IMA Journal of Numerical Analysis (2018) **00**, 1–47
doi:10.1093/imanum/dry014

A posteriori estimates for the two-step backward differentiation formula and discrete regularity for the time-dependent Stokes equations

Eberhard Bänsch and Andreas Brenner*
Applied Mathematics III, University Erlangen–Nürnberg, Cauerstr. 11, 91058 Erlangen, Germany
*Corresponding author: brenner@math.fau.de

[Received on 15 November 2016; revised on 4 September 2017]

We derive optimal order residual-based *a posteriori* error estimates for fully discrete finite element approximations to the time-dependent Stokes equations. The time discretization uses the two-step backward differentiation formula, and the space discretization is based on inf–sup stable pairs of finite elements, which are allowed to change with time. We show that the time error estimators are of optimal order. This proof of optimality uses time regularity of the semidiscrete (discrete in space) time-dependent Stokes equations. Computational examples are given to confirm the theoretical findings. For completeness, *a priori* estimates are also presented.

Keywords: *a posteriori* error analysis; Stokes equation; time-dependent; reconstruction; BDF2; finite elements.

1. Introduction

Adaptive algorithms have become very popular and powerful tools for finite element approximations of partial differential equations. They are naturally linked to error control, which analysis is provided by *a posteriori* estimates. While for elliptic problems the theory and practice of adaptive methods has reached sound maturity, cf. Verfürth (1996) and Ainsworth & Oden (2011), and the references therein, error control for time-dependent problems is still a challenging area. We mention Makridakis & Nochetto (2003), Verfürth (2003), Akrivis *et al.* (2006), Lakkis & Makridakis (2006), Akrivis *et al.* (2009) and Lozinski *et al.* (2009) for some recent contributions to parabolic problems. For time-dependent problems involving elliptic constraints like the incompressible Navier–Stokes equations, there is even less known.

In this paper we present *a posteriori* estimates for the time-dependent Stokes equations discretized in time by the *two-step backward differentiation formula* (BDF2) and in space by mixed finite elements, which are allowed to change in time. For linear parabolic problems discretized in time by BDF2 (and still continuous in space) we refer to the study by Akrivis & Chatzipantelidis (2010). To the best of our knowledge, however, there is no rigorous *a posteriori* analysis for the fully discrete Stokes equations using BDF2 in time so far. We round out our analysis by proving the optimality of our temporal error estimators (in case of nonchanging meshes). Here, optimality is understood in the sense that the *a posteriori* estimators are asymptotically of the same order of convergence as the corresponding *a priori* error estimates.

Our main tool in deriving *a posteriori* estimates is an application of the *elliptic reconstruction* technique introduced by Makridakis & Nochetto (2003) for parabolic problems. A main advantage of this approach is that (in our case) any available *a posteriori* estimate for the stationary Stokes equations can be used to control the main part of the spatial error.

The reason behind is that in deriving the error estimates an appropriate auxiliary function, called *reconstruction*, is constructed and the error between the exact and numerical solution is split into a part between error and reconstruction as well as an error of approximation type for the *stationary* problem. This approach is in some sense dual to using the Ritz projection (for parabolic problems) in the *a priori* analysis.

More precisely, the reconstruction exhibits two fundamental properties: (i) its difference to the numerical solution can be estimated via known elliptic *a posteriori* results, and (ii) its solution of an error equation (structurally the same as the Stokes equation) with a right-hand side that can be bounded by *a posteriori* estimators in an optimal way.

This technique has been used in many applications; let us mention Akrivis *et al.* (2006), Karakatsani & Makridakis (2007), Akrivis *et al.* (2009), Demlow *et al.* (2009), Georgoulis *et al.* (2011) and Bänsch *et al.* (2012). However, to the best of our knowledge there is no proof of optimality for the estimators for the corresponding *fully discrete* schemes for the Stokes equations. In the semidiscrete case (continuous in space) optimality theory is easily reduced to the regularity of the continuous solution and data, cf. Akrivis & Chatzipantelidis (2010, Chapter 3) for parabolic problems. For the fully discrete case, some additional technical effort has to be invested, for instance the semidiscrete solution (continuous in time, discrete in space) has to be considered as an intermediate step. Higher-order time derivatives of the semidiscrete solution have to be bounded and consequently we provide corresponding estimates, which we call *discrete regularity*, in Appendix B. However, due to technical difficulties we show the optimality result only for a constant mesh.

For the ease of presentation we follow the study by Akrivis & Chatzipantelidis (2010) and restrict our presented analysis to the case of constant time steps. Variable time steps and an adaptive strategy for time step control are treated in Section 7.

An adaptive strategy for-time dependent problems is still a challenging topic. A key problem is the required spatial mesh modifications with time. Their influences in the error estimates are often underestimated. Let us mention the classical article by Dupont (1982) showing that modifying the mesh too carelessly may lead to a wrong solution. In the context of Stokes equations the issue is even more complex. Using standard operators like interpolation or L^2 projection for transferring the velocity from one mesh to the next may lead to devastating effects at least for the pressure, cf. Besier & Wollner (2011) and Brenner *et al.* (2014). From the theoretical point of view a solenoidal projection like Helmholtz or Stokes projection should be used, which is, however, from the computational point of view too expensive. Hence, we use the L^2 projection for transferring the velocity from one mesh to the next and quantify the resulting error by our *changing mesh estimators*.

