynomial case. Whereas the initial velocity is part of the definition of the problem, the initial pressure is not. However, following [5], the initial pressure can be defined as the solution of an overdetermined Poisson problem with Neumann boundary conditions. It will be suggested in Remark 4.8 to consider as an initial approximation for the velocity the finite element function obtained by solving the corresponding steady-state Stokes problem with right-hand side $\tilde{\boldsymbol{f}}(0) - \partial_t \tilde{\boldsymbol{u}}(0)$ using the PSPG method. In practice, the right-hand side is replaced by a finite element approximation of $-\nu\Delta\tilde{\boldsymbol{u}}_0 + \nabla\tilde{p}(0)$, where $\tilde{\boldsymbol{u}}_0$ and $\tilde{p}(0)$ are the initial velocity and pressure. With this approach, it is proved in Theorems 4.10 and 4.11 that error bounds can be obtained that depend only on the velocity approximation error at $t = 0$ but not on the pressure approximation error at time $t = 0$. These results are analogous to those obtained for the Galerkin discretization using inf-sup stable elements, e.g., in [5, 6]. The difference is that in the inf-sup stable case the results hold for any standard initial approximation to the velocity at $t = 0$ such as the interpolant of $\tilde{\boldsymbol{u}}_0$ or the L^2 projection of $\tilde{\boldsymbol{u}}_0$.

After having worked out the key role of choosing the initial condition as described above, the finite element error analysis for the backward Euler/PSPG method is presented. It turns out that optimal estimates can be proved, for regular solutions, without restrictions on the length of the time step in the case of using the proposed initial condition, Theorem 5.1. Special care has to be spent in analyzing the first time step.

In the numerical simulations it will be shown that a pressure instability for small time steps does not occur for the proposed choice of the discrete initial velocity.

The outline of the paper is as follows. In section 2, preliminaries and notation are stated. Section 3 studies finite element error estimates of the PSPG method for the steady-state case. The numerical analysis for the continuous-in-time case is presented in section 4 and for the backward Euler/PSPG method in section 5. Section 6 presents numerical studies which discuss the instabilities reported in the literature and which support the analytical results. The paper finishes with a summary in section 7.

2. The PSPG method for the evolutionary Stokes problem. For performing the numerical analysis of (1.1), it is of advantage to apply the change of variables $(\boldsymbol{u}, p) = e^{-\alpha t}(\tilde{\boldsymbol{u}}, \tilde{p})$ with $\alpha \in \mathbb{R}_+$. A direct calculation shows that one obtains with this transform a problem with a positive zeroth order term:

$$
(2.1) \qquad \begin{aligned}
\partial_t \boldsymbol{u} - \nu\Delta\boldsymbol{u} + \alpha\boldsymbol{u} + \nabla p &= \boldsymbol{f} && \text{in} \quad (0,T]\times\Omega,\\
\nabla\cdot\boldsymbol{u} &= 0 && \text{in} \quad [0,T]\times\Omega,\\
\boldsymbol{u} &= \boldsymbol{0} && \text{on} \quad [0,T]\times\partial\Omega,\\
\boldsymbol{u}(0,\boldsymbol{x}) &= \boldsymbol{u}_0(\boldsymbol{x}) && \text{in} \quad \Omega,
\end{aligned}
$$

where $\boldsymbol{f} = e^{-\alpha t}\tilde{\boldsymbol{f}}$. For problems with bounded time intervals, as they are studied here, one can choose, e.g., $\alpha = 1/T$. By construction, the error bounds for $(\tilde{\boldsymbol{u}}, \tilde{p})$ are the error bounds for $(\boldsymbol{u}, p)$ multiplied by the factor $e^{\alpha t}$. Choosing $\alpha = 1/T$, this factor is bounded by e.

Throughout the paper, standard notation is used for Sobolev spaces and corresponding norms; see, e.g., [4]. In particular, given a measurable set $\omega \subset \mathbb{R}^d$, the inner product in $L^2(\omega)$ or $L^2(\omega)^d$ is denoted by $(\cdot,\cdot)_\omega$ and the notation $(\cdot,\cdot)$ is used instead of $(\cdot,\cdot)_\Omega$. The norm (seminorm) in $W^{m,p}(\omega)$ will be denoted by $\|\cdot\|_{m,p,\omega}$ $(|\cdot|_{m,p,\omega})$, with the conventions $\|\cdot\|_{m,\omega} = \|\cdot\|_{m,2,\omega}$ and $\|\cdot\|_m = \|\cdot\|_{m,2,\Omega}$.

A variational formulation of (2.1) reads as follows: Find $\boldsymbol{u} : [0,T] \to V = H_0^1(\Omega)$ and $p : (0,T] \to Q = L_0^2(\Omega)$ such that for all $(\boldsymbol{v}, q) \in V \times Q$

$$
(2.2) \qquad (\partial_t \boldsymbol{u}, \boldsymbol{v}) + \nu(\nabla\boldsymbol{u}, \nabla\boldsymbol{v}) + \alpha(\boldsymbol{u}, \boldsymbol{v}) - (\nabla\cdot\boldsymbol{v}, p) + (\nabla\cdot\boldsymbol{u}, q) = (\boldsymbol{f}, \boldsymbol{v}).
$$

In this paper, the analysis will be carried out for a family $\{\mathcal{T}_h\}_{h>0}$ of uniform triangulations. The consideration of uniform triangulations, instead of quasi-uniform triangulations, serves for concentrating on the main topic of this paper without overburdening the error analysis with technical details. It will be assumed that the family is regular in the sense of [4]. Let $h = h_K$ denote the diameter of a mesh cell $K \in \mathcal{T}_h$.

On $\mathcal{T}_h$, the fixed-in-time finite element spaces $V_h \subset V$ and $Q_h \subset L_0^2(\Omega) \cap H^1(\Omega)$ are defined. The regularity assumption $Q_h \subset H^1(\Omega)$ will be needed in the analysis.

Since the triangulations are assumed to be regular, the following inverse inequality holds for each $\boldsymbol{v}_h \in V_h$ and each mesh cell $K \in \mathcal{T}_h$ (e.g., see [4, Theorem 3.2.6]),

$$
(2.3) \quad \|\boldsymbol{v}_h\|_{m,q,K} \le c_{\mathrm{inv}} h_K^{\,l-m-d\left(\frac{1}{q'}-\frac{1}{q}\right)} \|\boldsymbol{v}_h\|_{l,q',K}, \quad 0 \le l \le m \le 1, \ 1 \le q' \le q \le \infty.
$$

Let $q \in [1,\infty]$ and let $s \in \{0,1\}$ with $s \le t \le r+1$. Then, I_h denotes a bounded linear interpolation operator $I_h : W^{t,q}(\Omega) \to V_h$ that satisfies for all $\boldsymbol{v} \in W^{t,q}(\Omega)$

and all mesh cells $K \in \mathcal{T}_h$

$$|\boldsymbol{v} - I_h\boldsymbol{v}|_{s,q,K} \le Ch_K^{t-s}|\boldsymbol{v}|_{t,q,K}; \tag{2.4}$$

e.g., see [4]. This interpolation bound also holds for the pressure space Q_h. In this case, the interpolation operator, denoted by J_h with $J_h : W^{t,q}(\Omega) \to Q_h$, is acting on scalar-valued functions instead of vector-valued functions.

The PSPG approximation of (2.1) that will be considered has the form find $\boldsymbol{u}_h : [0,T] \to V_h$ and $p_h : (0,T] \to Q_h$ satisfying for all $(\boldsymbol{v}_h, q_h) \in V_h \times Q_h$

$$\begin{aligned}(\partial_t\boldsymbol{u}_h, \boldsymbol{v}_h) + \nu(\nabla\boldsymbol{u}_h, \nabla\boldsymbol{v}_h) + \alpha(\boldsymbol{u}_h, \boldsymbol{v}_h)&\\ - (\nabla\cdot\boldsymbol{v}_h, p_h) + (\nabla\cdot\boldsymbol{u}_h, q_h) + \mu(\nabla\cdot\boldsymbol{u}_h, \nabla\cdot\boldsymbol{v}_h)&\\ = (\boldsymbol{f}, \boldsymbol{v}_h) + \sum_{K\in\mathcal{T}_h} \delta(\boldsymbol{f} - \partial_t\boldsymbol{u}_h + \nu\Delta\boldsymbol{u}_h - \alpha\boldsymbol{u}_h - \nabla p_h, \nabla q_h)_K&\end{aligned} \tag{2.5}$$

with an approximation $\boldsymbol{u}_h(0)$ of the initial velocity $\boldsymbol{u}_0(\boldsymbol{x})$. The actual choice of $\boldsymbol{u}_h(0)$ will be discussed in Remark 4.8. The so-called grad-div term, the last term on the left-hand side, is often used in combination with the PSPG method [2]. It is well known that the majority of commonly used finite element methods does not give divergence-free solutions and that the violation of the divergence constraint might be even large [11]. The grad-div term is a penalty term with respect to the continuity equation and it serves for improving the conservation of mass. It is included in the error analysis just for completeness. The analysis can be carried out, with only slight modifications, also without this term.

A common approach for performing a finite element error analysis of a transient problem consists in incorporating steady-state problems and utilizing estimates for the latter problems. This approach will be also applied here. To this end set $\boldsymbol{g} = \boldsymbol{f} - \partial_t\boldsymbol{u}$ and consider the following problem: Find $(\boldsymbol{s}_h, z_h)$ such that for all $(\boldsymbol{v}_h, q_h) \in V_h \times Q_h$

$$\begin{aligned}\nu(\nabla\boldsymbol{s}_h, \nabla\boldsymbol{v}_h) + \alpha(\boldsymbol{s}_h, \boldsymbol{v}_h) - (\nabla\cdot\boldsymbol{v}_h, z_h) + (\nabla\cdot\boldsymbol{s}_h, q_h) + \mu(\nabla\cdot\boldsymbol{s}_h, \nabla\cdot\boldsymbol{v}_h)&\\ = (\boldsymbol{g}, \boldsymbol{v}_h) + \sum_{K\in\mathcal{T}_h} \delta(\boldsymbol{g} + \nu\Delta\boldsymbol{s}_h - \alpha\boldsymbol{s}_h - \nabla z_h, \nabla q_h)_K.&\end{aligned} \tag{2.6}$$

The solution of the corresponding continuous problem is $(\boldsymbol{u}, p)$.

The result of the following lemma will be applied to bound the pressure error.

LEMMA 2.1. *Let $\mathcal{T}_h$ be a uniform triangulation and let $q_h \in Q_h \subset H^1(\Omega)$. Then it holds*

$$\|q_h\|_0 \le Ch\|\nabla q_h\|_0 + C\sup_{\boldsymbol{v}_h\in V_h}\frac{(q_h, \nabla\cdot\boldsymbol{v}_h)}{\|\boldsymbol{v}_h\|_1}. \tag{2.7}$$

Proof. The proof follows [3, Lemma 3]. Since $Q_h \subset L_0^2(\Omega)$, it is well known from the theory of saddle point problems that there is a unique $\boldsymbol{v} \in V$ such that $-\nabla\cdot\boldsymbol{v} = q_h$ and $\|\nabla\boldsymbol{v}\|_0 \le C\|q_h\|_0$. Let π_h denote the $L^2(\Omega)$ projection onto V_h, having the following properties:

$$\|\pi_h\boldsymbol{v} - \boldsymbol{v}\|_0 \le Ch\|\boldsymbol{v}\|_1, \quad \|\pi_h\boldsymbol{v}\|_1 \le C\|\boldsymbol{v}\|_1. \tag{2.8}$$

Using these properties, integration by parts, and Poincaré's inequality yields

$$\begin{aligned}\|q_h\|_0^2 = (q_h, -\nabla\cdot\boldsymbol{v}) &= (\nabla q_h, (\boldsymbol{v} - \pi_h\boldsymbol{v})) - (q_h, \nabla\cdot\pi_h\boldsymbol{v})\\ &\le Ch\|\nabla q_h\|_0\|\boldsymbol{v}\|_1 - (q_h, \nabla\cdot\pi_h\boldsymbol{v})\\ &\le Ch\|\nabla q_h\|_0\|q_h\|_0 + |(q_h, \nabla\cdot\pi_h\boldsymbol{v})|\,.\end{aligned}$$

Applying the estimate of the pressure from below by the velocity, Poincaré's inequality, and the stability of the projection gives

$$\frac{|(q_h, \nabla \cdot \pi_h \boldsymbol{v})|}{\|q_h\|_0} \le C \frac{|(q_h, \nabla \cdot \pi_h \boldsymbol{v})|}{\|\boldsymbol{v}\|_1} \le C \frac{|(q_h, \nabla \cdot \pi_h \boldsymbol{v})|}{\|\pi_h \boldsymbol{v}\|_1}$$

such that

$$\|q_h\|_0 \le Ch\|\nabla q_h\|_0 + C \frac{|(q_h, \nabla \cdot \pi_h \boldsymbol{v})|}{\|\pi_h \boldsymbol{v}\|_1}.$$

Since $\pi_h \boldsymbol{v} \in V_h$, (2.7) follows. □

3. Finite element error analysis for the steady-state problem. This section presents a finite element error analysis of the PSPG method for the steady-state Stokes problem with reactive term. Error bounds will be derived which will be applied in the analysis of the transient problem. The numerical analysis keeps track of the dependency of the error bounds on the coefficients of the problem ν, α and the parameters of the discretization δ, μ. It turns out that the error bounds for the pressure do not deteriorate for small values of the viscosity parameter ν.

The strong form of the steady-state problem looks as follows:

$$\begin{aligned} -\nu\Delta \boldsymbol{s} + \alpha \boldsymbol{s} + \nabla z &= \boldsymbol{g} \quad \text{in} \quad \Omega, \\ \nabla \cdot \boldsymbol{s} &= 0 \quad \text{in} \quad \Omega, \\ \boldsymbol{s} &= \boldsymbol{0} \quad \text{on} \quad \partial\Omega. \end{aligned} \tag{3.1}$$

Here $\boldsymbol{g}$ is a given function. This problem will be discretized with the PSPG method (2.6) giving the finite element solution $(\boldsymbol{s}_h, z_h)$.

LEMMA 3.1. *Let* $(\boldsymbol{s}, z) \in \left(H^{k+1}(\Omega)\right)^d \times H^{l+1}(\Omega)$ *with* $k \ge 1$ *and* $l \ge 0$ *and let* $(\boldsymbol{s}_h, z_h) \in V_h \times Q_h$ *be the solution of* (2.6). *If* $Q_h \subset H^1(\Omega)$ *and if*

$$\delta_0 h^2 \le \delta \le \min\left\{\frac{h^2}{8\nu C_{\text{inv}}^2}, \frac{1}{4\alpha}\right\} \tag{3.2}$$

with $\delta_0 > 0$, *where* C_{inv} *is the constant of the inverse inequality* (2.3), *then the following error bound holds:*

$$\begin{aligned} |||(\boldsymbol{s} - \boldsymbol{s}_h, z - z_h)|||_h &\le Ch^k \left(\nu^{1/2} + h\alpha^{1/2} + \mu^{1/2} + \delta_0^{-1/2}\right) \|\boldsymbol{s}\|_{k+1} \\ &\quad + Ch^l \left(\delta^{1/2} + h\mu^{-1/2}\right) \|z\|_{l+1}, \end{aligned} \tag{3.3}$$

where $|||(\boldsymbol{v}, q)|||_h = (\nu\|\nabla \boldsymbol{v}\|_0^2 + \alpha\|\boldsymbol{v}\|_0^2 + \mu\|\nabla \cdot \boldsymbol{v}\|_0^2 + \delta\|\nabla q\|_0^2)^{1/2}$.

Proof. Denote by $I_h \boldsymbol{s}$ the Lagrange interpolant of $\boldsymbol{s}_h$ in V_h and by $J_h z$ the Lagrange interpolant onto Q_h. The errors are split in the usual way:

$$\begin{aligned} \boldsymbol{s} - \boldsymbol{s}_h &= (\boldsymbol{s} - I_h \boldsymbol{s}) - (\boldsymbol{s}_h - I_h \boldsymbol{s}) = (\boldsymbol{s} - I_h \boldsymbol{s}) - \boldsymbol{E}_h, \\ z - z_h &= (z - J_h z) - (z_h - J_h z) = (z - J_h z) - R_h. \end{aligned} \tag{3.4}$$

Since the PSPG method is consistent, the solution $(\boldsymbol{s}, z)$ of (3.1) satisfies also (2.6), i.e., a Galerkin orthogonality holds. Using this Galerkin orthogonality and adding and subtracting $I_h \boldsymbol{s}$ and $J_h z$ yields, with a straightforward calculation, an error equation. Taking as test functions $(\boldsymbol{v}_h, q_h) = (\boldsymbol{E}_h, R_h)$ in this error equation leads to

$$
\begin{aligned}
&\nu\|\nabla \boldsymbol{E}_h\|_0^2 + \alpha\|\boldsymbol{E}_h\|_0^2 + \mu\|\nabla\cdot\boldsymbol{E}_h\|_0^2 + \delta\|\nabla R_h\|_0^2 \\
&\quad = \sum_{K\in\mathcal{T}_h} \delta(\nu\Delta\boldsymbol{E}_h - \alpha\boldsymbol{E}_h, \nabla R_h)_K + \nu(\nabla(\boldsymbol{s}-I_h\boldsymbol{s}), \nabla\boldsymbol{E}_h) + \alpha(\boldsymbol{s}-I_h\boldsymbol{s}, \boldsymbol{E}_h) \\
&\qquad + (\nabla\cdot\boldsymbol{E}_h, J_h z - z) + (\nabla\cdot(\boldsymbol{s}-I_h\boldsymbol{s}), R_h) + \mu(\nabla\cdot(\boldsymbol{s}-I_h\boldsymbol{s}), \nabla\cdot\boldsymbol{E}_h) \\
&\qquad + \sum_{K\in\mathcal{T}_h} \delta(\nu\Delta(I_h\boldsymbol{s}-\boldsymbol{s}) - \alpha(I_h\boldsymbol{s}-\boldsymbol{s}) - \nabla(J_h z - z), \nabla R_h)_K.
\end{aligned}
\tag{3.5}
$$

Now, the terms on the right-hand side of (3.5) have to be bounded. The bounds will be obtained for all terms individually, using always the Cauchy–Schwarz inequality and Young's inequality. For the first term, applying also the inverse inequality (2.3) and the condition on the stabilization parameter (3.2), one obtains

$$
\begin{aligned}
\sum_{K\in\mathcal{T}_h} \delta(\nu\Delta\boldsymbol{E}_h, \nabla R_h)_K &\le \delta \sum_{K\in\mathcal{T}_h} h^{-1}\nu C_{\mathrm{inv}}\|\nabla\boldsymbol{E}_h\|_{0,K}\|\nabla R_h\|_{0,K} \\
&\le \delta \sum_{K\in\mathcal{T}_h}\left(h^{-2}\nu^2 C_{\mathrm{inv}}^2\|\nabla\boldsymbol{E}_h\|_{0,K}^2 + \frac{1}{4}\|\nabla R_h\|_{0,K}^2\right) \\
&\le \frac{\nu}{8}\|\nabla\boldsymbol{E}_h\|_0^2 + \frac{\delta}{4}\|\nabla R_h\|_0^2.
\end{aligned}
$$

Using again (3.2), one gets for the second term

$$
\sum_{K\in\mathcal{T}_h} \delta(\alpha\boldsymbol{E}_h, R_h) \le \delta\alpha^2\|\boldsymbol{E}_h\|_0^2 + \frac{\delta}{4}\|\nabla R_h\|_0^2 \le \frac{\alpha}{4}\|\boldsymbol{E}_h\|_0^2 + \frac{\delta}{4}\|\nabla R_h\|_0^2.
$$

The right-hand sides of these two estimates can be absorbed into the left-hand side of (3.5). The remaining estimates will use the interpolation error estimates (2.4) and they will contribute to the error bound. Straightforward calculations lead to

$$
\begin{aligned}
\nu(\nabla(\boldsymbol{s}-I_h\boldsymbol{s}), \nabla\boldsymbol{E}_h) &\le C\nu h^{2k}\|\boldsymbol{s}\|_{k+1}^2 + \frac{\nu}{4}\|\nabla\boldsymbol{E}_h\|_0^2, \\
\alpha(\boldsymbol{s}-I_h\boldsymbol{s}, \boldsymbol{E}_h) &\le C\alpha h^{2k+2}\|\boldsymbol{s}\|_{k+1}^2 + \frac{\alpha}{4}\|\boldsymbol{E}_h\|_0^2, \\
(\nabla\cdot\boldsymbol{E}_h, J_h z - z) &\le \frac{\mu}{4}\|\nabla\cdot\boldsymbol{E}_h\|_0^2 + C\mu^{-1}h^{2l+2}\|z\|_{l+1}^2.
\end{aligned}
$$

The next estimate uses integration by parts, where $Q_h \subset H^1(\Omega)$ is exploited,

$$
(\nabla\cdot(\boldsymbol{s}-I_h\boldsymbol{s}), R_h) = -(\boldsymbol{s}-I_h\boldsymbol{s}, \nabla R_h) \le C\delta_0^{-1}h^{2k}\|\boldsymbol{s}\|_{k+1}^2 + \frac{\delta}{16}\|\nabla R_h\|_0^2.
$$

One obtains in a straightforward way

$$
\mu(\nabla\cdot(\boldsymbol{s}-I_h\boldsymbol{s}), \nabla\boldsymbol{E}_h) \le C\mu h^{2k}\|\boldsymbol{s}\|_{k+1}^2 + \frac{\mu}{4}\|\nabla\cdot\boldsymbol{E}_h\|_0^2.
$$

Using the definition (3.2) of the stabilization parameter yields

$$
\begin{aligned}
\sum_{K\in\mathcal{T}_h} \delta(\nu\Delta(I_h\boldsymbol{s}-\boldsymbol{s}), \nabla R_h)_K &\le \sum_{K\in\mathcal{T}_h}\left(4\delta\nu^2\|\Delta(I_h\boldsymbol{s}-\boldsymbol{s})\|_{0,K}^2 + \frac{\delta}{16}\|\nabla R_h\|_{0,K}^2\right) \\
&\le C\nu h^{2k}\|\boldsymbol{s}\|_{k+1}^2 + \frac{\delta}{16}\|\nabla R_h\|_0^2 \\
\sum_{K\in\mathcal{T}_h} \delta\alpha(I_h\boldsymbol{s}-\boldsymbol{s}), \nabla R_h)_K &\le C\alpha h^{2k+2}\|\boldsymbol{s}\|_{k+1}^2 + \frac{\delta}{16}\|\nabla R_h\|_0^2.
\end{aligned}
$$

Finally, one gets

$$\sum_{K\in\mathcal{T}_h} \delta(\nabla(J_h z - z), \nabla R_h)_K \le C\delta h^{2l}\|z\|_{l+1} + \frac{\delta}{16}\|\nabla R_h\|_0^2.$$

Collecting all estimates, one obtains

$$\begin{aligned}|||(\boldsymbol{E}_h, R_h)|||_h \le\ & Ch^k\left(\nu^{1/2} + \alpha^{1/2}h + \delta_0^{-1/2} + \mu^{1/2}\right)\|\boldsymbol{s}\|_{k+1} \\ & + Ch^l\left(\mu^{-1/2}h + \delta^{1/2}\right)\|z\|_{l+1}.\end{aligned} \tag{3.6}$$

Now, (3.3) is proved by applying the triangle inequality to the splitting of the errors given at the beginning of the proof and the interpolation estimates (2.4). ∎

Remark 3.2. From the error estimate (3.3), an appropriate asymptotic value for the stabilization parameter μ can be derived. One has to take into consideration that this parameter does not appear only on the right-hand side of (3.3), but also in the definition of the norm on the left-hand side of (3.3). It turns out that for obtaining an optimal order of convergence for $||\nabla\cdot(\boldsymbol{s}-\boldsymbol{s}_h)||_0$, one has to choose μ to be independent of the mesh width.

LEMMA 3.3. *With the assumptions of Lemma* 3.1, *the* $L^2(\Omega)$ *error of the pressure is bounded as follows:*

$$\begin{aligned}\|z - z_h\|_0 \le\ & Ch^k\left(\nu^{1/2} + h\alpha^{1/2} + \mu^{1/2} + \delta_0^{-1/2}\right)^2\|\boldsymbol{s}\|_{k+1} \\ & + Ch^l\left[\left(\nu^{1/2} + h\alpha^{1/2} + \mu^{1/2} + \delta_0^{-1/2}\right)\left(\delta^{1/2} + h\mu^{-1/2}\right) + h\right]\|z\|_{l+1}.\end{aligned} \tag{3.7}$$

Proof. The splitting of the error (3.4) is applied again. With (2.7), one obtains

$$\|R_h\|_0 \le C\left(h\|\nabla R_h\|_0 + \sup_{\boldsymbol{v}_h\in V_h}\frac{(R_h, \nabla\cdot\boldsymbol{v}_h)}{\|\boldsymbol{v}_h\|_1}\right).$$

Using the definition of the stabilization parameter (3.2), one gets for the first term

$$h\|\nabla R_h\|_0 \le \frac{h}{\delta^{1/2}}\delta^{1/2}\|\nabla R_h\|_0 \le \delta_0^{-1/2}|||(\boldsymbol{E}_h, R_h)|||_h.$$

To estimate the second term, $q_h = 0$ is used in the error equation, giving with the Cauchy–Schwarz and the Poincaré inequality

$$\begin{aligned}\sup_{\boldsymbol{v}_h\in V_h}\frac{(R_h, \nabla\cdot\boldsymbol{v}_h)}{\|\boldsymbol{v}_h\|_1} \le\ & C\left(\nu^{1/2} + \alpha^{1/2} + \mu^{1/2}\right)|||(\boldsymbol{E}_h, R_h)|||_h \\ & + Ch^k\left(\nu + \alpha h + \mu\right)\|\boldsymbol{s}\|_{k+1} + Ch^{l+1}\|z\|_{l+1}.\end{aligned}$$

Adding both estimates, using (3.6), and applying the triangle inequality gives the statement of the lemma. ∎

Remark 3.4. Considering the steady-state problem (3.1) with the right-hand side $\partial_t\boldsymbol{g}$, then the solution of this problem is $(\partial_t\boldsymbol{s}, \partial_t z)$, thanks to the linearity of the problem and the independency of time of the coefficients. Exactly the same analysis leads then to error estimates for $|||(\partial_t(\boldsymbol{s}-\boldsymbol{s}_h), \partial_t(z-z_h))|||_h$ and $\|\partial_t(z-z_h)\|_0$, where the error bounds depend on Sobolev norms of $(\partial_t\boldsymbol{s}, \partial_t z)$.

4. Finite element error analysis for the evolutionary problem. This section will present several convergence results. In Theorem 4.2, an error bound is derived whose proof uses a rather restrictive assumption such that the result applies only for piecewise linear or affine bilinear finite element spaces. In this estimate, the pressure appears only together with the time derivative of the velocity. The estimate of Theorem 4.5 is valid for higher order elements and it bounds also the pressure error in the norm appearing in $|||\cdot|||_h$. The proof requires stronger regularity assumptions and the bound involves a pressure error at the initial time. A way to handle this term, by an appropriate choice of the initial condition for the discrete velocity, is described in Remark 4.8. An estimate for the pressure error in $L^2(0,t;L^2)$, again for piecewise linear or affine bilinear finite element spaces, is presented in Theorem 4.7. Finally, it is shown in Theorems 4.10 and 4.11 that the initial pressure error can be removed from the bounds if the proposed initial discrete velocity is used.

Remark 4.1. For the finite element error analysis, the errors are split

$$\boldsymbol{u}-\boldsymbol{u}_h = (\boldsymbol{u}-\boldsymbol{s}_h)-(\boldsymbol{u}_h-\boldsymbol{s}_h) = (\boldsymbol{u}-\boldsymbol{s}_h)-\boldsymbol{e}_h,$$

$$p-p_h = (p-z_h)-(p_h-z_h) = (p-z_h)-r_h,$$

where $(\boldsymbol{s}_h, z_h)$ is the solution of (2.6) with right-hand side $\boldsymbol{g}=\boldsymbol{f}-\partial_t\boldsymbol{u}$. Hence, the solution of the corresponding continuous problem is $(\boldsymbol{s}, z)=(\boldsymbol{u}, p)$ and the first terms on the right-hand sides can be bounded with the estimates derived in section 3.

THEOREM 4.2. *Let $(\boldsymbol{u}, p)$ be the solution of* (2.2) *with*

- $\boldsymbol{u}\in L^\infty(0,T;H^{k+1}(\Omega))$, $\boldsymbol{u}\in L^2(0,t;H^{k+1})$, $\partial_t\boldsymbol{u}\in L^2(0,t;H^{k+1})$,
- $p\in L^\infty(0,T;H^{l+1}(\Omega))$, $p\in L^2(0,t;H^{l+1})$, $\partial_t p\in L^2(0,t;H^{l+1})$

with $k\ge 1$, $l\ge 0$. Let $(\boldsymbol{u}_h, p_h)$ be the solution obtained with the PSPG method (2.5), *let the stabilization parameter satisfy* (3.2), *and let $\delta\le 1/4$. If the velocity finite element space satisfies*

$$\Delta \boldsymbol{v}_h = \boldsymbol{0} \quad \forall\, K\in\mathcal{T}_h,\ \forall\, \boldsymbol{v}_h\in V_h, \tag{4.1}$$

then the following error estimate holds for all $t\in(0,T]$

$$\begin{aligned}
&\|(\boldsymbol{u}-\boldsymbol{u}_h)(t)\|_0^2+\nu\delta\|\nabla(\boldsymbol{u}-\boldsymbol{u}_h)(t)\|_0^2+\mu\delta\|\nabla\cdot(\boldsymbol{u}-\boldsymbol{u}_h)(t)\|_0^2\\
&\quad+\nu\|\nabla(\boldsymbol{u}-\boldsymbol{u}_h)\|_{L^2(0,t;L^2)}^2+\alpha\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(0,t;L^2)}^2+\mu\|\nabla\cdot(\boldsymbol{u}-\boldsymbol{u}_h)\|_{L^2(0,t;L^2)}^2\\
&\quad+\delta\|\partial_t(\boldsymbol{u}-\boldsymbol{u}_h)+\nabla(p-p_h)\|_{L^2(0,t;L^2)}^2\\
&\le C\left(\|(\boldsymbol{u}-\boldsymbol{u}_h)(0)\|_0^2+\nu\delta\|\nabla(\boldsymbol{u}-\boldsymbol{u}_h)(0)\|_0^2+\mu\delta\|\nabla\cdot(\boldsymbol{u}-\boldsymbol{u}_h)(0)\|_0^2\right)\\
&\quad+C_1h^{2k}\left(\nu+h^2\alpha+\mu+\delta_0^{-1}\right)+C_2h^{2l}\left(\delta+h^2\mu^{-1}\right),
\end{aligned}\tag{4.2}$$

with

$$C_1 = C_1\left(\delta,\alpha^{-1},\|\boldsymbol{u}\|_{L^2(0,t;H^{k+1})}^2,\|\partial_t\boldsymbol{u}\|_{L^2(0,t;H^{k+1})}^2,\|\boldsymbol{u}(t)\|_{k+1}^2,\|\boldsymbol{u}_0\|_{k+1}^2\right),$$

$$C_2 = C_2\left(\delta,\alpha^{-1},\|p\|_{L^2(0,t;H^{l+1})}^2,\|\partial_t p\|_{L^2(0,t;H^{l+1})}^2,\|p(t)\|_{l+1}^2,\|p(0)\|_{l+1}^2\right).$$

Proof. The proof starts by introducing the decomposition of the error given in Remark 4.1. A straightforward calculation, subtracting (2.6) from (2.5), shows that for all $(\boldsymbol{v}_h, q_h) \in V_h \times Q_h$

$$
\begin{aligned}
&(\partial_t \boldsymbol{e}_h, \boldsymbol{v}_h) + \nu(\nabla \boldsymbol{e}_h, \nabla \boldsymbol{v}_h) + \alpha(\boldsymbol{e}_h, \boldsymbol{v}_h) \\
&\quad - (\nabla \cdot \boldsymbol{v}_h, r_h) + (\nabla \cdot \boldsymbol{e}_h, q_h) + \mu(\nabla \cdot \boldsymbol{e}_h, \nabla \cdot \boldsymbol{v}_h) \\
&\quad = (\boldsymbol{T}_{\mathrm{tr}}, \boldsymbol{v}_h) + \delta(\boldsymbol{T}_{\mathrm{tr}}, \nabla q_h) - \sum_{K \in \mathcal{T}_h} \delta(\partial_t \boldsymbol{e}_h, \nabla q_h)_K \\
&\qquad + \sum_{K \in \mathcal{T}_h} \delta(\nu \Delta \boldsymbol{e}_h - \alpha \boldsymbol{e}_h - \nabla r_h, \nabla q_h)_K,
\end{aligned}
\tag{4.3}
$$

with the truncation error $\boldsymbol{T}_{\mathrm{tr}} = \partial_t \boldsymbol{u} - \partial_t \boldsymbol{s}_h$.