The rest of this article is organized as follows: in Section 2 we introduce notation, recall some results for the Stokes projection and state corresponding *a priori* estimates. The main results, the *a posteriori* estimates and their optimality, are stated in Section 3. The corresponding proofs are given in Sections 4 and 5. The performance of the proposed error estimators are discussed in Section 6 by computational experiments. An extension for variable time steps will be presented in Section 7. Finally, for completeness *a priori* estimates are proved in Appendix A and the theory of discrete regularity is presented in Appendix B.

2. Notation and preliminaries

Let $\Omega \subset \mathbb{R}^d$ ($d = 2, 3$) be an open, connected and bounded domain with boundary $\Gamma := \partial\Omega$. We consider the incompressible time-dependent Stokes equations on a finite time interval $]0, T[$: Find the velocity u and the pressure p fulfilling

$$\partial_t u - \nu \Delta u + \nabla p = f, \qquad \operatorname{div} u = 0 \text{ in } \Omega \times]0, T[, \qquad u_{|\Gamma} = 0, \qquad u(t=0) = u_0 \text{ in } \Omega. \tag{2.1}$$

For the sake of simplicity we set $\nu = 1$; u_0 is the initial, sufficiently smooth velocity. In the following we assume that the unique solution (u, p) and the right-hand side f to the above system is sufficiently smooth in time and space (compare also the discussion in Appendix B regarding regularity for the continuous system).

As usual the standard Sobolev spaces $H^m(\Omega)$ $(m = 0, \pm 1, \dots)$ are used, whose norms are denoted by $\|\cdot\|_m$. The norm and inner product of $L^2(\Omega) = H^0(\Omega)$ is denoted by $\|\cdot\|$ and $\langle \cdot, \cdot \rangle$, respectively. Moreover, define $L_0^2(\Omega) := \{q \in L^2(\Omega) \mid \int_\Omega q(x)\,\mathrm{d}x = 0\}$. To account for the homogeneous Dirichlet boundary condition we introduce $H_0^1(\Omega) := \{v \in H^1(\Omega) \,:\, v_{|\Gamma} = 0\}$. Thanks to the Poincaré inequality, for $v \in H_0^1(\Omega)^d$, the norms $\|\nabla v\|$ and $\|v\|_1$ are equivalent, and thus in what follows $\|\nabla v\|$ is taken as norm in $H_0^1(\Omega)^d$.

Set $X := H_0^1(\Omega)^d$ and $Y := L_0^2(\Omega)$, satisfying the continuous inf–sup condition:

$$\exists\, c > 0, \quad \inf_{q \in Y} \sup_{v \in X} \frac{\langle q, \operatorname{div} v \rangle}{\|q\| \|\nabla v\|} \ \geq\ c.$$

Furthermore, define the spaces of solenoidal functions in X and $L^2(\Omega)^d$, respectively:

$$\begin{aligned} J &:= \left\{ v \in X \ :\ \operatorname{div} v = 0 \right\}, \\ H &:= \left\{ v \in L^2(\Omega)^d \ : \operatorname{div} v = 0,\ v \cdot n_{|\Gamma} = 0 \right\}. \end{aligned}$$

2.1. *Finite element spaces*

To discretize in time let $0 = t_0 < t_1 < \cdots < t_N = T$ be a partition of $[0, T]$, $I_n := (t_{n-1}, t_n]$ a time sub-interval and $\tau_n := t_n - t_{n-1}$ the time step. In the following we assume constant time steps. An extension for variable time steps will be presented in Section 7.

For the space discretization we use a standard finite element method and introduce a family $\{\mathcal{T}_n\}_{n=0}^N$ of conforming shape regular triangulations of the domain Ω. Each triangulation $\mathcal{T}_n$ corresponds to a time interval I_n. This allows arbitrary mesh changes. Let $h_n(x)$ be the *mesh-size function* for each given triangulation $\mathcal{T}_n$ defined by

$$h_n(x) := \operatorname{diam}(K), \quad K \in \mathcal{T}_n,\ x \in K$$

and denote by $h_n := \max_{x \in \Omega_n} h_n(x)$. For each triangulation $\mathcal{T}_n$ we choose a pair of finite element spaces $X_n := X_{h_n} \subseteq X$ and $Y_n := Y_{h_n} \subseteq Y$ of X and Y.

Assume for all n that $X_n \times Y_n$ is of order $k \in \mathbb{N}$, i.e. for each $v \in X \cap H^{k+1}(\Omega)^d$ and $q \in Y \cap H^k(\Omega)$ there exists approximations $\mathcal{I}_n v \in X_n$ and $\mathcal{J}_n q \in Y_n$ such that

$$\|v - \mathcal{I}_n v\| + h_n \|\nabla(v - \mathcal{I}_n v)\| \leq C h_n^{k+1} \|v\|_{k+1}, \qquad \|q - \mathcal{J}_n q\|_{L_0^2(\Omega)} \leq C h_n^k \|q\|_k. \tag{2.2}$$

These are standard conditions satisfied for instance by Taylor–Hood elements $\mathcal{P}_k - \mathcal{P}_{k-1}$, $k \geq 2$.