Arguing like in [3], one picks two sets of test functions in (4.3) and the resulting equations are added. The first set is $(\boldsymbol{v}_h, q_h) = (\boldsymbol{e}_h, r_h)$ which gives

$$
\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\|\boldsymbol{e}_h\|_0^2 + \nu\|\nabla \boldsymbol{e}_h\|_0^2 + \alpha\|\boldsymbol{e}_h\|_0^2 + \mu\|\nabla \cdot \boldsymbol{e}_h\|_0^2 + \delta(\partial_t \boldsymbol{e}_h, \nabla r_h) + \delta\|\nabla r_h\|_0^2 \\
&\quad = (\boldsymbol{T}_{\mathrm{tr}}, \boldsymbol{e}_h) + \delta(\boldsymbol{T}_{\mathrm{tr}}, \nabla r_h) + \sum_{K \in \mathcal{T}_h} \delta(\nu \Delta \boldsymbol{e}_h - \alpha \boldsymbol{e}_h, \nabla r_h)_K.
\end{aligned}
\tag{4.4}
$$

Taking next $(\boldsymbol{v}_h, q_h) = (\delta \partial_t \boldsymbol{e}_h, 0)$ and applying integration by parts, using that $Q_h \subset H^1(\Omega)$, leads to

$$
\begin{aligned}
\delta\|\partial_t \boldsymbol{e}_h\|_0^2 + \frac{\nu\delta}{2}\frac{d}{dt}\|\nabla \boldsymbol{e}_h\|_0^2 + \frac{\alpha\delta}{2}\frac{d}{dt}\|\boldsymbol{e}_h\|_0^2 + \delta(\partial_t \boldsymbol{e}_h, \nabla r_h) + \frac{\mu\delta}{2}\frac{d}{dt}\|\nabla \cdot \boldsymbol{e}_h\|_0^2 \\
= \delta(\boldsymbol{T}_{\mathrm{tr}}, \partial_t \boldsymbol{e}_h).
\end{aligned}
\tag{4.5}
$$

Adding (4.4) and (4.5) and taking into account that

$$
\delta\|\partial_t \boldsymbol{e}_h + \nabla r_h\|_0^2 = \delta\|\partial_t \boldsymbol{e}_h\|_0^2 + 2\delta(\partial_t \boldsymbol{e}_h, \nabla r_h) + \delta\|\nabla r_h\|_0^2
$$

yields

$$
\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\left(\|\boldsymbol{e}_h\|_0^2 + \nu\delta\|\nabla \boldsymbol{e}_h\|_0^2 + \alpha\delta\|\boldsymbol{e}_h\|_0^2 + \mu\delta\|\nabla \cdot \boldsymbol{e}_h\|_0^2\right) \\
&\quad + \nu\|\nabla \boldsymbol{e}_h\|_0^2 + \alpha\|\boldsymbol{e}_h\|_0^2 + \mu\|\nabla \cdot \boldsymbol{e}_h\|_0^2 + \delta\|\partial_t \boldsymbol{e}_h + \nabla r_h\|_0^2 \\
&\quad = (\boldsymbol{T}_{\mathrm{tr}}, \boldsymbol{e}_h) + \delta(\boldsymbol{T}_{\mathrm{tr}}, \partial_t \boldsymbol{e}_h + \nabla r_h) + \sum_{K \in \mathcal{T}_h} \delta(\nu \Delta \boldsymbol{e}_h - \alpha \boldsymbol{e}_h, \nabla r_h)_K.
\end{aligned}
\tag{4.6}
$$

By assumption (4.1), the viscous term vanishes. For the reactive term, one has

$$
\begin{aligned}
-\sum_{K \in \mathcal{T}_h} \delta(\alpha \boldsymbol{e}_h, \nabla r_h)_K &= -\sum_{K \in \mathcal{T}_h} \delta(\alpha \boldsymbol{e}_h, \partial_t \boldsymbol{e}_h + \nabla r_h)_K + \sum_{K \in \mathcal{T}_h} \delta(\alpha \boldsymbol{e}_h, \partial_t \boldsymbol{e}_h)_K \\
&\le \delta\alpha\|\boldsymbol{e}_h\|_0^2 + \frac{1}{4}\delta\|\partial_t \boldsymbol{e}_h + \nabla r_h\|_0^2 + \frac{\alpha\delta}{2}\frac{d}{dt}\|\boldsymbol{e}_h\|_0^2.
\end{aligned}
$$

Using $\delta \le 1/4$, inserting this estimate into (4.6), and estimating the other terms on the right-hand side with standard arguments gives

$$
\begin{aligned}
\frac{1}{2}\frac{d}{dt}\left(\|\boldsymbol{e}_h\|_0^2 + \nu\delta\|\nabla \boldsymbol{e}_h\|_0^2 + \mu\delta\|\nabla \cdot \boldsymbol{e}_h\|_0^2\right) + \nu\|\nabla \boldsymbol{e}_h\|_0^2 + \frac{\alpha}{2}\|\boldsymbol{e}_h\|_0^2 \\
+ \mu\|\nabla \cdot \boldsymbol{e}_h\|_0^2 + \frac{\delta}{2}\|\partial_t \boldsymbol{e}_h + \nabla r_h\|_0^2 \le \left(\frac{1}{\alpha} + \delta\right)\|\boldsymbol{T}_{\mathrm{tr}}\|_0^2.
\end{aligned}
\tag{4.7}
$$

Applying now Remark 3.4 in combination with estimate (3.3) yields

$$\|\boldsymbol{T}_{\mathrm{tr}}\|_0^2 \le \frac{C}{\alpha}\left[h^{2k}\left(\nu + h^2\alpha + \mu + \delta_0^{-1}\right)\|\partial_t \boldsymbol{u}\|_{k+1}^2 + h^{2l}\left(\delta + h^2\mu^{-1}\right)\|\partial_t p\|_{l+1}^2\right]. \tag{4.8}$$

Inserting (4.8) into (4.7) and integrating between 0 and t leads to

$$\begin{aligned}
&\frac{1}{2}\left(\|\boldsymbol{e}_h(t)\|_0^2 + \nu\delta\|\nabla \boldsymbol{e}_h(t)\|_0^2 + \mu\delta\|\nabla\cdot \boldsymbol{e}_h(t)\|_0^2\right) + \nu\|\nabla \boldsymbol{e}_h\|_{L^2(0,t;L^2)}^2 \\
&\quad + \frac{\alpha}{2}\|\boldsymbol{e}_h\|_{L^2(0,t;L^2)}^2 + \mu\|\nabla\cdot \boldsymbol{e}_h\|_{L^2(0,t;L^2)}^2 + \frac{\delta}{2}\|\partial_t \boldsymbol{e}_h + \nabla r_h\|_{L^2(0,t;L^2)}^2 \\
&\le \frac{1}{2}\left(\|\boldsymbol{e}_h(0)\|_0^2 + \nu\delta\|\nabla \boldsymbol{e}_h(0)\|_0^2 + \mu\delta\|\nabla\cdot \boldsymbol{e}_h(0)\|_0^2\right) \\
&\quad + C\left(\frac{1}{\alpha^2} + \frac{\delta}{\alpha}\right)\Big[h^{2k}\left(\nu + h^2\alpha + \mu + \delta_0^{-1}\right)\|\partial_t \boldsymbol{u}\|_{L^2(0,t;H^{k+1})}^2 \\
&\qquad + h^{2l}\left(\delta + h^2\mu^{-1}\right)\|\partial_t p\|_{L^2(0,t;H^{l+1})}^2\Big].
\end{aligned} \tag{4.9}$$

The final step of the proof consists in using the decomposition of the errors given in Remark 4.1 and applying the triangle inequality. □

Remark 4.3. Assumption (4.1) restricts the analysis to $V_h = P_1$, to the space of unmapped Q_1 finite elements (which are defined directly on K), or to mapped Q_1 finite elements on grids consisting of mesh cells which are obtained by an affine transformation of the reference cell $[0,1]^d$ or $[-1,1]^d$. Thus, estimate (4.2) is particularly of interest for $k = l = 1$. If the initial condition is approximated sufficiently well, then the terms on the right-hand side of (4.2) are of order $\mathcal{O}(h^2)$. Thus, for all errors on the left-hand side, a first order convergence was shown.

Remark 4.4. As is well known (see [5]), the required regularity for the solution assumed in Theorem 4.2 (and the regularity that will be assumed in Theorems 4.5, 4.7, 4.10, and 4.11 below) holds only in the presence of nonlocal compatibility conditions of various orders. For example, if the norm $\|\partial_t \boldsymbol{u}(t)\|_1$ remains bounded as $t \to 0$, then the nonlocal compatibility condition $\nabla p(0)\mid_{\partial\Omega} = (\nu\Delta\boldsymbol{u}_0 + \boldsymbol{f}(\cdot,0))\mid_{\partial\Omega}$ must hold, where $p(0)$ is the solution of an overdetermined Neumann problem; see (4.31) below.

In the case that the compatibility conditions are not assumed, as in [5], error bounds can be obtained that contain negative powers of t, such that convergence is achieved except at time $t = 0$. The extension of the technique from [5] to get error bounds for the PSPG method, valid in the case in which nonlocal compatibility conditions will not be assumed, is outside the scope of the present paper. In all numerical studies presented in this paper, the analytical solution satisfies the compatibility conditions.

The unsatisfactory aspects of Theorem 4.2 are that it does not provide an error estimate for the pressure and that assumption (4.1) restricts the analysis to lowest order pairs of finite element spaces; see Remark 4.3. The pressure occurs only in combination with the temporal derivative of the velocity. In what follows, error bounds for the pressure and for higher order finite elements will be studied. The derivation of these bounds requires us, however, to assume higher regularity of the solution and one obtains terms on the right-hand side which involve the pressure error at the initial time. The last issue will be discussed below in Remark 4.8 and resolved for a special choice of the discrete initial velocity in Theorems 4.10 and 4.11.

THEOREM 4.5. *Let* $(\boldsymbol{u}, p)$ *be the solution of* (2.2) *with*

- $\boldsymbol{u} \in L^\infty(0,T;H^{k+1}(\Omega))$, $\boldsymbol{u} \in L^2(0,t;H^{k+1})$, $\partial_t\boldsymbol{u} \in L^2(0,t;H^{k+1})$, $\partial_{tt}\boldsymbol{u} \in L^2(0,t;H^{\max\{1,k-1\}})$,

- $p \in L^\infty(0,T;H^{l+1}(\Omega))$, $p \in L^2(0,t;H^{l+1})$, $\partial_t p \in L^2(0,t;H^{l+1})$, $\partial_{tt} p \in L^2(0,t;H^{\max\{1,l-1\}})$

with $k \geq 1$, $l \geq 0$. Let $(\boldsymbol{u}_h, p_h)$ be the solution obtained with the PSPG method (2.5)*, let the stabilization parameter satisfy* (3.2)*, and let*

$$\delta(1+8\alpha^2\delta) \leq \frac{1}{16}, \quad \delta \leq \frac{h^2}{4\mu C_{\mathrm{inv}}^2}. \tag{4.10}$$

Then the following error estimate holds for all $t \in (0,T]$:

$$\begin{aligned}
&\|(\boldsymbol{u}-\boldsymbol{u}_h)(t)\|_0^2 + \nu\|\nabla(\boldsymbol{u}-\boldsymbol{u}_h)\|_{L^2(0,t;L^2)}^2 + \alpha\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(0,t;L^2)}^2 \\
&\quad + \mu\|\nabla\cdot(\boldsymbol{u}-\boldsymbol{u}_h)\|_{L^2(0,t;L^2)}^2 + \delta\|\nabla(p-p_h)\|_{L^2(0,t;L^2)}^2 \\
&\quad \leq C\Big[\|(\boldsymbol{u}-\boldsymbol{u}_h)(0)\|_0^2 + \delta\nu\|\nabla(\boldsymbol{u}-\boldsymbol{u}_h)(0)\|_0^2 \\
&\qquad + \delta\mu\|\nabla\cdot(\boldsymbol{u}-\boldsymbol{u}_h)(0)\|_0^2 + \delta^2\|\nabla(p-p_h)(0)\|_0^2\Big] \\
&\quad + C_1 h^{2k}\left(\nu + h^2\alpha + \mu + \delta_0^{-1}\right) + C_2 h^{2l}\left(\delta + h^2\mu^{-1}\right)
\end{aligned} \tag{4.11}$$

with

$$\begin{aligned}
C_1 = C_1\Big(&\delta, \alpha, \alpha^{-1}, \mu^{-1}, \|\boldsymbol{u}\|_{L^2(0,t;H^{k+1})}^2, \|\partial_t\boldsymbol{u}\|_{L^2(0,t;H^{k+1})}^2, \\
&\|\partial_{tt}\boldsymbol{u}\|_{L^2(0,t;H^{\max\{1,k-1\}})}^2, \|\boldsymbol{u}(t)\|_{k+1}^2, \|\boldsymbol{u}_0\|_{k+1}^2, \|\partial_t\boldsymbol{u}(0)\|_{\max\{1,k-1\}}^2\Big),
\end{aligned} \tag{4.12}$$

$$\begin{aligned}
C_2 = C_2\Big(&\delta, \alpha, \alpha^{-1}, \mu^{-1}, \|p\|_{L^2(0,t;H^{l+1})}^2, \|\partial_t p\|_{L^2(0,t;H^{l+1})}^2, \\
&\|\partial_{tt} p\|_{L^2(0,t;H^{\max\{1,l-1\}})}^2, \|p(t)\|_{l+1}^2, \|p(0)\|_{l+1}^2, \|\partial_t p(0)\|_{\max\{1,l-1\}}^2\Big).
\end{aligned} \tag{4.13}$$

Proof. Applying standard estimates to (4.4), like the Cauchy–Schwarz inequality, Young's inequality, the inverse inequality (2.3) and (3.2), and integration in $(0,t)$ leads directly to

$$\begin{aligned}
&\|\boldsymbol{e}_h(t)\|_0^2 + \nu\|\nabla\boldsymbol{e}_h\|_{L^2(0,t;L^2)}^2 + \frac{\alpha}{2}\|\boldsymbol{e}_h\|_{L^2(0,t;L^2)}^2 + \mu\|\nabla\cdot\boldsymbol{e}_h\|_{L^2(0,t;L^2)}^2 \\
&\quad + \delta\|\nabla r_h\|_{L^2(0,t;L^2)}^2 \leq \|\boldsymbol{e}_h(0)\|_0^2 + 4\left(\delta + \frac{1}{\alpha}\right)\|\boldsymbol{T}_{\mathrm{tr}}\|_{L^2(0,t;L^2)}^2 + 4\delta\|\partial_t\boldsymbol{e}_h\|_{L^2(0,t;L^2)}^2.
\end{aligned} \tag{4.14}$$

The last term has to be bounded now. Using $\boldsymbol{v}_h = \delta\partial_t\boldsymbol{e}_h$ and $q_h = 0$ in (4.3) leads to

$$\begin{aligned}
&\delta\|\partial_t\boldsymbol{e}_h\|_0^2 + \frac{\nu\delta}{2}\frac{d}{dt}\|\nabla\boldsymbol{e}_h\|_0^2 + \frac{\alpha\delta}{2}\frac{d}{dt}\|\boldsymbol{e}_h\|_0^2 + \frac{\mu\delta}{2}\frac{d}{dt}\|\nabla\cdot\boldsymbol{e}_h\|_0^2 \\
&\quad = (\boldsymbol{T}_{\mathrm{tr}}, \delta\partial_t\boldsymbol{e}_h) + \delta(\nabla\cdot\partial_t\boldsymbol{e}_h, r_h) \\
&\quad \leq 4\delta\|\boldsymbol{T}_{\mathrm{tr}}\|_0^2 + \frac{\delta}{16}\|\partial_t\boldsymbol{e}_h\|_0^2 + |\delta(\nabla\cdot\partial_t\boldsymbol{e}_h, r_h)|.
\end{aligned} \tag{4.15}$$

To estimate the terms on the right-hand side of (4.15), differentiate (4.3) with respect to time, take $\boldsymbol{v}_h = \boldsymbol{0}$ and $q_h = \delta r_h$ as test functions to obtain

$$\begin{aligned}
\delta(\nabla\cdot\partial_t\boldsymbol{e}_h, r_h) &= \delta(\partial_t\boldsymbol{T}_{\mathrm{tr}}, \delta\nabla r_h) + \delta(-\partial_{tt}\boldsymbol{e}_h - \alpha\partial_t\boldsymbol{e}_h - \nabla\partial_t r_h, \delta\nabla r_h) \\
&\quad + \sum_{K\in\mathcal{T}_h}\delta(\nu\Delta\partial_t\boldsymbol{e}_h, \delta\nabla r_h)_K.
\end{aligned} \tag{4.16}$$

The first term on the right-hand side of (4.16) is bounded by adding and subtracting $\delta\partial_t \boldsymbol{e}_h$ and applying standard estimates

$$\begin{aligned}\delta(\partial_t \boldsymbol{T}_{\mathrm{tr}}, \delta\nabla r_h) &= \delta(\partial_t \boldsymbol{T}_{\mathrm{tr}}, \delta(\partial_t \boldsymbol{e}_h + \nabla r_h)) - \delta(\partial_t \boldsymbol{T}_{\mathrm{tr}}, \delta\partial_t \boldsymbol{e}_h)\\ &\le \frac{\delta^2}{2}\|\partial_t \boldsymbol{T}_{\mathrm{tr}}\|_0^2 + \frac{\delta^2}{2}\|\partial_t \boldsymbol{e}_h + \nabla r_h\|_0^2 + 4\delta^3\|\partial_t \boldsymbol{T}_{\mathrm{tr}}\|_0^2 + \frac{\delta}{16}\|\partial_t \boldsymbol{e}_h\|_0^2.\end{aligned}$$