To handle mesh changes let us denote by $\varPi_n := \varPi_{X_n} : X \to X_n$ the L^2 projection onto X_n.

We introduce the linear operator $-\varDelta_n : X \to X_n$, such that for all $u \in X$, $v_n \in X_n$, $-\varDelta_n u \in X_n$ with $\langle -\varDelta_n u, v_n \rangle = \langle \nabla u, \nabla v_n \rangle$, and the linear operator $B_n : X_n \to Y_n$ and its transpose $B_n^\top : Y_n \to X_n$, so that for every couple $(v_n, q_n) \in X_n \times Y_n$ the identities $\langle B_n v_n, q_n \rangle = -\langle \mathrm{div}\, v_n, q_n \rangle$ and $\langle v_n, B_n^\top q_n \rangle = -\langle \mathrm{div}\, v_n, q_n \rangle$ hold. In the sequel we assume that the mixed approximation satisfies the discrete inf–sup condition

$$\exists\, c > 0, \quad \inf_{q_n \in Y_n} \sup_{v_n \in X_n} \frac{\langle B_n v_n, q_n \rangle}{\|\nabla v_n\| \, \|q_n\|} \geq c \tag{2.3}$$

with a constant c not depending on n.

Remark 2.1 Our analysis is also valid for nonconforming finite elements spaces. In this case the definitions of the discrete operators has to be extended to

$$-\triangle_n : X \to X_n + X, \quad B_n : X_n + X \to Y_n, \quad B_n^\top : Y_n \to X_n + X, \quad \varPi_n : X \to X_n + X,$$

and we must require the equivalence of the norms $\|\nabla u\|$ and $\|u\|_1$ for $u \in X_n$. This is fulfilled for instance for the Crouzeix–Raviart element of lowest order, cf. Crouzeix & Raviart (1973).

The space of discretely divergence-free vector fields is denoted by

$$J_n := \{v_n \in X_n \; : \; \langle \mathrm{div}\, v_n, q_n \rangle = 0 \;\; \forall\, q_n \in Y_n\} = \ker B_n$$

and for $u \in H^{k+1}(\varOmega)^d \cap X$ and $p \in H^k(\varOmega) \cap Y$ the discrete *Stokes projection* $\mathcal{S}_n^u$ is given by $\mathcal{S}_n(u,p) = (\mathcal{S}_n^u(u,p), \mathcal{S}_n^p(u,p)) \in X_n \times Y_n$, solution of the problem

$$\begin{aligned} \left\langle \nabla \mathcal{S}_n^u(u,p), \nabla v_n \right\rangle - \left\langle \mathcal{S}_n^p(u,p), \mathrm{div}\, v_n \right\rangle &= \langle \nabla u, \nabla v_n \rangle + \langle \nabla p, v_n \rangle && \forall\, v_n \in X_n, \\ \left\langle \mathrm{div}\, \mathcal{S}_n^u(u,p), q_n \right\rangle &= 0 && \forall\, q_n \in Y_n. \end{aligned} \tag{2.4}$$

Note that in the case of a discrete pair $(V_n, Q_n) \in X_n \times Y_n$ one has $(\mathcal{S}_n^u(V_n,Q_n), \mathcal{S}_n^p(V_n,Q_n)) = (V_n, Q_n)$.

In what follows C will denote a generic constant (possibly changing values at different places) that is independent of h and τ, but may depend on the solution and the domain. Furthermore, we write $a \lesssim b$ for two functions or quantities a, b, whenever there is a generic constant C, such that $a \leq C\, b$.

In the study by Heywood & Rannacher (1988) the following estimates are proved:

Lemma 2.2 Let $m_0 \in \mathbb{N}$ and $\partial_t^{(m)} u \in H^{k+1}(\varOmega)^d \cap X$ and $\partial_t^{(m)} p \in H^k(\varOmega) \cap Y$ for $0 \leq m \leq m_0$. Then

$$\begin{gathered} \left\| u - \mathcal{S}_n^u(u,p) \right\| + h_n \left\| \nabla (u - \mathcal{S}_n^u(u,p)) \right\| \lesssim h_n^{k+1} \left(\|u\|_{k+1} + \|p\|_k \right), \\ \left\| \partial_t^{(m)} \left(u - \mathcal{S}_n^u(u,p) \right) \right\| + h_n \left\| \partial_t^{(m)} \nabla \left(u - \mathcal{S}_n^u(u,p) \right) \right\| \lesssim h_n^{k+1} \left(\left\| \partial_t^{(m)} u \right\|_{k+1} + \left\| \partial_t^{(m)} p \right\|_k \right). \end{gathered}$$

Additionally we need an estimate for the associated pressure.