For the second term on the right-hand side of (4.16), one also adds and subtracts $\delta\partial_t \boldsymbol{e}_h$ to obtain

$$\begin{aligned}&\delta^2(-\partial_{tt}\boldsymbol{e}_h - \alpha\partial_t \boldsymbol{e}_h - \nabla\partial_t r_h, \nabla r_h)\\ &\quad = \delta^2(-\partial_t(\partial_t \boldsymbol{e}_h + \nabla r_h), \partial_t \boldsymbol{e}_h + \nabla r_h) + \delta^2(\partial_{tt}\boldsymbol{e}_h + \alpha\partial_t \boldsymbol{e}_h + \nabla\partial_t r_h, \partial_t \boldsymbol{e}_h)\\ &\qquad - \delta^2(\alpha\partial_t \boldsymbol{e}_h, \partial_t \boldsymbol{e}_h + \nabla r_h)\\ &\quad \le -\delta^2\frac{d}{dt}\|\partial_t \boldsymbol{e}_h + \nabla r_h\|_0^2 + \frac{\delta}{16}\|\partial_t \boldsymbol{e}_h\|_0^2 + 4\alpha^2\delta^3\|\partial_t \boldsymbol{e}_h + \nabla r_h\|_0^2\\ &\qquad + \delta^2\left|(\partial_t(\partial_t \boldsymbol{e}_h + \alpha\boldsymbol{e}_h + \nabla r_h), \partial_t \boldsymbol{e}_h)\right|.\end{aligned}\tag{4.17}$$

Differentiating (4.3) with respect to time, applying the test functions $\boldsymbol{v}_h = \delta^2\partial_t \boldsymbol{e}_h$ and $q_h = 0$, and using integration by parts yields for the last term on the right-hand side of (4.17)

$$\begin{aligned}&\delta^2(\partial_t(\partial_t \boldsymbol{e}_h + \nabla r_h + \alpha\boldsymbol{e}_h), \partial_t \boldsymbol{e}_h)\\ &\quad = -\delta^2\nu\|\nabla\partial_t \boldsymbol{e}_h\|_0^2 - \delta^2\mu\|\nabla\cdot\partial_t \boldsymbol{e}_h\|_0^2 + \delta^2(\partial_t \boldsymbol{T}_{\mathrm{tr}}, \partial_t \boldsymbol{e}_h)\\ &\quad \le -\delta^2\nu\|\nabla\partial_t \boldsymbol{e}_h\|_0^2 - \delta^2\mu\|\nabla\cdot\partial_t \boldsymbol{e}_h\|_0^2 + 4\delta^3\|\partial_t \boldsymbol{T}_{\mathrm{tr}}\|_0^2 + \frac{\delta}{16}\|\partial_t \boldsymbol{e}_h\|_0^2.\end{aligned}$$

The last term on the right-hand side of (4.16) is bounded by using the Cauchy–Schwarz inequality, the inverse inequality (2.3), the definition (3.2) of δ, and Young's inequality

$$\begin{aligned}\sum_{K\in\mathcal{T}_h}\delta(\nu\Delta\partial_t \boldsymbol{e}_h, \delta\nabla r_h)_K &\le \delta^2\nu\frac{C_{\mathrm{inv}}^2}{h^2}\|\partial_t \boldsymbol{e}_h\|_0\|\nabla r_h\|_0\\ &\le \frac{\delta}{8}\|\partial_t \boldsymbol{e}_h\|_0\|\nabla r_h\|_0 \le \frac{\delta}{16}\|\partial_t \boldsymbol{e}_h\|_0^2 + \frac{\delta}{16}\|\nabla r_h\|_0^2.\end{aligned}\tag{4.18}$$

A straightforward calculation inserting the estimates leads to

$$\begin{aligned}&\frac{11}{16}\delta\|\partial_t \boldsymbol{e}_h\|_0^2 + \frac{\nu\delta}{2}\frac{d}{dt}\|\nabla\boldsymbol{e}_h\|_0^2 + \frac{\alpha\delta}{2}\frac{d}{dt}\|\boldsymbol{e}_h\|_0^2 + \frac{\mu\delta}{2}\frac{d}{dt}\|\nabla\cdot\boldsymbol{e}_h\|_0^2\\ &\qquad + \frac{\delta^2}{2}\frac{d}{dt}\|\partial_t \boldsymbol{e}_h + \nabla r_h\|_0^2 + \delta^2\nu\|\nabla\partial_t \boldsymbol{e}_h\|_0^2 + \delta^2\mu\|\nabla\cdot\partial_t \boldsymbol{e}_h\|_0^2\\ &\quad \le 4\delta\|\boldsymbol{T}_{\mathrm{tr}}\|_0^2 + \frac{\delta^2}{2}(1+8\delta)\|\partial_t \boldsymbol{T}_{\mathrm{tr}}\|_0^2\\ &\qquad + \frac{\delta^2}{2}\left(1+8\alpha^2\delta\right)\|\partial_t \boldsymbol{e}_h + \nabla r_h\|_0^2 + \frac{\delta}{16}\|\nabla r_h\|_0^2.\end{aligned}\tag{4.19}$$

For the next to last term, the triangle inequality and the upper bound (4.10) for the stabilization parameter are applied,

$$\begin{aligned}\frac{\delta^2}{2}\left(1+8\alpha^2\delta\right)\|\partial_t \boldsymbol{e}_h + \nabla r_h\|_0^2 &\le \delta^2\left(1+8\alpha^2\delta\right)\left(\|\partial_t \boldsymbol{e}_h\|_0^2 + \|\nabla r_h\|_0^2\right)\\ &\le \frac{\delta}{16}\|\partial_t \boldsymbol{e}_h\|_0^2 + \frac{\delta}{16}\|\nabla r_h\|_0^2.\end{aligned}\tag{4.20}$$

Absorbing the first term into the left-hand side of (4.19) gives

$$
\begin{aligned}
&\frac{5}{8}\delta\|\partial_t \boldsymbol{e}_h\|_0^2 + \frac{\nu\delta}{2}\frac{d}{dt}\|\nabla \boldsymbol{e}_h\|_0^2 + \frac{\alpha\delta}{2}\frac{d}{dt}\|\boldsymbol{e}_h\|_0^2 + \frac{\mu\delta}{2}\frac{d}{dt}\|\nabla\cdot \boldsymbol{e}_h\|_0^2 \\
&\quad + \frac{\delta^2}{2}\frac{d}{dt}\|\partial_t \boldsymbol{e}_h + \nabla r_h\|_0^2 + \delta^2\nu\|\nabla\partial_t \boldsymbol{e}_h\|_0^2 + \delta^2\mu\|\nabla\cdot\partial_t \boldsymbol{e}_h\|_0^2 \\
&\le 4\delta\|\boldsymbol{T}_{\mathrm{tr}}\|_0^2 + \frac{\delta^2}{2}(1+8\delta)\|\partial_t \boldsymbol{T}_{\mathrm{tr}}\|_0^2 + \frac{\delta}{8}\|\nabla r_h\|_0^2.
\end{aligned} \tag{4.21}
$$

Integrating (4.21) between 0 and t, applying an appropriate scaling, denoting unimportant constants by C, and restricting on the only important term on the left-hand side gives

$$
\begin{aligned}
&4\delta\|\partial_t \boldsymbol{e}_h\|_{L^2(0,t;L^2)}^2 \\
&\quad\le C\big(\nu\delta\|\nabla \boldsymbol{e}_h(0)\|_0^2 + \alpha\delta\|\boldsymbol{e}_h(0)\|_0^2 + \mu\delta\|\nabla\cdot \boldsymbol{e}_h(0)\|_0^2 + \delta^2\|(\partial_t \boldsymbol{e}_h + \nabla r_h)(0)\|_0^2 \\
&\qquad + \delta\|\boldsymbol{T}_{\mathrm{tr}}\|_{L^2(0,t;L^2)}^2 + \delta^2(1+\delta)\|\partial_t \boldsymbol{T}_{\mathrm{tr}}\|_{L^2(0,t;L^2)}^2\big) + \frac{4\delta}{5}\|\nabla r_h\|_{L^2(0,t;L^2)}^2.
\end{aligned} \tag{4.22}
$$

The last term on the right-hand side will be absorbed into the left-hand side of (4.14). The first factor in front of $\|\partial_t \boldsymbol{T}_{\mathrm{tr}}\|_{L^2(0,t;L^2)}^2$ can be estimated with (4.10) and then the terms with $\|\boldsymbol{T}_{\mathrm{tr}}\|_{L^2(0,t;L^2)}$ and $\|\partial_t \boldsymbol{T}_{\mathrm{tr}}\|_{L^2(0,t;L^2)}$ are estimated with (4.8), noting that (4.8) was derived without using assumption (4.1).

For the errors at time $t = 0$, the crucial term is the last one. First, the triangle inequality is applied. Taking in (4.3) $\boldsymbol{v}_h = \delta\partial_t \boldsymbol{e}_h$ and $q_h = 0$, and using (3.2) and (4.10), one gets

$$
\delta\|\partial_t \boldsymbol{e}_h\|_0^2 \le C\left(\nu\|\nabla \boldsymbol{e}_h\|_0^2 + \delta\alpha^2\|\boldsymbol{e}_h\|_0^2 + \delta\|\nabla r_h\|_0^2 + \mu\|\nabla\cdot \boldsymbol{e}_h\|_0^2 + \delta\|\boldsymbol{T}_{\mathrm{tr}}\|_0^2\right). \tag{4.23}
$$

Using this estimate for $t = 0$ yields

$$
\begin{aligned}
&\delta^2\|(\partial_t \boldsymbol{e}_h + \nabla r_h)(0)\|_0^2 \\
&\quad\le C\delta\left(\nu\|\nabla \boldsymbol{e}_h(0)\|_0^2 + \delta\alpha^2\|\boldsymbol{e}_h(0)\|_0^2 + \mu\|\nabla\cdot \boldsymbol{e}_h(0)\|_0^2 + \delta\|\boldsymbol{T}_{\mathrm{tr}}(0)\|_0^2 + \delta\|\nabla r_h(0)\|_0^2\right).
\end{aligned} \tag{4.24}
$$

Inserting (4.24) into (4.22), then inserting the result into (4.14), applying (4.8) in combination with (4.10) to estimate δ^2, the triangle inequality, (3.3), and noting that $\delta\alpha \le 1/4$ gives the statement of the theorem. ◻

Remark 4.6. With assumption (4.1) it is not necessary to perform the estimate (4.18) such that the last term on the right-hand side of (4.19) does not appear. Also, the term on the left-hand side of (4.20) can be estimated in a different way. Integrating (4.19) between 0 and t and applying (4.10) yields

$$
\begin{aligned}
&\delta^2\|\partial_t \boldsymbol{e}_h + \nabla r_h\|_0^2 \\
&\quad\le \nu\delta\|\nabla \boldsymbol{e}_h(0)\|_0^2 + \alpha\delta\|\boldsymbol{e}_h(0)\|_0^2 + \mu\delta\|\nabla\cdot \boldsymbol{e}_h(0)\|_0^2 + \delta^2\|(\partial_t \boldsymbol{e}_h + \nabla r_h)(0)\|_0^2 \\
&\qquad + 8\delta\|\boldsymbol{T}_{\mathrm{tr}}\|_{L^2(0,t;L^2)}^2 + \delta^2(1+8\delta)\|\partial_t \boldsymbol{T}_{\mathrm{tr}}\|_{L^2(0,t;L^2)}^2 + \frac{\delta}{16}\|\partial_t \boldsymbol{e}_h + \nabla r_h\|_{L^2(0,t;L^2)}^2.
\end{aligned}
$$

Now, the last term on the right-hand side can be bounded by (4.9). This alternative way leads to an estimate of form (4.22) without the last term on the right-hand side.

THEOREM 4.7. *Let all assumptions of Theorem* 4.5 *be satisfied and assume moreover that* (4.1) *holds, then*

$$\begin{aligned}\|p-p_h\|^2_{L^2(0,t;L^2)} \le C\Big[&\|(\boldsymbol{u}-\boldsymbol{u}_h)(0)\|_0^2 + \nu\|\nabla(\boldsymbol{u}-\boldsymbol{u}_h)(0)\|_0^2 \\ &+ \mu\|\nabla\cdot(\boldsymbol{u}-\boldsymbol{u}_h)(0)\|_0^2 + \delta\|\nabla(p-p_h)(0)\|_0^2\Big] \\ &+ C_1 h^{2k}\left(\nu + h^2\alpha + \mu + \delta_0^{-1}\right) + C_2 h^{2l}\left(\delta + h^2\mu^{-1}\right) \qquad (4.25)\end{aligned}$$

with

$$\begin{aligned}C_1 = C_1\Big(&\|\boldsymbol{u}\|^2_{L^2(0,t;H^{k+1})}, \|\partial_t\boldsymbol{u}\|^2_{L^2(0,t;H^{k+1})}, \|\partial_{tt}\boldsymbol{u}\|^2_{L^2(0,t;H^k)}, \\ &\|\boldsymbol{u}(t)\|^2_{k+1}, \|\boldsymbol{u}_0\|^2_{k+1}, \|\partial_t\boldsymbol{u}(0)\|^2_{\max\{1,k-1\}}\Big), \qquad (4.26)\\ C_2 = C_2\Big(&\|p\|^2_{L^2(0,t;H^{l+1})}, \|\partial_t p\|^2_{L^2(0,t;H^{l+1})}, \|\partial_{tt} p\|^2_{L^2(0,t;H^{\max\{1,l\}})}, \\ &\|p(t)\|^2_{l+1}, \|p(0)\|^2_{l+1}, \|\partial_t p(0)\|_{\max\{1,l-1\}}\Big). \qquad (4.27)\end{aligned}$$

All constants depend on $\nu, \delta, \delta_0^{-1}, \alpha, \alpha^{-1}, \mu, \mu^{-1}$, *but not on negative powers of* ν *and* δ.

Proof. Using (2.7) and (3.2) gives

$$\|r_h\|_0 \le C\delta_0^{-1/2}\delta^{1/2}\|\nabla r_h\|_0 + C\sup_{\boldsymbol{v}_h\in V_h}\frac{(\nabla\cdot\boldsymbol{v}_h, r_h)}{\|\boldsymbol{v}_h\|_1}. \qquad (4.28)$$

From (4.3) with $q_h = 0$, one obtains

$$\sup_{\boldsymbol{v}_h\in V_h}\frac{(r_h, \nabla\cdot\boldsymbol{v}_h)}{\|\boldsymbol{v}_h\|_1} \le \|\partial_t\boldsymbol{e}_h\|_0 + \nu\|\nabla\boldsymbol{e}_h\|_0 + \alpha\|\boldsymbol{e}_h\|_0 + \mu\|\nabla\cdot\boldsymbol{e}_h\|_0 + \|\boldsymbol{T}_{\mathrm{tr}}\|_0.$$

Inserting this estimate into (4.28) and integrating in $(0,T)$ yields

$$\begin{aligned}\|r_h\|^2_{L^2(0,t;L^2)} \le C\Big(&\delta_0^{-1}\delta\|\nabla r_h\|^2_{L^2(0,t;L^2)} + \nu^2\|\nabla\boldsymbol{e}_h\|^2_{L^2(0,t;L^2)} \\ &+ \alpha^2\|\boldsymbol{e}_h\|^2_{L^2(0,t;L^2)} + \mu^2\|\nabla\cdot\boldsymbol{e}_h\|^2_{L^2(0,t;L^2)} \\ &+ \|\boldsymbol{T}_{\mathrm{tr}}\|^2_{L^2(0,t;L^2)} + \|\partial_t\boldsymbol{e}_h\|^2_{L^2(0,t;L^2)}\Big).\end{aligned}$$

Applying (4.14) leads to

$$\begin{aligned}&\|r_h\|^2_{L^2(0,t;L^2)} \\ &\le C\max\{\delta_0^{-1},\nu,\alpha,\mu\}\left(\|\boldsymbol{e}_h(0)\|_0^2 + \left(\delta + \frac{1}{\alpha}\right)\|\boldsymbol{T}_{\mathrm{tr}}\|^2_{L^2(0,t;L^2)} + \delta\|\partial_t\boldsymbol{e}_h\|^2_{L^2(0,t;L^2)}\right) \\ &\quad + C\left(\|\boldsymbol{T}_{\mathrm{tr}}\|^2_{L^2(0,t;L^2)} + \|\partial_t\boldsymbol{e}_h\|^2_{L^2(0,t;L^2)}\right) \\ &\le C\left(\|\boldsymbol{e}_h(0)\|_0^2 + \left(\delta + \frac{1}{\alpha} + 1\right)\|\boldsymbol{T}_{\mathrm{tr}}\|^2_{L^2(0,t;L^2)} + (1+\delta)\|\partial_t\boldsymbol{e}_h\|^2_{L^2(0,t;L^2)}\right),\end{aligned}$$

where the constant in the last line depends on $\max\{\delta_0^{-1},\nu,\alpha,\mu\}$. To conclude the proof, one needs to bound $\|\partial_t\boldsymbol{e}_h\|^2_{L^2(0,t;L^2)}$. This term was estimated already in (4.22), where one has to take into account that with assumption (4.1) the last term on the right-hand side of (4.22) vanishes; cf., Remark 4.6. Now the proof finishes as the proof of Theorem 4.5. □

Remark 4.8. In the bounds (4.11) and (4.25) an error of the initial pressure $(p - p_h)(0)$ is contained. Whereas the initial velocity is part of the definition of the problem, an initial pressure is not given. However, there are temporal discretizations which require an initial pressure, like higher order Runge–Kutta schemes (e.g., see [10]), such that this issue does not appear only in the analysis presented in this paper. Next, a way for computing an initial approximation to the velocity will be described that has the advantage that the error of the initial pressure on the right-hand side of (4.11) and (4.25) disappears from the bounds.

As an approximation for the initial velocity it is suggested to take $\boldsymbol{u}_h(0) = \boldsymbol{s}_h(0)$. From the definition of $\boldsymbol{s}_h$ (see (2.6)), it follows that a steady-state Stokes problem at the initial time has to be solved with the PSPG method. More precisely, the problem consists in computing $(\boldsymbol{s}_h(0), z_h(0)) \in V_h \times Q_h$ such that for all $(\boldsymbol{v}_h, q_h) \in V_h \times Q_h$

$$\begin{aligned}
&\nu(\nabla \boldsymbol{s}_h(0), \nabla \boldsymbol{v}_h) + \alpha(\boldsymbol{s}_h(0), \boldsymbol{v}_h) - (\nabla \cdot \boldsymbol{v}_h, z_h(0)) + (\nabla \cdot \boldsymbol{s}_h(0), q_h) \\
&\quad + \mu(\nabla \cdot \boldsymbol{s}_h(0), \nabla \cdot \boldsymbol{v}_h) \\
&\quad = (\mathbf{f}(0) - \partial_t \boldsymbol{u}(0), \boldsymbol{v}_h) \\
&\qquad + \sum_{K \in \mathcal{T}_h} \delta(\mathbf{f}(0) - \partial_t \boldsymbol{u}(0) + \nu \Delta \mathbf{s}_h(0) - \alpha \boldsymbol{s}_h(0) - \nabla z_h(0), \nabla q_h)_K.
\end{aligned} \tag{4.29}$$

Since $\partial_t \boldsymbol{u}(0)$ is generally not known, the term $\mathbf{f}(0) - \partial_t \boldsymbol{u}(0)$ is replaced by the limit of the momentum balance equation for $t \to 0$

$$\mathbf{f}(0) - \partial_t \boldsymbol{u}(0) = -\nu \Delta \boldsymbol{u}_0 + \alpha \mathbf{u}_0 + \nabla p(0), \tag{4.30}$$

where $p(0)$ is the pressure at time $t = 0$. Following [5], the initial pressure $p(0)$ is the solution of the problem

$$\begin{aligned}
\Delta p(0) &= \nabla \cdot \boldsymbol{f}(0) && \text{in } \Omega, \\
\frac{\partial p(0)}{\partial \boldsymbol{n}} &= (\nu \Delta \boldsymbol{u}_0 + \boldsymbol{f}(0)) \cdot \boldsymbol{n} && \text{on } \partial\Omega,
\end{aligned} \tag{4.31}$$

where $\boldsymbol{n}$ is the outward pointing unit normal vector. Problem (4.31) defines a unique pressure up to a constant, which is however not important if the pressure is inserted into (4.30). The solution of (4.31) can be approximated by solving a discrete problem.

It will be proved in Theorems 4.10 and 4.11 that choosing as initial discrete velocity $\boldsymbol{u}_h(0) = \boldsymbol{s}_h(0)$ has the advantage that error bounds of forms (4.11) and (4.25) can be obtained where the bounds depend only on the initial error of the velocity but not on the initial error of the pressure.

Remark 4.9. It should be noted that if any other discrete initial velocity is considered, the so far presented error analysis can also be applied. However, the error bounds will depend on the error of the pressure at the initial time.

THEOREM 4.10. *Let the assumptions of Theorem* 4.5 *be satisfied and assume moreover that* $\boldsymbol{u}_h(0) = \boldsymbol{s}_h(0)$*, then*

$$\begin{aligned}
&\|(\boldsymbol{u} - \boldsymbol{u}_h)(t)\|_0^2 + \nu \|\nabla(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(0,t;L^2)}^2 + \alpha \|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(0,t;L^2)}^2 \\
&\quad + \mu \|\nabla \cdot (\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(0,t;L^2)}^2 + \delta \|\nabla(p - p_h)\|_{L^2(0,t;L^2)}^2 \\
&\quad \le C\Big[\|(\boldsymbol{u} - \boldsymbol{u}_h)(0)\|_0^2 + \delta\nu \|\nabla(\boldsymbol{u} - \boldsymbol{u}_h)(0)\|_0^2 + \delta\mu \|\nabla \cdot (\boldsymbol{u} - \boldsymbol{u}_h)(0)\|_0^2\Big] \\
&\qquad + C_1 h^{2k}\left(\nu + h^2\alpha + \mu + \delta_0^{-1}\right) + C_2 h^{2l}\left(\delta + h^2\mu^{-1}\right),
\end{aligned} \tag{4.32}$$

where the constants' dependencies are the same as given in (4.12) *and* (4.13).

Proof. With the initial condition $\boldsymbol{u}_h(0)=\boldsymbol{s}_h(0)$, the term $\|(\partial_t\boldsymbol{e}_h+\nabla r_h)(0)\|_0$ can be estimated in a different way as in the proof of Theorem 4.5.

Using $\boldsymbol{u}_h(0)=\boldsymbol{s}_h(0)$ and considering (2.5) with $q_h=0$ at time $t=0$ yields

$$\begin{aligned}\nu(\nabla\boldsymbol{s}_h(0),\nabla\boldsymbol{v}_h)+\alpha(\boldsymbol{s}_h(0),\boldsymbol{v}_h)+(\nabla p_h(0),\boldsymbol{v}_h)+\mu(\nabla\cdot\boldsymbol{s}_h(0),\nabla\cdot\boldsymbol{v}_h)&\\=(\boldsymbol{f}(0)-\partial_t\boldsymbol{u}_h(0),\boldsymbol{v}_h).&\end{aligned}\tag{4.33}$$

Similarly, one obtains from (4.29) with $q_h=0$

$$\begin{aligned}\nu(\nabla\boldsymbol{s}_h(0),\nabla\boldsymbol{v}_h)+\alpha(\boldsymbol{s}_h(0),\boldsymbol{v}_h)+(\nabla z_h(0),\boldsymbol{v}_h)+\mu(\nabla\cdot\boldsymbol{s}_h(0),\nabla\cdot\boldsymbol{v}_h)&\\=(\boldsymbol{f}(0)-\partial_t\boldsymbol{u}(0),\boldsymbol{v}_h).&\end{aligned}\tag{4.34}$$

Subtracting (4.34) from (4.33) leads to

$$(\nabla p_h(0)+\partial_t\boldsymbol{u}_h(0),\boldsymbol{v}_h)=(\nabla z_h(0)+\partial_t\boldsymbol{u}(0),\boldsymbol{v}_h).\tag{4.35}$$

Considering again (2.5) at time $t=0$ with $\boldsymbol{v}_h=\boldsymbol{0}$ gives

$$\begin{aligned}&(\nabla\cdot\boldsymbol{s}_h(0),q_h)\\&\quad=\sum_{K\in\mathcal{T}_h}\delta(\mathbf{f}(0)-\partial_t\boldsymbol{u}_h(0)+\nu\Delta\boldsymbol{s}_h(0)-\alpha\boldsymbol{s}_h(0)-\nabla p_h(0),\nabla q_h)_K.\end{aligned}\tag{4.36}$$

Likewise (4.29) with $\boldsymbol{v}_h=\boldsymbol{0}$ yields

$$\begin{aligned}&(\nabla\cdot\boldsymbol{s}_h(0),q_h)\\&\quad=\sum_{K\in\mathcal{T}_h}\delta(\mathbf{f}(0)-\partial_t\boldsymbol{u}(0)+\nu\Delta\boldsymbol{s}_h(0)-\alpha\boldsymbol{s}_h(0)-\nabla z_h(0),\nabla q_h)_K,\end{aligned}\tag{4.37}$$

such that the difference of (4.37) and (4.36) is

$$(\nabla p_h(0)+\partial_t\boldsymbol{u}_h(0),\nabla q_h)=(\nabla z_h(0)+\partial_t\boldsymbol{u}(0),\nabla q_h).\tag{4.38}$$

Adding (4.35) and (4.38), and then adding and subtracting $\partial_t\boldsymbol{s}_h(0)$ gives

$$((\partial_t\boldsymbol{e}_h+\nabla r_h)(0),\boldsymbol{v}_h+\nabla q_h)=(\partial_t\boldsymbol{u}(0)-\partial_t\boldsymbol{s}_h(0),\boldsymbol{v}_h+\nabla q_h)$$

for all $(\boldsymbol{v}_h,q_h)$. Choosing $\boldsymbol{v}_h=\partial_t\boldsymbol{e}_h(0)$ and $q_h=r_h(0)$, one obtains

$$\|(\partial_t\boldsymbol{e}_h+\nabla r_h)(0)\|_0\le\|\partial_t\boldsymbol{u}(0)-\partial_t\boldsymbol{s}_h(0)\|_0=\|\boldsymbol{T}_{\mathrm{tr}}(0)\|_0.\tag{4.39}$$

Now, one can follow the steps of the proof of Theorem 4.5 until (4.22). Instead of the estimate used in Theorem 4.5 to bound $\delta^2\|(\partial_t\boldsymbol{e}_h+\nabla r_h)(0)\|_0$, one applies (4.39) together with (4.10) and then (4.32) follows immediately. ∎

THEOREM 4.11. *Let all assumptions of Theorem* 4.10 *be satisfied and assume moreover that* (4.1) *holds, then*

$$\begin{aligned}\|p-p_h\|^2_{L^2(0,t;L^2)}\le C\Big[&\|(\boldsymbol{u}-\boldsymbol{u}_h)(0)\|_0^2+\nu\|\nabla(\boldsymbol{u}-\boldsymbol{u}_h)(0)\|_0^2+\mu\|\nabla\cdot(\boldsymbol{u}-\boldsymbol{u}_h)(0)\|_0^2\Big]\\&+C_1h^{2k}\left(\nu+h^2\alpha+\mu+\delta_0^{-1}\right)+C_2h^{2l}\left(\delta+h^2\mu^{-1}\right),\end{aligned}$$

where the constants posses the same dependencies as given in (4.26) *and* (4.27) *and all constants depend on* $\nu,\delta,\delta_0^{-1},\alpha,\alpha^{-1},\mu,\mu^{-1}$, *but not on negative powers of* ν *and* δ.

Proof. The proof follows the steps of the proof of Theorem 4.7 with the only difference that one now applies (4.39) to estimate $\delta^2\|(\partial_t\boldsymbol{e}_h+\nabla r_h)(0)\|_0$ in (4.22). ∎

5. Fully discrete case with backward Euler scheme. This section transfers the results of Theorem 4.5 to the fully discrete situation, where the backward Euler scheme is used as the time integrator. It turns out that the analysis of the first time step needs special care and that again the use of the initial condition as proposed in Remark 4.8 is important to derive the desired results.

The fully discrete approximation reads as follows: Find $(\boldsymbol{u}_h^n, p_h^n) \in V_h \times Q_h$ such that for all $\boldsymbol{v}_h \in V_h, q_h \in Q_h \subset H^1(\Omega)$

$$
\begin{aligned}
&\left(\frac{\boldsymbol{u}_h^n - \boldsymbol{u}_h^{n-1}}{\tau}, \boldsymbol{v}_h\right) + \nu(\nabla \boldsymbol{u}_h^n, \nabla \boldsymbol{v}_h) + \alpha(\boldsymbol{u}_h^n, \boldsymbol{v}_h) - (\nabla \cdot \boldsymbol{v}_h, p_h^n) + (\nabla \cdot \boldsymbol{u}_h^n, q_h) \\
&\qquad + \mu(\nabla \cdot \boldsymbol{u}_h^n, \nabla \cdot \boldsymbol{v}_h) \\
&\quad = (\boldsymbol{f}^n, \boldsymbol{v}_h) + \sum_{K \in \mathcal{T}_h} \delta \left(\boldsymbol{f}^n - \frac{\boldsymbol{u}_h^n - \boldsymbol{u}_h^{n-1}}{\tau} + \nu \Delta \boldsymbol{u}_h^n - \alpha \boldsymbol{u}_h^n - \nabla p_h^n, \nabla q_h\right)_K .
\end{aligned} \tag{5.1}
$$

Here, $\boldsymbol{u}_h^n = \boldsymbol{u}_h(t_n)$ and τ denotes the length of the equidistant time step.

Derivation of an error equation. The error is decomposed in the following way:

$$
\begin{aligned}
\boldsymbol{u}(t_n) - \boldsymbol{u}_h^n &= (\boldsymbol{u}(t_n) - \boldsymbol{s}_h^n) - (\boldsymbol{u}_h^n - \boldsymbol{s}_h^n) = (\boldsymbol{u}(t_n) - \boldsymbol{s}_h^n) - \boldsymbol{e}_h^n, \\
p(t_n) - p_h^n &= (p(t_n) - z_h^n) - (p_h^n - z_h^n) = (p(t_n) - z_h^n) - r_h^n,
\end{aligned} \tag{5.2}
$$

where $(\boldsymbol{s}_h^n, r_h^n)$ is the PSPG solution to (2.6) with right-hand side $\boldsymbol{g} = \boldsymbol{f}(t_n) - \partial_t \boldsymbol{u}(t_n)$. Using (2.6), one finds readily that

$$
\begin{aligned}
&\left(\frac{\boldsymbol{s}_h^n - \boldsymbol{s}_h^{n-1}}{\tau}, \boldsymbol{v}_h\right) + \nu(\nabla \boldsymbol{s}_h^n, \nabla \boldsymbol{v}_h) + \alpha(\boldsymbol{s}_h^n, \boldsymbol{v}_h) \\
&\quad - (\nabla \cdot \boldsymbol{v}_h, z_h^n) + (\nabla \cdot \boldsymbol{s}_h^n, q_h) + \mu(\nabla \cdot \boldsymbol{s}_h^n, \nabla \cdot \boldsymbol{v}_h) \\
&= (\boldsymbol{f}^n, \boldsymbol{v}_h) + (-T_{\mathrm{tr}}^n, \boldsymbol{v}_h) + \sum_{K \in \mathcal{T}_h} \delta \left(\boldsymbol{f}^n - \frac{\boldsymbol{s}_h^n - \boldsymbol{s}_h^{n-1}}{\tau} + \nu \Delta \boldsymbol{s}_h^n - \alpha \boldsymbol{s}_h^n - \nabla z_h^n, \nabla q_h\right)_K \\
&\quad + \sum_{K \in \mathcal{T}_h} \delta(-T_{\mathrm{tr}}^n, \nabla q_h)_K,
\end{aligned} \tag{5.3}
$$

where the truncation error is given by

$$
T_{\mathrm{tr}}^n = \partial_t \boldsymbol{u}(t_n) - \frac{\boldsymbol{s}_h^n - \boldsymbol{s}_h^{n-1}}{\tau} = (\partial_t \boldsymbol{u}(t_n) - \partial_t \boldsymbol{s}_h^n) + \left(\partial_t \boldsymbol{s}_h^n - \frac{\boldsymbol{s}_h^n - \boldsymbol{s}_h^{n-1}}{\tau}\right). \tag{5.4}
$$

The backward difference between values of functions at subsequent time steps will be denoted with the subscript τ, e.g., $\boldsymbol{e}_{h,\tau}^n = (\boldsymbol{e}_h^n - \boldsymbol{e}_h^{n-1})/\tau$. Subtracting (5.3) from (5.1) yields the error equation

$$
\begin{aligned}
&(\boldsymbol{e}_{h,\tau}^n, \boldsymbol{v}_h) + \nu(\nabla \boldsymbol{e}_h^n, \nabla \boldsymbol{v}_h) + \alpha(\boldsymbol{e}_h^n, \boldsymbol{v}_h) - (\nabla \cdot \boldsymbol{v}_h, r_h^n) + (\nabla \cdot \boldsymbol{e}_h^n, q_h) + \mu(\nabla \cdot \boldsymbol{e}_h^n, \nabla \cdot \boldsymbol{v}_h) \\
&\quad = (T_{\mathrm{tr}}^n, \boldsymbol{v}_h) + \delta(T_{\mathrm{tr}}^n, \nabla q_h) - \delta(\boldsymbol{e}_{h,\tau}^n, \nabla q_h) + \sum_{K \in \mathcal{T}_h} \delta(\nu \Delta \boldsymbol{e}_h^n - \alpha \boldsymbol{e}_h^n - \nabla r_h^n, \nabla q_h)_K .
\end{aligned} \tag{5.5}
$$