LEMMA 2.3 Let $u \in H^{k+1}(\Omega)^d \cap X$ and $p \in H^k(\Omega) \cap Y$. Then the following pressure estimate holds:

$$\left\|\mathcal{S}_n^p(u,p) - p\right\| \leq h_n^k\,(\|u\|_{k+1} + \|p\|_k).$$

Proof. Use the inf–sup stability (2.3) and the interpolation operator $\mathcal{J}_h$ to obtain

$$\begin{aligned}
\left\|\mathcal{S}_n^p(u,p) - p\right\| &\leq \left\|\mathcal{S}_n^p(u,p) - \mathcal{J}_n p\right\| + \|(I-\mathcal{J}_n)p\| \\
&\lesssim \sup_{\substack{v_n \in X_n \\ \|\nabla v_n\|=1}} \left\langle \mathcal{S}_n^p(u,p) - \mathcal{J}_n p, \operatorname{div} v_n\right\rangle + \|(I-\mathcal{J}_n)p\| \\
&= \sup_{\substack{v_n \in X_n \\ \|\nabla v_n\|=1}} \left\langle \mathcal{S}_n^p(u,p) - p, \operatorname{div} v_n\right\rangle + \sup_{\substack{v_n \in X_n \\ \|\nabla v_n\|=1}} \langle (I-\mathcal{J}_n)p, \operatorname{div} v_n\rangle + \|(I-\mathcal{J}_n)p\| \\
&\lesssim \left\|\nabla\left(u - \mathcal{S}_n^u(u,p)\right)\right\| + \|(I-\mathcal{J}_n)p\|,
\end{aligned}$$

where in the last step equation (2.4) was used. Then the lemma is concluded by equation (2.2) and Lemma 2.2. □

Let us define the *Stokes operator*

$$\mathcal{A}u := -\Pi_H \triangle u \quad \forall\, u \in D(\mathcal{A}) = J \cap H^2(\Omega)^d, \tag{2.5}$$

where $\Pi_H : L^2(\Omega)^d \to H$ is the Helmholtz projection, see for instance Sohr (2001, III.2). For a given $g \in L^2(\Omega)^d$, $w = \mathcal{A}^{-1}g$ is solution of the Stokes system

$$-\Delta w + \nabla r = g,\ \operatorname{div} w = 0 \quad \text{in } \Omega, \qquad w = 0 \quad \text{on } \partial\Omega. \tag{2.6}$$

In what follows, we assume that the Stokes problem is $H^2 \times H^1$-regular. More precisely:

ASSUMPTION 2.4 Let (w, r) be the solution of equation (2.6). Then $(w, r) \in H^2(\Omega)^d \cap X \times H^1(\Omega) \cap Y$ and

$$\|w\|_2 + \|\nabla r\| \lesssim \|g\|. \tag{2.7}$$

2.2 *Time discretization*

For the time discretization we introduce the finite difference operators $\bar{\partial}^k$. For a sequence $(v^n)_{n=0}^N$ with $v^n \in X_n$ we set recursively

$$\bar{\partial} v^n := \frac{v^n - v^{n-1}}{\tau}, \qquad \bar{\partial}^2 v^n := \bar{\partial}(\bar{\partial} v^n), \qquad \bar{\partial}^k v^n = \bar{\partial}(\bar{\partial}^{k-1} v^n),$$

and the BDF2 time discretization operator by

$$\bar{\partial}^B v^n := \frac{\tau}{2}\bar{\partial}^2 v^n + \bar{\partial} v^n = \frac{3v^n - 4v^{n-1} + v^{n-2}}{2\tau}. \tag{2.8}$$

An extension of $\bar{\partial}^B$ to arbitrary time steps will be presented in Section 7. Note that in general $\bar{\partial} v^n \notin X_n$. Hence, define

$$\bar{\partial}_n^k v^n := \Pi_n \bar{\partial}^k v^n \ \in \ X_n, \qquad \bar{\partial}_n^B v^n := \Pi_n \bar{\partial}^B v^n \ \in \ X_n.$$

For any Banach space Z, $L^p((t_1,t_2);Z)$ is the space of Lebesgue measurable Z-valued functions $u : (t_1, t_2) \to Z$ with

$$\|u\|_{L^p((t_1,t_2);Z)} := \left(\int_{t_1}^{t_2} \|u(t)\|_Z^p \, \mathrm{d}t \right)^{1/p}, \qquad \|u\|_{L^\infty((t_1,t_2);Z)} := \operatorname{ess} \sup_{t_1 < t < t_2} \|u(t)\|_Z,$$

$1 \le p < \infty$, and $\|\cdot\|_{L^p(Z)} := \|\cdot\|_{L^p((0,T);Z)}$. Furthermore, for a sequence $v^{i_1}, \ldots, v^{i_m} \in Z$, $i_1, \ldots, i_m \in \mathbb{N}$, $i_1 \le \cdots \le i_m$, let

$$\|u\|_{\ell^p(i_1,i_m;Z)} := \left(\sum_{n=i_1}^{i_m} \tau \|u^n\|_Z^p \right)^{1/p}, \qquad \|u\|_{\ell^\infty(i_1,i_m;Z)} := \max_{i_1 \le n \le i_m} \|u^n\|_Z,$$

$1 \le p < \infty$, and set $\|u\|_{\ell^p(Z)} := \|u\|_{\ell^p(1,N;Z)}$.

Throughout this paper we denote space discrete quantities by capitals and continuous quantities by small letters.

2.3 *The scheme: BDF2 for the fully discrete Stokes equation*

In order to discretize equation (2.1) define a sequence of approximate velocities $U^n \in X_n$ and pressures $P^n \in Y_n$ fulfilling: let $X_0 = X_1 = X_2$ and $U^0 \in X_0$, $U^1 \in X_1$ be given; for $n \ge 2$ determine $\left(U^n, P^n\right) \in X_n \times Y_n$ from

$$\boxed{\bar{\partial}_n^B U^n - \Delta_n U^n + B_n^\top P^n = F^n, \qquad B_n U^n = 0,} \tag{2.9}$$

where $F^n := \Pi_n f(t_n)$.