Inserting the error as test function. Taking $\boldsymbol{v}_h = \boldsymbol{e}_h^n$ and $q_h = r_h^n$ in (5.5) gives

$$
\begin{aligned}
&(\boldsymbol{e}_{h,\tau}^n, \boldsymbol{e}_h^n) + \nu(\nabla \boldsymbol{e}_h^n, \nabla \boldsymbol{e}_h^n) + \alpha(\boldsymbol{e}_h^n, \boldsymbol{e}_h^n) - (\nabla \cdot \boldsymbol{e}_h^n, r_h^n) + (\nabla \cdot \boldsymbol{e}_h^n, r_h^n) + \mu(\nabla \cdot \boldsymbol{e}_h^n, \nabla \cdot \boldsymbol{e}_h^n) \\
&\quad = (T_{\mathrm{tr}}^n, \boldsymbol{e}_h^n) + \delta(T_{\mathrm{tr}}^n, \nabla r_h^n) - \delta(\boldsymbol{e}_{h,\tau}^n, \nabla r_h^n) + \sum_{K \in \mathcal{T}_h} \delta(\nu \Delta \boldsymbol{e}_h^n - \alpha \boldsymbol{e}_h^n - \nabla r_h^n, \nabla r_h^n)_K .
\end{aligned} \tag{5.6}
$$

With the Cauchy–Schwarz inequality, the inverse estimate, Young's inequality, and (3.2), one obtains

$$(\boldsymbol{e}_{h,\tau}^n, \boldsymbol{e}_h^n) + \frac{3\nu}{4}\|\nabla \boldsymbol{e}_h^n\|_0^2 + \frac{\alpha}{4}\|\boldsymbol{e}_h^n\|_0^2 + \mu\|\nabla\cdot \boldsymbol{e}_h^n\|_0^2 + \frac{\delta}{2}\|\nabla r_h^n\|_0^2 \le \left(2\delta + \frac{1}{\alpha}\right)\|T_{\rm tr}^n\|_0^2 + 2\delta\|\boldsymbol{e}_{h,\tau}\|_0^2.$$

Taking into account that $(\boldsymbol{e}_{h,\tau}^n, \boldsymbol{e}_h^n) = (\|\boldsymbol{e}_h^n\|_0^2 - \|\boldsymbol{e}_h^{n-1}\|_0^2 + \|\boldsymbol{e}_h^n - \boldsymbol{e}_h^{n-1}\|_0^2)/(2\tau)$ and summing up from $j=1$ to n, one gets

$$\begin{aligned}(5.7)\quad &\|\boldsymbol{e}_h^n\|_0^2 + \frac{3\tau}{2}\sum_{j=1}^n \nu\|\nabla \boldsymbol{e}_h^j\|_0^2 + \frac{\tau\alpha}{2}\sum_{j=1}^n\|\boldsymbol{e}_h^j\|_0^2 + 2\tau\mu\sum_{j=1}^n\|\nabla\cdot \boldsymbol{e}_h^j\|_0^2 \\ &+ \tau\delta\sum_{j=1}^n\|\nabla r_h^j\|_0^2 \le \|\boldsymbol{e}_h^0\|_0^2 + 2\tau\left(2\delta+\frac{1}{\alpha}\right)\sum_{j=1}^n\|T_{\rm tr}^j\|_0^2 + 4\tau\delta\sum_{j=1}^n\|\boldsymbol{e}_{h,\tau}^j\|_0^2.\end{aligned}$$

Estimate of the last term on the right-hand side of (5.7) *for* $n \ge 2$. Using $\boldsymbol{v}_h = \delta \boldsymbol{e}_{h,\tau}^n$ and $q_h = 0$ in (5.5) leads to

$$\begin{aligned}(5.8)\quad &\delta\|\boldsymbol{e}_{h,\tau}^n\|_0^2 + \nu\delta(\nabla \boldsymbol{e}_h^n, \nabla \boldsymbol{e}_{h,\tau}^n) + \alpha\delta(\boldsymbol{e}_h^n, \boldsymbol{e}_{h,\tau}^n) + \mu\delta(\nabla\cdot \boldsymbol{e}_h^n, \nabla\cdot \boldsymbol{e}_{h,\tau}^n) \\ &\le 4\delta\|T_{\rm tr}^n\|_0^2 + \frac{\delta}{16}\|\boldsymbol{e}_{h,\tau}^n\|_0^2 + |\delta(\nabla\cdot \boldsymbol{e}_{h,\tau}^n, r_h^n)|.\end{aligned}$$

Now, one needs to estimate the last term on the right-hand side of (5.8). Taking first $\boldsymbol{v}_h = \boldsymbol{0}$ in (5.5) yields

$$(5.9)\quad (\nabla\cdot \boldsymbol{e}_h^n, q_h) = \delta(T_{\rm tr}^n, \nabla q_h) - \delta(\boldsymbol{e}_{h,\tau}^n + \alpha \boldsymbol{e}_h^n + \nabla r_h^n, \nabla q_h) + \delta\sum_{K\in\mathcal{T}_h}(\nu\Delta \boldsymbol{e}_h^n, \nabla q_h)_K.$$

The same equation is considered for t_{n-1}, both equations are subtracted, the result is divided by τ, and $q_h = \delta r_h^n$ is taken. Denoting $\boldsymbol{z}_h^n = \boldsymbol{e}_{h,\tau}^n$ gives for $n\ge 2$

$$(5.10)\quad \begin{aligned}\delta(\nabla\cdot \boldsymbol{e}_{h,\tau}^n, r_h^n) &= \delta(T_{{\rm tr},\tau}^n, \delta\nabla r_h^n) - \delta(\boldsymbol{z}_{h,\tau}^n + \alpha \boldsymbol{e}_{h,\tau}^n + \nabla r_{h,\tau}^n, \delta\nabla r_h^n) \\ &\quad + \sum_{K\in\mathcal{T}_h}\delta(\nu\Delta \boldsymbol{e}_{h,\tau}^n, \delta\nabla r_h^n)_K.\end{aligned}$$

Note that with the initial conditions proposed in Remark 4.8, one has $\boldsymbol{e}_h^0 = \boldsymbol{0}$. The first term on the right-hand side of (5.10) is bounded by adding and subtracting $\delta \boldsymbol{e}_{h,\tau}$ and applying the Cauchy–Schwarz inequality and Young's inequality

$$\begin{aligned}(5.11)\quad \delta(T_{{\rm tr},\tau}^n, \delta\nabla r_h^n) &= \delta(T_{{\rm tr},\tau}^n, \delta(\boldsymbol{e}_{h,\tau}^n + \nabla r_h^n)) - \delta(T_{{\rm tr},\tau}^n, \delta \boldsymbol{e}_{h,\tau}^n) \\ &\le \frac{\delta^2}{2}\|T_{{\rm tr},\tau}^n\|_0^2 + \frac{\delta^2}{2}\|\boldsymbol{e}_{h,\tau} + \nabla r_h^n\|_0^2 + 4\delta^3\|T_{{\rm tr},\tau}^n\|_0^2 + \frac{\delta}{16}\|\boldsymbol{e}_{h,\tau}^n\|_0^2.\end{aligned}$$

For the second term on the right-hand side of (5.10), one also adds and subtracts $\delta \boldsymbol{e}_{h,\tau}$ in the second argument and uses the definition of $\boldsymbol{z}_h^n$:

$$(5.12)\quad \begin{aligned}&-\delta^2(\boldsymbol{z}_{h,\tau}^n + \alpha \boldsymbol{e}_{h,\tau}^n + \nabla r_{h,\tau}^n, \nabla r_h^n) \\ &\le -\delta^2((\boldsymbol{z}_h^n + \nabla r_h^n)_\tau, \boldsymbol{z}_h^n + \nabla r_h^n) + \frac{\delta}{16}\|\boldsymbol{e}_{h,\tau}^n\|_0^2 + 4\alpha^2\delta^3\|\boldsymbol{e}_{h,\tau}^n + \nabla r_h^n\|_0^2 \\ &\quad + \delta^2((\boldsymbol{e}_{h,\tau}^n + \alpha \boldsymbol{e}_h^n + \nabla r_h^n)_\tau, \boldsymbol{e}_{h,\tau}^n).\end{aligned}$$

Subtracting (5.5) with $q_h = 0$ for t_{n-1} from the same equation for t_n, and dividing the result by τ yields

$$(5.13)\quad (\boldsymbol{z}_{h,\tau}^n, \boldsymbol{v}_h) + \nu(\nabla \boldsymbol{e}_{h,\tau}^n, \nabla \boldsymbol{v}_h) + \alpha(\boldsymbol{e}_{h,\tau}^n, \boldsymbol{v}_h) - (\nabla \cdot \boldsymbol{v}_h, r_{h,\tau}^n) + \mu(\nabla \cdot \boldsymbol{e}_{h,\tau}^n, \nabla \cdot \boldsymbol{v}_h) = (T_{\mathrm{tr},\tau}^n, \boldsymbol{v}_h).$$

Choosing $\boldsymbol{v}_h = \delta^2 \boldsymbol{e}_{h,\tau}^n$ in (5.13) and applying integration by parts gives

$$(5.14)\quad \delta^2((\boldsymbol{e}_{h,\tau}^n + \nabla r_h^n + \alpha \boldsymbol{e}_h^n)_\tau, \boldsymbol{e}_{h,\tau}^n) \le -\delta^2\nu\|\nabla \boldsymbol{e}_{h,\tau}^n\|_0^2 - \delta^2\mu\|\nabla\cdot \boldsymbol{e}_{h,\tau}^n\|_0^2 + 4\delta^3\|T_{\mathrm{tr},\tau}^n\|_0^2 + \frac{\delta}{16}\|\boldsymbol{e}_{h,\tau}^n\|_0^2.$$

The last term on the right-hand side of (5.10) is bounded by using the Cauchy–Schwarz inequality, the inverse inequality, the definition (3.2) δ, and Young's inequality

$$(5.15)\quad \sum_{K\in\mathcal{T}_h} \delta(\nu\Delta \boldsymbol{e}_{h,\tau}^n, \delta\nabla r_h^n)_K \le \delta^2\nu\frac{C_{\mathrm{inv}}^2}{h^2}\|\boldsymbol{e}_{h,\tau}^n\|_0\|\nabla r_h^n\|_0 \le \frac{\delta}{16}\|\boldsymbol{e}_{h,\tau}^n\|_0^2 + \frac{\delta}{16}\|\nabla r_h^n\|_0^2.$$

Inserting (5.10), (5.11), (5.12), (5.14), and (5.15) into (5.8) leads to

$$(5.16)\quad \begin{aligned} &\frac{11}{16}\delta\|\boldsymbol{e}_{h,\tau}^n\|_0^2 + \nu\delta(\nabla \boldsymbol{e}_{h,\tau}^n, \nabla \boldsymbol{e}_h^n) + \alpha\delta(\boldsymbol{e}_{h,\tau}^n, \boldsymbol{e}_h^n) + \mu\delta(\nabla\cdot \boldsymbol{e}_{h,\tau}^n, \nabla\cdot \boldsymbol{e}_h^n)\\ &\quad + \delta^2((\boldsymbol{e}_{h,\tau}^n + \nabla r_h^n)_\tau, \boldsymbol{e}_{h,\tau}^n + \nabla r_h^n) + \delta^2\nu\|\nabla \boldsymbol{e}_{h,\tau}^n\|_0^2 + \delta^2\mu\|\nabla\cdot \boldsymbol{e}_{h,\tau}^n\|_0^2\\ &\le 4\delta\|T_{\mathrm{tr}}^n\|_0^2 + \frac{\delta^2}{2}(1+16\delta)\|T_{\mathrm{tr},\tau}^n\|_0^2 + \frac{\delta^2}{2}(1+8\alpha^2\delta)\|\boldsymbol{e}_{h,\tau}^n + \nabla r_h^n\|_0^2 + \frac{\delta}{16}\|\nabla r_h^n\|_0^2. \end{aligned}$$

The next to last term is bounded by applying the triangle inequality and (4.10)

$$\frac{\delta^2}{2}(1+8\alpha^2\delta)\|\boldsymbol{e}_{h,\tau}^n + \nabla r_h^n\|_0^2 \le \frac{\delta}{16}\|\boldsymbol{e}_{h,\tau}^n\|_0^2 + \frac{\delta}{16}\|\nabla r_h^n\|_0^2.$$

Absorbing the first term into the left-hand side of (5.16) gives

$$(5.17)\quad \begin{aligned} &\frac{5}{8}\delta\|\boldsymbol{e}_{h,\tau}^n\|_0^2 + \nu\delta(\nabla \boldsymbol{e}_{h,\tau}^n, \nabla \boldsymbol{e}_h^n) + \alpha\delta(\boldsymbol{e}_{h,\tau}^n, \boldsymbol{e}_h^n) + \mu\delta(\nabla\cdot \boldsymbol{e}_{h,\tau}^n, \nabla\cdot \boldsymbol{e}_h^n)\\ &\quad + \delta^2((\boldsymbol{e}_{h,\tau}^n + \nabla r_h^n)_\tau, \boldsymbol{e}_{h,\tau}^n + \nabla r_h^n) + \delta^2\nu\|\nabla \boldsymbol{e}_{h,\tau}^n\|_0^2 + \delta^2\mu\|\nabla\cdot \boldsymbol{e}_{h,\tau}^n\|_0^2\\ &\le 4\delta\|T_{\mathrm{tr}}^n\|_0^2 + \frac{\delta^2}{2}(1+16\delta)\|T_{\mathrm{tr},\tau}^n\|_0^2 + \frac{\delta}{8}\|\nabla r_h^n\|_0^2. \end{aligned}$$

Summing up from $j = 2$ to n, observing that, e.g.,

$$(5.18)\quad \sum_{j=2}^n (\boldsymbol{e}_{h,\tau}^j, \boldsymbol{e}_h^j) \ge \frac{1}{2\tau}\sum_{j=2}^n (\|\boldsymbol{e}_h^j\|_0^2 - \|\boldsymbol{e}_h^{j-1}\|_0^2) = \frac{1}{2\tau}(\|\boldsymbol{e}_h^n\|_0^2 - \|\boldsymbol{e}_h^1\|_0^2),$$

applying an appropriate scaling, and restricting on the important term on the left-hand side gives

(5.19)
$$4\delta\tau\sum_{j=2}^{n}\|\boldsymbol{e}_{h,\tau}^{j}\|_0^2 \le C\Bigg(\nu\delta\|\nabla\boldsymbol{e}_h^1\|_0^2+\alpha\delta\|\boldsymbol{e}_h^1\|_0^2+\mu\delta\|\nabla\cdot\boldsymbol{e}_h^1\|_0^2 + \delta^2\|\boldsymbol{e}_{h,\tau}^1+\nabla r_h^1\|_0^2+\delta\sum_{j=2}^{n}\tau\|T_{\mathrm{tr}}^j\|_0^2+\delta^2(1+\delta)\sum_{j=2}^{n}\tau\|T_{\mathrm{tr},\tau}^j\|_0^2\Bigg) + \sum_{j=2}^{n}\frac{4\delta}{5}\tau\|\nabla r_h^j\|_0^2.$$

The last term on the right-hand side of (5.19) will be absorbed into the left-hand side of (5.7).

Estimate the terms for the first time step. The other terms on the right-hand side of (5.19) and the term

$$4\delta\tau\|\boldsymbol{e}_{h,\tau}^1\|_0^2=\frac{4\delta}{\tau}\|\boldsymbol{e}_h^1\|_0^2 \tag{5.20}$$

from (5.7), where $\boldsymbol{e}_h^0=\mathbf{0}$ has been used, have to be bounded.

Considering (5.5) for $n=1$, $\boldsymbol{v}_h=\delta\boldsymbol{e}_h^1$, and $q_h=0$, and applying integration by parts yields

$$\frac{\delta}{\tau}\|\boldsymbol{e}_h^1\|_0^2+\delta\nu\|\nabla\boldsymbol{e}_h^1\|_0^2+\delta\alpha\|\boldsymbol{e}_h^1\|_0^2+\delta\mu\|\nabla\cdot\boldsymbol{e}_h^1\|_0^2=\delta(T_{\mathrm{tr}}^1,\boldsymbol{e}_h^1)-\delta(\boldsymbol{e}_h^1,\nabla r_h^1). \tag{5.21}$$

With the Cauchy–Schwarz inequality and Young's inequality, one gets for the terms from (5.19) and (5.20)

(5.22)
$$\frac{4\delta}{\tau}\|\boldsymbol{e}_h^1\|_0^2+C\Big(\nu\delta\|\nabla\boldsymbol{e}_h^1\|_0^2+\alpha\delta\|\boldsymbol{e}_h^1\|_0^2+\mu\delta\|\nabla\cdot\boldsymbol{e}_h^1\|_0^2\Big)\le C\left(\delta\tau\|T_{\mathrm{tr}}^1\|_0^2+\delta\tau\|\nabla r_h^1\|_0^2\right).$$

Then, from (5.19) and (5.22), apart from the truncation errors, one needs to bound

$$\delta\tau\|\nabla r_h^1\|_0^2,\quad \delta^2\|\boldsymbol{e}_{h,\tau}^1+\nabla r_h^1\|_0^2. \tag{5.23}$$

For $n=1$, and using $\boldsymbol{e}_h^0=\mathbf{0}$ one obtains from (5.6)

$$\begin{aligned}\frac{1}{\tau}\|\boldsymbol{e}_h^1\|_0^2&+\nu\|\nabla\boldsymbol{e}_h^1\|_0^2+\alpha\|\boldsymbol{e}_h^1\|_0^2+\mu\|\nabla\cdot\boldsymbol{e}_h^1\|_0^2+\delta\|\nabla r_h^1\|_0^2\\&=(T_{\mathrm{tr}}^1,\boldsymbol{e}_h^1)+\delta(T_{\mathrm{tr}}^1,\nabla r_h^1)-\frac{\delta}{\tau}(\boldsymbol{e}_h^1,\nabla r_h^1)+\sum_{K\in\mathcal{T}_h}\delta(\nu\Delta\boldsymbol{e}_h^1-\alpha\boldsymbol{e}_h^1,\nabla r_h^1)_K.\end{aligned} \tag{5.24}$$

From (5.5) with $\boldsymbol{v}_h=\mathbf{0}$ and $q^h=\delta r_h^1$, one gets

$$\delta(\nabla\cdot\boldsymbol{e}_h^1,r_h^1)=\delta^2(T_{\mathrm{tr}}^1,\nabla r_h^1)-\frac{\delta^2}{\tau}(\boldsymbol{e}_h^1,\nabla r_h^1)+\sum_{K\in\mathcal{T}_h}\delta^2(\nu\Delta\boldsymbol{e}_h^1-\alpha\boldsymbol{e}_h^1-\nabla r_h^1,\nabla r_h^1)_K. \tag{5.25}$$

Inserting (5.25) into (5.24) gives

$$
\begin{aligned}
&\frac{1}{\tau}\|\boldsymbol{e}_h^1\|_0^2 + \nu\|\nabla \boldsymbol{e}_h^1\|_0^2 + \alpha\|\boldsymbol{e}_h^1\|_0^2 + \mu\|\nabla\cdot \boldsymbol{e}_h^1\|_0^2 + \delta\|\nabla r_h^1\|_0^2 \\
&\quad = (T_{\mathrm{tr}}^1, \boldsymbol{e}_h^1) + \delta(T_{\mathrm{tr}}^1, \nabla r_h^1) + \frac{\delta^2}{\tau}(T_{\mathrm{tr}}^1, \nabla r_h^1) + \sum_{K\in\mathcal{T}_h} \delta(\nu\Delta \boldsymbol{e}_h^1 - \alpha \boldsymbol{e}_h^1, \nabla r_h^1)_K \\
&\qquad + \sum_{K\in\mathcal{T}_h} \frac{\delta^2}{\tau}(\nu\Delta \boldsymbol{e}_h^1 - \alpha \boldsymbol{e}_h^1, \nabla r_h^1)_K - \frac{\delta^2}{\tau^2}(\boldsymbol{e}_h^1 + \tau\nabla r_h^1, \nabla r_h^1).
\end{aligned} \tag{5.26}
$$

The last term can be expanded in the form

$$
-\frac{\delta^2}{\tau}\left(\frac{\boldsymbol{e}_h^1}{\tau} + \nabla r_h^1, \frac{\boldsymbol{e}_h^1}{\tau} + \nabla r_h^1\right) + \frac{\delta^2}{\tau^2}\left(\frac{\boldsymbol{e}_h^1}{\tau} + \nabla r_h^1, \boldsymbol{e}_h^1\right).
$$

The first term goes to the left-hand side of (5.26). Choosing in (5.5) $\boldsymbol{v}_h = \frac{\delta^2}{\tau^2}\boldsymbol{e}_h^1$, $q_h = 0$, and applying integration by parts gives

$$
\frac{\delta^2}{\tau^2}\left(\frac{\boldsymbol{e}_h^1}{\tau} + \nabla r_h^1, \boldsymbol{e}_h^1\right) = -\frac{\delta^2\nu}{\tau^2}\|\nabla \boldsymbol{e}_h^1\|_0^2 - \frac{\delta^2\alpha}{\tau^2}\|\boldsymbol{e}_h^1\|_0^2 - \frac{\delta^2\mu}{\tau^2}\|\nabla\cdot \boldsymbol{e}_h^1\|_0^2 + \frac{\delta^2}{\tau^2}(T_{\mathrm{tr}}^1, \boldsymbol{e}_h^1).
$$

Inserting the last terms into (5.26) and multiplying with τ leads to

$$
\begin{aligned}
&\|\boldsymbol{e}_h^1\|_0^2 + \tau\nu\|\nabla \boldsymbol{e}_h^1\|_0^2 + \tau\alpha\|\boldsymbol{e}_h^1\|_0^2 + \tau\mu\|\nabla\cdot \boldsymbol{e}_h^1\|_0^2 + \tau\delta\|\nabla r_h^1\|_0^2 \\
&\qquad + \delta^2\left\|\frac{\boldsymbol{e}_h^1}{\tau} + \nabla r_h^1\right\|_0^2 + \frac{\delta^2\nu}{\tau}\|\nabla \boldsymbol{e}_h^1\|_0^2 + \frac{\delta^2\alpha}{\tau}\|\boldsymbol{e}_h^1\|_0^2 + \frac{\delta^2\mu}{\tau}\|\nabla\cdot \boldsymbol{e}_h^1\|_0^2 \\
&\quad = \tau(T_{\mathrm{tr}}^1, \boldsymbol{e}_h^1) + \tau\delta(T_{\mathrm{tr}}^1, \nabla r_h^1) + \sum_{K\in\mathcal{T}_h} \tau\delta(\nu\Delta \boldsymbol{e}_h^1 - \alpha \boldsymbol{e}_h^1, \nabla r_h^1) \\
&\qquad + \delta^2 \sum_{K\in\mathcal{T}_h} \left(\nu\Delta \boldsymbol{e}_h^1 - \alpha \boldsymbol{e}_h^1, \nabla r_h^1\right) + \delta^2\left(T_{\mathrm{tr}}^1, \frac{\boldsymbol{e}_h^1}{\tau} + \nabla r_h^1\right).
\end{aligned} \tag{5.27}
$$

The terms on the right-hand side of (5.27) are bounded with the Cauchy–Schwarz inequality, Young's inequality, and (3.2):

$$
\begin{aligned}
\tau(T_{\mathrm{tr}}^1, \boldsymbol{e}_h^1) &\le \frac{\tau}{\alpha}\|T_{\mathrm{tr}}^1\|_0^2 + \frac{\tau\alpha}{4}\|\boldsymbol{e}_h^1\|_0^2, \\
\tau\delta(T_{\mathrm{tr}}^1, \nabla r_h^1) &\le 2\tau\delta\|T_{\mathrm{tr}}^1\|_0^2 + \frac{\tau\delta}{8}\|\nabla r_h^1\|_0^2, \\
\sum_{K\in\mathcal{T}_h} \tau\delta(\nu\Delta \boldsymbol{e}_h^1, \nabla r_h^1) &\le \frac{2\tau\delta C_{\mathrm{inv}}^2\nu^2}{h^2}\|\nabla \boldsymbol{e}_h^1\|_0^2 + \frac{\tau\delta}{8}\|\nabla r_h^1\|_0^2 \le \frac{\tau\nu}{4}\|\nabla \boldsymbol{e}_h^1\|_0^2 + \frac{\tau\delta}{8}\|\nabla r_h^1\|_0^2, \\
\tau\delta(\alpha \boldsymbol{e}_h^1, \nabla r_h^1) &\le 2\tau\delta\alpha^2\|\boldsymbol{e}_h^1\|_0^2 + \frac{\tau\delta}{8}\|\nabla r_h^1\|_0^2 \le \frac{\tau\alpha}{2}\|\boldsymbol{e}_h^1\|_0^2 + \frac{\tau\delta}{8}\|\nabla r_h^1\|_0^2, \\
\sum_{K\in\mathcal{T}_h} \delta^2(\nu\Delta \boldsymbol{e}_h^1, \nabla r_h^1) &\le \frac{2\delta^3 C_{\mathrm{inv}}^2\nu^2}{\tau h^2}\|\nabla \boldsymbol{e}_h^1\|_0^2 + \frac{\tau\delta}{8}\|\nabla r_h^1\|_0^2 \le \frac{\delta^2\nu}{4\tau}\|\nabla \boldsymbol{e}_h^1\|_0^2 + \frac{\tau\delta}{8}\|\nabla r_h^1\|_0^2, \\
\delta^2(\alpha \boldsymbol{e}_h^1, \nabla r_h^1) &\le \frac{2\delta^3\alpha^2}{\tau}\|\boldsymbol{e}_h^1\|_0^2 + \frac{\tau\delta}{8}\|\nabla r_h^1\|_0^2 \le \frac{\delta^2\alpha}{2\tau}\|\boldsymbol{e}_h^1\|_0^2 + \frac{\delta\tau}{8}\|\nabla r_h^1\|_0^2, \\
\delta^2\left(T_{\mathrm{tr}}^1, \frac{\boldsymbol{e}_h^1}{\tau} + \nabla r_h^1\right) &\le \delta^2\|T_{\mathrm{tr}}^1\|_0^2 + \frac{\delta^2}{4}\left\|\frac{\boldsymbol{e}_h^1}{\tau} + \nabla r_h^1\right\|_0^2.
\end{aligned}
$$

Inserting these bounds into (5.27) yields

$$
(5.28)\quad \begin{aligned}
&\|\boldsymbol{e}_h^1\|_0^2 + \frac{3}{4}\tau\nu\|\nabla \boldsymbol{e}_h^1\|_0^2 + \frac{\tau\alpha}{4}\|\boldsymbol{e}_h^1\|_0^2 + \tau\mu\|\nabla\cdot \boldsymbol{e}_h^1\|_0^2 + \frac{3\tau\delta}{8}\|\nabla r_h^1\|_0^2\\
&\quad + \frac{3\delta^2}{4}\left\|\frac{\boldsymbol{e}_h^1}{\tau} + \nabla r_h^1\right\|_0^2 + \frac{3\delta^2\nu}{4\tau}\|\nabla \boldsymbol{e}_h^1\|_0^2 + \frac{\delta^2\alpha}{2\tau}\|\boldsymbol{e}_h^1\|_0^2 + \frac{\delta^2\mu}{\tau}\|\nabla\cdot \boldsymbol{e}_h^1\|_0^2\\
&\le \frac{\tau}{\alpha}\|T_{\rm tr}^1\|_0^2 + (2\tau\delta + \delta^2)\|T_{\rm tr}^1\|_0^2.
\end{aligned}
$$

The right-hand side of (5.28) is bounded following [9, (5.30)]. One obtains

$$
(5.29)\quad \begin{aligned}
\|T_{\rm tr}^1\|_0^2 &\le \frac{C}{\alpha}\Big[h^{2k}(\nu + h^2\alpha + \mu + \delta_0^{-1})\|\partial_t \boldsymbol{u}\|^2_{L^\infty(0,t_1;H^{k+1})}\\
&\qquad + h^{2l}(\delta + h^2\mu^{-1})\|\partial_t p\|^2_{L^\infty(0,t_1;H^{l+1})}\Big] + C\tau\|\partial_{tt}\boldsymbol{s}_h\|^2_{L^2(0,t_1;L^2)}\\
&\le \frac{C}{\alpha}\Big[h^{2k}(\nu + h^2\alpha + \mu + \delta_0^{-1})\|\partial_t \boldsymbol{u}\|^2_{L^\infty(0,t_1;H^{k+1})}\\
&\qquad + h^{2l}(\delta + h^2\mu^{-1})\|\partial_t p\|^2_{L^\infty(0,t_1;H^{l+1})}\Big] + C\tau^2\|\partial_{tt}\boldsymbol{s}_h\|^2_{L^\infty(0,t_1;L^2)}.
\end{aligned}
$$

In this way, the terms in (5.23) are bounded, since both appear on the left-hand side of (5.28).

Estimates of the truncation errors. The truncation errors in (5.7) and (5.19) have still to be bounded. One obtains, again following [9, (5.30)]

$$
(5.30)\quad \begin{aligned}
\sum_{j=1}^n \tau\|T_{\rm tr}^j\|_0^2 &\le 2\sum_{j=1}^n \tau\left(\|\partial_t \boldsymbol{u}(t_j) - \partial_t \boldsymbol{s}_h(t_j)\|_0^2 + \left\|\partial_t \boldsymbol{s}_h^j - \frac{\boldsymbol{s}_h^j - \boldsymbol{s}_h^{j-1}}{\tau}\right\|_0^2\right)\\
&\le \frac{C}{\alpha}\Big(\tau n\Big[h^{2k}(\nu + h^2\alpha + \mu + \delta_0^{-1})\|\partial_t \boldsymbol{u}\|^2_{L^\infty(0,t_n;H^{k+1})}\\
&\qquad + h^{2l}(\delta + h^2\mu^{-1})\|\partial_t p\|^2_{L^\infty(0,t_n;H^{l+1})}\Big]\Big) + C\tau^2\|\partial_{tt}\boldsymbol{s}_h\|^2_{L^2(0,t_n;L^2)},
\end{aligned}
$$

where $\tau n \le T$. Using the definition (5.4) of $T_{\mathrm{tr},\tau}^n$ leads to the decomposition

$$
\begin{aligned}
T_{\mathrm{tr},\tau}^j &= \left(\frac{\partial_t \boldsymbol{u}(t_j) - \partial_t \boldsymbol{u}(t_{j-1})}{\tau} - \partial_{tt}\boldsymbol{u}(t_j)\right) + (\partial_{tt}(\boldsymbol{u} - \boldsymbol{s}_h)(t_j))\\
&\quad + \left(\partial_{tt}\boldsymbol{s}_h(t_j) - \frac{\boldsymbol{s}_h^j - 2\boldsymbol{s}_h^{j-1} + \boldsymbol{s}_h^{j-2}}{\tau^2}\right).
\end{aligned}
$$

The first and second term can be bounded as before and the third one is an $\mathcal{O}(\tau)$ approximation to $\partial_{tt}\boldsymbol{s}_h(t_j)$. Hence, one gets with (4.10)

$$
\begin{aligned}
\sum_{j=2}^n \tau\delta^2\|T_{\mathrm{tr},\tau}^j\|_0^2 &\le \frac{C}{\alpha}\Big(\tau n\Big[h^{2k}(\nu + h^2\alpha + \mu + \delta_0^{-1})\|\partial_{tt}\boldsymbol{u}\|^2_{L^\infty(0,t_n;H^{\max\{1,k-1\}})}\\
&\qquad + h^{2l}(\delta + h^2\mu^{-1})\|\partial_{tt} p\|^2_{L^\infty(0,t_n;H^{\max\{1,l-1\}})}\Big]\Big)\\
&\quad + C\delta^2\tau^2\left(\|\partial_{ttt}\boldsymbol{u}\|^2_{L^2(0,t_n;L^2)} + \|\partial_{ttt}\boldsymbol{s}_h\|^2_{L^2(0,t_n;L^2)}\right).
\end{aligned}
$$

Final step of the proof. To obtain the error estimate, the decomposition (5.2) is used, the triangle inequality is applied, and norms of $\boldsymbol{u}(t_n) - \boldsymbol{s}_h^n$ and $p(t_n) - z_h^n$ are estimated with (3.3). Note that $(\boldsymbol{s}, z) = (\boldsymbol{u}, p)$.

THEOREM 5.1. *Let $(\boldsymbol{u},p)$ be the solution of* (2.2) *with*
- $\boldsymbol{u}, \partial_t \boldsymbol{u} \in L^\infty(0,t_n,H^k(\Omega))$, $\partial_{tt}\boldsymbol{u} \in L^\infty(0,t_n;H^{\max\{1,k-1\}}(\Omega))$, $k \ge 1$, $\partial_{ttt}\boldsymbol{u} \in L^2(0,t_n;L^2(\Omega))$,
- $\partial_{tt}\boldsymbol{s}_h, \partial_{ttt}\boldsymbol{s}_h \in L^2(0,t_n;L^2(\Omega))$, $\partial_{tt}\boldsymbol{s}_h \in L^\infty(0,t_1;L^2(\Omega))$,
- $p, \partial_t p \in L^\infty(0,t_n,H^l(\Omega))$, $\partial_{tt}p \in L^\infty(0,t_n;H^{\max\{1,l-1\}})$, $l \ge 0$.