To initialize the scheme, values for U^0, P^0 and U^1 are needed. To this end determine $p_0 := p(0)$ from u_0 by taking the divergence of the first equation of equation (2.1) for $t = 0$ resulting in a Poisson problem for p_0 with, say, Neumann boundary values obtained by testing the first equation of equation (2.1) (evaluated on the boundary) with the normal vector. Set $(U^0, P^0) := \mathcal{S}_0(u_0, p_0) = (\mathcal{S}_0^u(u_0, p_0), \mathcal{S}_0^p(u_0, p_0))$ and calculate $(U^1, P^1) \in X_1 \times Y_1$ by the trapezoidal scheme:

$$\bar{\partial}_1 U^1 - \Delta_1 U^{1/2} + B_1^\top P^{1/2} = F^{1/2}, \qquad B_1 U^1 = 0, \tag{2.10}$$

where $U^{1/2} := \frac{1}{2}(U^1 + U^0)$.

The above discretization yields the following *a priori* estimates for constant meshes:

THEOREM 2.5 (*A priori* estimates for the time-dependent Stokes equations) Suppose the exact solution (u, p) of equation (2.1) is sufficiently regular. Let $X_n \equiv X_h \subseteq X$ for all $n \ge 0$ with mesh size h and

(U^n, P^n) be the solution of equation (2.9) initialized by (2.10). Then

$$\|U-u\|^2_{\ell^\infty(L^2)} \lesssim h^{2(k+1)}(A^0+A^1) \;+\; \tau^4 \left\|\partial_t^3 u\right\|^2_{L^1(H^{-1})}, \tag{2.11}$$

$$\|\nabla(U-u)\|^2_{\ell^2(L^2)} \lesssim h^{2k}(A^0+h^2A^1) \;+\; \tau^4 \left\|\partial_t^3 u\right\|^2_{L^1(H^{-1})}, \tag{2.12}$$

$$\|\nabla(U-u)\|^2_{\ell^\infty(L^2)} + \|P-p\|^2_{\ell^2(L^2)} \lesssim h^{2k}(A^0+h^2A^1) \;+\; \tau^4 \left\|\partial_t^3 u\right\|^2_{L^1(L^2)}, \tag{2.13}$$

where $A^i := \|\partial_t^i u\|^2_{L^\infty(H^{k+1})} + \|\partial_t^i p\|^2_{L^\infty(H^k)}$, $i = 0, 1$.

The proof is outlined in Appendix A.

3. *A posteriori* estimates for the fully discrete scheme

In this section we introduce *a posteriori* error estimators and state corresponding *a posteriori* estimates in Theorem 3.4. The proof of the theorem is outlined in Section 4. Furthermore, in Theorem 3.6 we state that our time estimators are optimal, i.e. that they have the same asymptotic behavior as the *a priori* estimates. The theorem is proven in Section 5 and utilizes discrete regularity estimates, which are provided in Appendix B.

In order to derive optimal *a posteriori* estimates we need an approximation $U(t)$ to $u(t)$ $\forall\ t \in\ ]0, T[$ to the right order. To this end we introduce a piecewise linear function, interpolating a sequence $\{U^n\}_{n=0,\dots,N}$ at the time nodes, by

$$U(t) := U^n + (t-t_n)\bar{\partial} U^n, \quad t \in I_n,\ n = 1, \dots, N. \tag{3.1}$$

Even though the interpolant $U(t)$, defined in equation (3.1), is a second-order approximation, its residual is of first order only, cf. the study by Akrivis & Chatzipantelidis (2010) for parabolic problems. We follow the study by Akrivis & Chatzipantelidis (2010) and modify the interpolant $U(t)$ using the piecewise quadratic *three-point reconstruction* $\hat{U}(t)$, interpolating a sequence $\{U^n\}_{n=0,\dots,N}$ in t_{n-2}, t_{n-1}, t_n:

$$\hat{U}(t) := U(t) + \frac{1}{2}(t-t_n)(t-t_{n-1})\bar{\partial}^2 U^n, \quad t \in I_n,\ n = 2, \dots, N. \tag{3.2}$$

Its time derivative is given by

$$\partial_t \hat{U}(t) = \bar{\partial}^B U^n + (t-t_n)\bar{\partial}^2 U^n, \quad t \in I_n,\ n = 2, \dots, N. \tag{3.3}$$

This three-point reconstruction was also proposed in the study by Lozinski *et al.* (2009) for the Crank–Nicolson scheme.

Our main tool in deriving estimates is an application of the *elliptic reconstruction* technique introduced by Makridakis & Nochetto (2003) for parabolic problems and extended to the Stokes equations in the study by Karakatsani & Makridakis (2007).

To this end define the *Stokes Reconstruction* $\mathcal{R}_u^n$ of a discrete function $(U^n, P^n) \in X_n \times Y_n$ by

$$\mathcal{R}_u^n := \mathcal{A}^{-1}\left(-\Delta_n U^n + B_n^\top P^n - F^n + f(t_n)\right),$$

i.e. there is a unique $\mathcal{R}_p^n \in Y$ such that $(\mathcal{R}_u^n, \mathcal{R}_p^n) \in X \times Y$ is the *continuous* solution of

$$-\Delta\mathcal{R}_u^n + \nabla\mathcal{R}_p^n = -\Delta_n U^n + B_n^\top P^n - F^n + f(t_n), \qquad \operatorname{div}\mathcal{R}_u^n = 0, \qquad \mathcal{R}_u^n|_\Gamma = 0. \tag{3.4}$$

The next lemma, which is straightforward, though crucial for our analysis, shows that the finite element approximation of $(\mathcal{R}_u^n, \mathcal{R}_p^n)$ in equation (3.4) is nothing else but (U^n, P^n) itself.

LEMMA 3.1 Let $(V^n, R^n) \in X_n \times Y_n$ be the finite element solution of equation (3.4), namely

$$-\Delta_n V^n + B_n^\top R^n = -\Delta\mathcal{R}_u^n + \nabla\mathcal{R}_p^n = -\Delta_n U^n + B_n^\top P^n - F^n + f(t_n), \qquad B_n V^n = 0.$$

Then,

$$V^n = U^n \qquad \text{and} \qquad R^n = P^n.$$

Proof. See Karakatsani & Makridakis (2007, Lemma 1.3). □

The above lemma implies that the difference of a given finite element function and its Stokes reconstruction can be estimated by well-known estimates for the stationary Stokes system. We thus make the assumption on the availability of such *a posteriori* error estimates.

ASSUMPTION 3.2 Let $(\mathcal{R}_u^n, \mathcal{R}_p^n) \in X \times Y$ and $(U^n, P^n) \in X_n \times Y_n$ be as in equation (3.4). We assume that there exist *a posteriori* estimator functions $\eta_u^n = \eta_u^n(U^n; Z)$, $\eta_p^n = \eta_p^n(P^n; Y)$, such that

$$\begin{aligned} \left\|U^n - \mathcal{R}_u^n\right\|_Z^2 &\le \eta_u^n(U^n; Z), \qquad Z \in \left\{J^*, L^2, H_0^1\right\}, \\ \left\|P^n - \mathcal{R}_p^n\right\|^2 &\le \eta_p^n(P^n; Y). \end{aligned} \tag{3.5 }$$

Additionally there exists an estimator function $\eta^0(Z)$ of the initial approximation error, such that

$$\left\|u_0 - \mathcal{S}_0^u(u_0, p_0)\right\|_Z^2 \le \eta^0(Z).$$

Note that for a large class of finite element spaces X_n, Y_n such estimators are available, see for example Verführt (2013). In the study by Bank & Welfert (1990) a comparison of several *a posteriori* error estimates for the stationary Stokes problem is given.

Let us now summarize the notation. First we denote space discrete quantities by capitals and continuous quantities by small letters. $U^n \in X_n$ denotes the solution of equation (2.9) at discrete time t_n, whereas $U(t)$ and $\hat{U}(t)$ are interpolants of U^n and U^{n-1} in the corresponding time interval I_n. Equally we apply this notation for the Stokes reconstruction $\mathcal{R}_u^n$, which is defined in equation (3.4) at discrete time t_n and $\mathcal{R}_u(t)$ or $\hat{\mathcal{R}}_u(t)$ are the interpolation, respectively three-point reconstruction, in the time interval I_n.

Now we are able to define the error estimators.

DEFINITION 3.3 (Error estimators). We assume $X_2 = X_1 = X_0$ and set to simplify the presentation $U^{-1} := U^1 - 2\tau_0(\Delta_0 U^0 - B_0^\top P^0 + F^0) \in X_0$ and $U^{-2} := -6\tau_0(\Delta_0 U^0 - B_0 P^0 + F^0) - 3U^0 + 4U^1 \in X_0$, cf. Lemma 4.5.

Let us introduce the **time error estimators**

$$\mathcal{E}_1 := \sum_{n=2}^{N} \tau \left\| \bar{\partial}_n^2 U^n \right\|^2, \qquad \mathcal{E}_2 := \sum_{n=2}^{N} \frac{1}{\tau} \left\| \bar{\partial}_n^2 U^n - \bar{\partial}_{n-1}^2 U^{n-1} \right\|_{J^*}^2, \quad \mathcal{E}_3 := \max_{2\le n\le N} \left\| \bar{\partial}_n^2 U^n \right\|^2,$$

$$\mathcal{E}_6 := \sum_{n=2}^{N} \frac{1}{\tau} \left\| \bar{\partial}_n^2 U^n - \bar{\partial}_{n-1}^2 U^{n-1} \right\|^2, \quad \mathcal{E}_7 := \max_{2\le n\le N} \left\| \nabla \bar{\partial}_n^2 U^n \right\|^2,$$

the **Stokes reconstruction error estimators**

$$\mathcal{E}_4 := \sum_{n=2}^{N} \tau \left\| \bar{\partial}^2 f(t_n) - \bar{\partial}_n^2 \bar{\partial}^B U^n \right\|_{J^*}^2, \qquad \mathcal{E}_5 := \sum_{n=2}^{N} \tau \left\| \bar{\partial}^2 f(t_n) - \bar{\partial}_n^2 \bar{\partial}^B U^n \right\|^2,$$