Let $(\boldsymbol{u}_h^n, p_h^n)$ *be the solution computed with the backward Euler–PSPG method* (5.1)*, let the stabilization parameter satisfy* (3.2) *and* (4.10)*, and let the initial conditions be computed as proposed in Remark* 4.8*, then the following error estimate holds:*

$$
\begin{aligned}
&\|\boldsymbol{u}(t_n)-\boldsymbol{u}_h^n\|_0^2 + \tau\nu\sum_{j=1}^n \|\nabla(\boldsymbol{u}(t_j)-\boldsymbol{u}_h^j)\|_0^2 + \tau\alpha\sum_{j=1}^n \|\boldsymbol{u}(t_j)-\boldsymbol{u}_h^j\|_0^2 \\
&\quad + \tau\mu\sum_{j=1}^n \|\nabla\cdot(\boldsymbol{u}(t_j)-\boldsymbol{u}_h^j)\|_0^2 + \tau\delta\sum_{j=1}^n \|\nabla(p(t_j)-p_h^j)\|_0^2 \\
&\le C\left(h^{2k}+h^{2l}+\tau^2\right),
\end{aligned}
$$

where

$$
\begin{aligned}
C = C\Big(&\delta,\alpha,\alpha^{-1}, \|\boldsymbol{u}\|^2_{L^\infty(0,t_n;H^{k+1})}, \|\partial_t\boldsymbol{u}\|^2_{L^\infty(0,t_n;H^{k+1})}, \|\partial_{tt}\boldsymbol{u}\|^2_{L^\infty(0,t_n;H^{\max\{1,k-1\}})}, \\
&\|\partial_{ttt}\boldsymbol{u}\|^2_{L^2(0,t_n;L^2)}, \|\partial_{tt}\boldsymbol{s}_h\|^2_{L^2(0,t_n;L^2)}, \|\partial_{tt}\boldsymbol{s}_h\|^2_{L^\infty(0,t_1;L^2)}, \|\partial_{ttt}\boldsymbol{s}_h\|^2_{L^2(0,t_n;L^2)}, \\
&\|p\|^2_{L^\infty(0,t_n;H^{l+1})}, \|\partial_t p\|^2_{L^\infty(0,t_n;H^{l+1})}, \|\partial_{tt}p\|^2_{L^\infty(0,t_n;H^{\max\{1,l-1\}})}\Big).
\end{aligned}
$$

Note that the norms with $\boldsymbol{s}_h$ can be estimated with norms of $(\boldsymbol{u},p) = (\boldsymbol{s},z)$ by applying the triangle inequality and using (3.3). At the initial time, there is only an error between $\boldsymbol{u}_0$ and $\boldsymbol{s}_h^0$.

6. Numerical studies. This section first studies two examples for which "instabilities" for small time steps were reported in the literature. It is shown that with the choice of the initial condition proposed in Remark 4.8, these "instabilities" do not arise. It is also discussed that for other initial conditions, the results for small time steps are not unstable in the sense that the error explodes. Finally, an example is presented that supports the analytical results of section 4. We also checked the analytical results of section 5 numerically, but for the sake of brevity, we omit their presentation. All numerical studies were performed with the research code MooNMD [8].

Example 6.1. First, the motivating computational experiment of [1] will be studied. In [1], problem (2.1) with $\Omega = (0,1)^2$, $\nu = 1$, and $\alpha = 0$ was considered with a steady-state solution, whose P_1/P_1 finite element approximation computed with the PSPG method (2.5) with $\mu = 0$ will be denoted by $(\boldsymbol{u}_h, p_h)$. Simulations were performed on uniform grids and different levels of refinement. Starting with the $L^2(\Omega)$ projection of $\boldsymbol{u}_0$ as initial finite element velocity, it was observed that for small time steps, after the first time step, the finite element pressure p_h^1 does not resemble p_h. We repeated this study with different initial conditions. Using the Lagrangian interpolant of $\boldsymbol{u}_0$ as $\boldsymbol{u}_h(0)$, we could observe the same behavior as in [1]. With $\boldsymbol{u}_h(0) = \boldsymbol{s}_h(0)$ (see Remark 4.8), we obtained $(\boldsymbol{u}_h^1, p_h^1) = (\boldsymbol{u}_h, p_h)$ for all considered refinement levels and time steps, with $\Delta t \in [10^{-10}, 10^{-1}]$. Performing simulations in a longer time interval, it could be observed for the Lagrangian interpolant as initial condition and for all time steps and refinement levels a fast convergence of the computed solution to $(\boldsymbol{u}_h, p_h)$. This behavior corresponds to the analytical results derived in section 4.

In [1, p. 581] it is stated "after one time step, the unsteady approximation $(\boldsymbol{u}_h^1, p_h^1)$ should be an $\mathcal{O}(\Delta t)$ perturbation of the steady-state solution $(\boldsymbol{u}_h, p_h)$." We think that

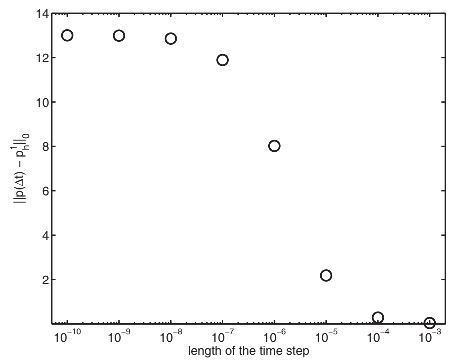

FIG. 1. *Example* 6.1, P_1/P_1, $h = \sqrt{2}/64$, *backward Euler scheme*, $\delta = h^2/4$. *Error of the pressure in* $L^2(\Omega)$ *after the first time step.*

this expectation is not justified if the $L^2(\Omega)$ projection or the Lagrangian interpolant of the steady-state solution is used as initial condition. Neither the first nor the second choice fulfills the discrete equations together with p_h. Thus, the corresponding finite element pressure at the initial time is not known. In the first time step, one expects that the initial velocity does not change much (as it was always observed), but the result of this step will not yet give the steady-state solution. In particular, after the first time step, a finite element pressure p_h^1 is computed such that the approximation of the continuity equation by the PSPG method is satisfied. For small time steps, one expects an approximation of the pressure which accompanies the chosen initial velocity $\boldsymbol{u}_h(0)$. In fact, we could not observe an instability in the sense that $\|p - p_h^1\|_0$ explodes as $\Delta t \to 0$. In our simulations, this error was bounded and the values of the error seem to converge; see Figure 1. This behavior indicates that p_h^1 converges to some (unknown) function as $\Delta t \to 0$.

In [1], the instability of the pressure for small time steps was not observed by solving the equations with the Galerkin discretization using (inf-sup stable) Taylor–Hood finite elements. The analysis for inf-sup stable mixed finite elements can be found, e.g., in [5, 6], where error bounds for mixed finite element approximations to the Navier–Stokes equations were obtained without assuming that nonlocal compatibility conditions are satisfied. In contrast to considering non inf-sup stable elements, the error bounds depend only on the initial approximation of the velocity and not on the initial approximation of the pressure. The analysis is performed by projecting the equations into the space of discretely divergence-free functions, getting an error estimate for the velocity in this space, and then using the discrete inf-sup condition to get the error bound for the pressure. From our point of view, the absence of the error for the pressure at the initial time is the basic difference between both the case of inf-sup stable finite elements and the case in which an appropriate initial approximation to the velocity is chosen (Theorems 4.10 and 4.11) and the estimates of the pressure errors in Theorems 4.5 and 4.7 for a general initial velocity approximation.

Example 6.2. This example is taken from [3] (note that there is a misprint in the definition of the velocity field). The domain is $\Omega = (0,1)^2$ and the prescribed solution has the form

$$\boldsymbol{u} = \cos(t)\begin{pmatrix}\sin(\pi x - 0.7)\sin(\pi y + 0.2)\\ \cos(\pi x - 0.7)\cos(\pi y + 0.2)\end{pmatrix},$$
$$p = \cos(t)(\sin(x)\cos(y) + (\cos(1) - 1)\sin(1)).$$

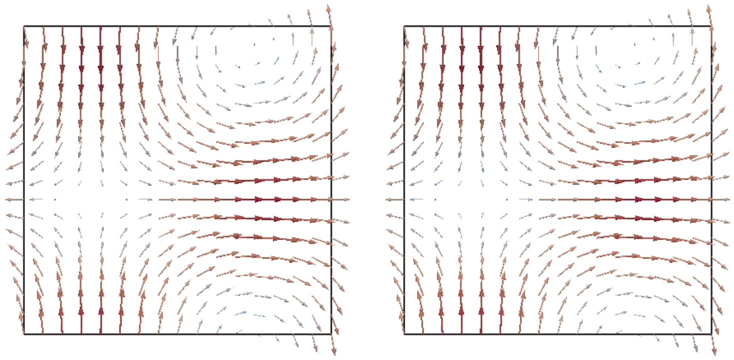

FIG. 2. *Example 6.2. Simulations with $\Delta t = 10^{-8}$ and P_3/P_3 finite elements, $h = \sqrt{2}/16$, velocity after 50 time steps. Left: with Lagrange interpolant as initial velocity; right: with solution of steady-state PSPG problem at $t = 0$ as initial velocity.*

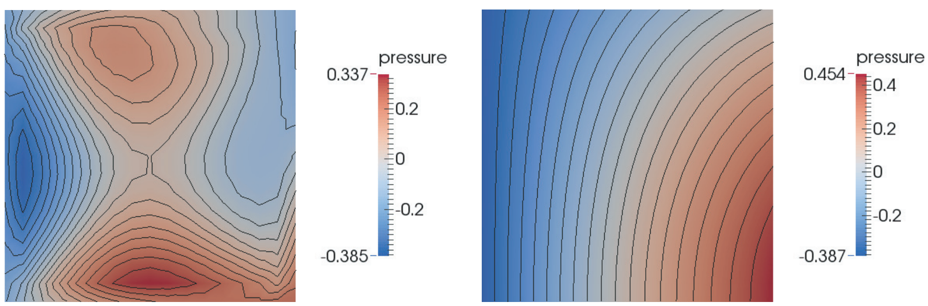

FIG. 3. *Example 6.2. Simulations with $\Delta t = 10^{-8}$ and P_3/P_3 finite elements, $h = \sqrt{2}/16$, pressure after the first time step. Left: with Lagrange interpolant as initial velocity; right: with solution of steady-state PSPG problem at $t = 0$ as initial velocity.*

Appropriate Dirichlet boundary conditions were applied. The viscosity was set to be $\nu = 1$ (there is no value given for ν in [3]) and $\alpha = 0$ was used.

In [3], an instability of the velocity field was observed for very small time steps and the P_3/P_3 finite element method. We tried to reproduce this result. To this end, a uniform triangular mesh with diagonals from lower left to upper right was used with $h = \sqrt{2}/16$. The mesh resulted in 4802 degrees of freedom for the velocity (including Dirichlet nodes) and 2401 degrees of freedom for the pressure. The Crank–Nicolson scheme was applied as temporal discretization and the PSPG method was used with $\delta = 0.005h^2/\nu = 0.005h^2$. The grad-div stabilization was not applied.

Results after having performed 50 steps (as in [3]) with $\Delta t = 10^{-8}$ with the Lagrange interpolation of $\boldsymbol{u}_0$ as initial velocity as well as with the solution $(\boldsymbol{s}_h(0), z_h(0))$ of (2.6) (see Remark 4.8) as initial velocity are presented in Figure 2. In contrast to [3], there are absolutely no instabilities. Also for long term simulations, e.g., with 100 000 time steps, we could not observe instabilities. However, we like to note that for larger stabilization parameters, e.g., $\delta = h^2$, the time stepping scheme diverged quickly in our studies. But according to [3], also in this paper small values of the stabilization parameter were tested.

Next, the effect of different discrete initial velocities on the discrete pressure after the first time step will be illustrated; see Figure 3. Using the Lagrange interpolant, then after a very short time step, the pressure is quite different from the actual solu-

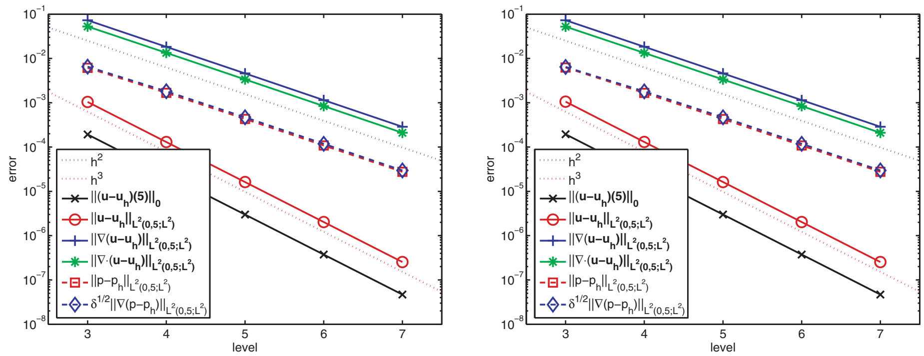

FIG. 4. *Example* 6.3. *Simulations with* P_2/P_2*,* $\Delta t = 5 \cdot 10^{-5}$*,* $\delta = 0.01h^2$*. Convergence of several errors. Left:* $\alpha = 0, \mu = 0$*; right:* $\alpha = 0.2$*,* $\mu = 1$*.*

tion. As explained in Example 6.1, this effect comes from the fact that the Lagrange interpolant is not related to the PSPG discretization of the problem. Using instead $(\boldsymbol{s}_h(0), z_h(0))$ as initial solution, the discrete pressure at $t = 10^{-8}$ is an approximation of the continuous pressure which is as good as the underlying grid admits.

Altogether, the behavior of the discrete solutions observed here is in agreement with the analytical results from section 4.

Example 6.3. Finally, an example will be presented which serves for supporting the analytical results with respect to the order of convergence. Again, the solution from Example 6.2 will be considered. For the sake of brevity, only results for the P_2/P_2 finite element and $\nu = 1$ will be shown. Simulations were performed in the interval $[0, 5]$ and the initial velocity suggested in Remark 4.8 was used. The PSPG stabilization parameter was set to be $\delta = 0.01h^2$. In one series of simulations, $\alpha = \mu = 0$ was used and in a second series $\alpha = 0.2$ and $\mu = 1$. As temporal discretization, the Crank–Nicolson scheme was applied with the small time step $\Delta t = 5 \cdot 10^{-5}$. With this length of the time step, the spatial errors dominate. Level 3 of the mesh refinement has a mesh width of $h = \sqrt{2}/8$ (578 velocity degrees of freedom, 289 pressure degrees of freedom) and all other meshes were obtained with a uniform red refinement.

Results for different errors are presented in Figure 4. Most of the errors are second order convergent, exactly as the theory predicts. The errors that involve the $L^2(\Omega)$ norm of the velocity are even third order convergent. It is well known that this higher order of convergence cannot be proved within the framework of the energy argument used in the analysis of section 4.

7. Summary. The finite element error analysis of the PSPG stabilization for the evolutionary Stokes equations was studied. An optimal error estimate was derived in the continuous-in-time case under the assumption of a regular solution, which holds also for higher order finite elements. An important feature of this estimate is the appearance of the pressure error in $L^2(\Omega)$ at the initial time. The construction of a discrete initial velocity was suggested that allows us to remove this error from the bound. For this discrete initial velocity, an optimal estimate for the fully discrete case with the backward Euler scheme as time integrator was derived. Using the proposed discrete initial velocity, no instabilities of the pressure for small time steps were observed in the numerical simulations. The observations reported in the literature were explained on the basis of the derived error estimates. The analytically predicted results were confirmed in numerical studies.

Acknowledgment. We would like to thank Gabriel Barrenechea (University of Strathclyde) for helpful discussions which were the motivation for proving Theorems 4.10 and 4.11.

REFERENCES

[1] P. B. Bochev, M. D. Gunzburger, and R. B. Lehoucq, *On stabilized finite element methods for the Stokes problem in the small time step limit*, Internat. J. Numer. Methods Fluids, 53 (2007), pp. 573–597.

[2] M. Braack, E. Burman, V. John, and G. Lube, *Stabilized finite element methods for the generalized Oseen problem*, Comput. Methods Appl. Mech. Engrg., 196 (2007), pp. 853–866.

[3] E. Burman and M. A. Fernández, *Analysis of the PSPG method for the transient Stokes' problem*, Comput. Methods Appl. Mech. Engrg., 200 (2011), pp. 2882–2890.

[4] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.

[5] J. G. Heywood and R. Rannacher, *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 (1982), pp. 275–311.

[6] J. G. Heywood and R. Rannacher, *Finite element approximation of the nonstationary Navier-Stokes problem.* III. *Smoothing property and higher order error estimates for spatial discretization*, SIAM J. Numer. Anal., 25 (1988), pp. 489–512.

[7] T. J. R. Hughes, L. P. Franca, and M. Balestra, *A new finite element formulation for computational fluid dynamics.* V. *Circumventing the Babuška-Brezzi condition: A stable Petrov-Galerkin formulation of the Stokes problem accommodating equal-order interpolations*, Comput. Methods Appl. Mech. Engrg., 59 (1986), pp. 85–99.

[8] V. John and G. Matthies, *MooNMD—a program package based on mapped finite element methods*, Comput. Vis. Sci., 6 (2004), pp. 163–169.

[9] V. John and J. Novo, *Error analysis of the SUPG finite element discretization of evolutionary convection-diffusion-reaction equations*, SIAM J. Numer. Anal., 49 (2011), pp. 1149–1176.

[10] V. John and J. Rang, *Adaptive time step control for the incompressible Navier-Stokes equations*, Comput. Methods Appl. Mech. Engrg., 199 (2010), pp. 514–524.

[11] A. Linke, *Collision in a cross-shaped domain—a steady 2d Navier-Stokes example demonstrating the importance of mass conservation in CFD*, Comput. Methods Appl. Mech. Engrg., 198 (2009), pp. 3278–3286.

[12] L. Tobiska and R. Verfürth, *Analysis of a streamline diffusion finite element method for the Stokes and Navier–Stokes equations*, SIAM J. Numer. Anal., 33 (1996), pp. 107–127.