the **space error estimators**

$$\Lambda_1 := \sum_{n=0}^{N} \tau\ \eta_u^n(U^n; L^2), \qquad \Lambda_2 := \max_{0\le n\le N} \eta_u^n(U^n; L^2),$$

$$\Lambda_3 := \sum_{n=1}^{N} \tau\ \eta_u^n\left(\bar{\partial}_n U^n; J^*\right), \qquad \Lambda_4 := \sum_{n=1}^{N} \tau\ \eta_u^n\left(U^n; H_0^1\right),$$

$$\Lambda_5 := \max_{0\le n\le N} \eta_u^n\left(U^n; H_0^1\right), \qquad \Lambda_6 := \sum_{n=1}^{N} \tau\ \eta_u^n\left(\bar{\partial}_n U^n; L^2\right),$$

$$\Lambda_7 := \sum_{n=0}^{N} \tau\ \eta_p^n(P^n; Y),$$

the **changing mesh estimators**

$$\Xi_1 := \sum_{n=2}^{N} \tau \xi_B^{n,n}, \qquad \Xi_4 := \sum_{n=2}^{N} \frac{h_n^2}{\tau} \left(\xi_B^{n,n} + \xi_B^{n-1,n} + \xi_B^{n-2,n}\right), \quad \Xi_6 := \sum_{n=2}^{N} \frac{1}{\tau} \xi_B^{n,n},$$

$$\Xi_2 := \sum_{n=2}^{N} \frac{h_n^2}{\tau} \xi_B^{n,n}, \quad \Xi_5 := \sum_{n=2}^{N} \frac{1}{\tau} \left(\xi_B^{n,n} + \xi_B^{n-1,n} + \xi_B^{n-2,n}\right), \quad \Xi_7 := \tau^4 \max_{2\le n\le N} \left\| \nabla \left(\bar{\partial}^2 - \bar{\partial}_n^2\right) U^n \right\|^2,$$

$$\Xi_3 := \max_{2\le n\le N} \xi_B^{n,n},$$

where the local BDF2-changing mesh operator $\xi_B^{a,b}$ is defined by

$$\xi_B^{a,b} := \left(\|U^a\|^2 - \|\Pi_b U^a\|^2\right) + \left(\|U^{a-1}\|^2 - \|\Pi_b U^{a-1}\|^2\right) + \left(\|U^{a-2}\|^2 - \|\Pi_b U^{a-2}\|^2\right), \tag{3.6}$$

and the **data approximation error estimators**

$$\mathcal{F}_1 := \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \left\| f(t_n) + (t - t_n)\bar{\partial} f(t_n) - f(t) \right\|_{J^*}^2 \mathrm{d}t, \quad \mathcal{F}_2 := \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \left\| f(t_n) + (t - t_n)\bar{\partial} f(t_n) - f(t) \right\|^2,$$

$$\mathcal{I}_1 := \tau^3 \|\Psi_1\|_{J^*}^2, \qquad \mathcal{I}_2 := \tau^3 \|\Psi_1\|^2,$$

where $\Psi_1 := \bar{\partial}^2 U^2 - \Delta_1 \bar{\partial} U^1 + B_1^\top \bar{\partial} P^1 - \bar{\partial} F^1$.

In the following theorem *a posteriori* estimates in various norms are stated.

THEOREM 3.4 (*A posteriori* error estimates for the time-dependent Stokes system). Let (U,P) be the solution of equation (2.9) initialized by equation (2.10) and (u,p) the exact solution of equation (2.1). Then

$$\|U - u\|_{L^2(L^2)}^2 \lesssim \eta^0(J^*) + \tau^4(\mathcal{E}_1 + \mathcal{E}_2) + \Lambda_1 + \Xi_1 + \Xi_2 + \mathcal{F}_1 + \mathcal{I}_1, \tag{3.7}$$

$$\|U - u\|_{L^\infty(L^2)}^2 \lesssim \eta^0(L^2) + \tau^4(\mathcal{E}_2 + \mathcal{E}_3 + \mathcal{E}_4) + \Lambda_2 + \Lambda_3 + \Xi_2 + \Xi_3 + \Xi_4 + \mathcal{F}_1 + \mathcal{I}_1, \tag{3.8}$$

$$\|\nabla(U - u)\|_{L^2(L^2)}^2 \lesssim \eta^0(L^2) + \tau^4(\mathcal{E}_2 + \mathcal{E}_4) + \Lambda_3 + \Lambda_4 + \Xi_2 + \Xi_4 + \mathcal{F}_1 + \mathcal{I}_1, \tag{3.9}$$

$$\|\nabla(U - u)(t)\|_{L^\infty(L^2)}^2 \lesssim \eta^0\left(H_0^1\right) + \tau^4(\mathcal{E}_5 + \mathcal{E}_6 + \mathcal{E}_7) + \Lambda_5 + \Lambda_6 + \Xi_5 + \Xi_6 + \Xi_7 + \mathcal{F}_2 + \mathcal{I}_2, \tag{3.10}$$

$$\|P - p\|_{L^2(L^2)}^2 \lesssim \eta^0\left(H_0^1\right) + \tau^4(\mathcal{E}_5 + \mathcal{E}_6) + \Lambda_6 + \Lambda_7 + \Xi_5 + \Xi_6 + \mathcal{F}_2 + \mathcal{I}_2, \tag{3.11}$$

where the estimators are defined in Definition 3.3 and $\eta^0(Z)$ fulfills Assumption 3.2.

REMARK 3.5 (Negative norms). The time error estimator $\mathcal{E}_2$, Stokes reconstruction estimator $\mathcal{E}_4$ and data approximation error estimator $\mathcal{F}_1$ are defined with respect to a negative norm, as usual for error estimates of BDF2 time discretizations, cf. Akrivis & Chatzipantelidis (2010). From a computational point of view this is completely prohibitive, since its evaluation may be computationally more expensive than solving the discrete scheme. However, there are possibilities for estimating it. The simplest way is to bound $\|\cdot\|_{J^*}$ by $\|\cdot\|$, which can be pessimistic in practice. Another way is to approximate the J^* norm in an adequate way. In our case the estimator $\mathcal{E}_4$ (for the sake of convenience assume constant meshes) can be handled further in the following way. We approximate the norm J^* by the discrete norm J_h^*:

$$\left\| \bar{\partial}^2 f(t_n) - \bar{\partial}^2 \bar{\partial}^B U^n \right\|_{J^*} \approx \left\| \bar{\partial}^2 f(t_n) - \bar{\partial}^2 \bar{\partial}^B U^n \right\|_{J_h^*}.$$

Now we use the triangle inequality and insert the scheme to get

$$\begin{aligned}\left\|\bar{\partial}^2 f(t_n) - \bar{\partial}^2\bar{\partial}^B U^n\right\|_{J_h^*} &\leq \left\|\bar{\partial}^2\left(f(t_n) - F^n\right)\right\|_{J_h^*} + \left\|\bar{\partial}^2(F^n - \bar{\partial}^B U^n)\right\|_{J_h^*}\\ &= \left\|\bar{\partial}^2\left(f(t_n) - F^n\right)\right\|_{J_h^*} + \left\|\bar{\partial}^2(-\Delta_h U^n)\right\|_{J_h^*}\\ &\leq \left\|\bar{\partial}^2\left(f(t_n) - F^n\right)\right\|_{J_h^*} + \|\nabla\bar{\partial}^2 U^n\|.\end{aligned}$$

For the handling of the data approximation errors in negative norms let us mention the discussion in the study by Ern *et al.* (2016).

Our next theorem states the optimality of the time error estimators $\mathcal{E}_i$. We understand optimality in the sense that the time error estimators have the same asymptotic behavior as corresponding *a priori* estimates. In view of Theorem 3.4 the time error estimators are of second order, which corresponds to the *a priori* estimates in Theorem 2.5. However, this requires that the estimators are independent of the time step and mesh size, i.e. there exists a constant C independent of τ and h, such that $\mathcal{E}_i \leq C$.

THEOREM 3.6 (Stability of the time estimators). Suppose the exact solution (u, p) and data f, u_0 are sufficiently regular. Let $(X_h, Y_h) \subseteq X \times Y$ be a mixed finite element approximation on a quasi-uniform triangulation $\mathcal{T}_h$ of order $k \geq 2$ and let (U^n, P^n) be solution of equation (2.9) with $(X_n, Y_n) = (X_h, Y_h)$ $\forall$ $n = 0, \ldots, N$. Then the time error estimators $\mathcal{E}_1$–$\mathcal{E}_7$ defined in (3.3) are stable, i.e.

$$\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3 + \mathcal{E}_4 + \mathcal{E}_5 + \mathcal{E}_6 + \mathcal{E}_7 \leq C(u,p,f),$$

where C is independent of the time step τ and the mesh size h.

The proof is outlined in Section 5. Due to its technical difficulties we have to restrict our approximations in equation (2.9) to a constant, quasi-uniform mesh. This assumption, of course, does not go together with spatial adaptivity. However, this restriction is only necessary for the theoretical result of stability, not for the proof of the *a posteriori* estimates in Theorem 3.4.

4. Proofs of the *a posteriori* estimates

In this section the proof of the *a posteriori* error estimates in Theorem 3.4 is presented. Let us start with some auxiliary statements.

In the first step we deduce the error equation using the three-point reconstruction $\hat{U}(t)$:

LEMMA 4.1 (Error equation). Let $n \geq 2$ and $t \in I_n$. Then we have the following identity in H^{-1}:

$$\partial_t(\hat{U} - u)(t) - \Delta(U - u)(t) + \nabla(P - p)(t) = -\Delta(U - \mathcal{R}_u)(t) + \nabla(P - \mathcal{R}_p)(t) + \Upsilon(t), \tag{4.1}$$

where for $t \in I_n$

$$\Upsilon(t) := f(t_n) - f(t) + (t - t_n)\left(\bar{\partial}^2 U^n - \Delta\bar{\partial}\mathcal{R}_u^n + \nabla\bar{\partial}\mathcal{R}_p^n\right) + \left(\bar{\partial}^B - \bar{\partial}_n^B\right) U^n. \tag{4.2}$$

Proof. Using the definitions of $\hat{U}(t)$ in equation (3.2) and the Stokes reconstruction inequation (3.4) one has for $n \geq 2$,

$$
\begin{aligned}
\partial_t \hat{U}(t) - \Delta U(t) + \nabla P(t) &= \bar{\partial}^B U^n + (t-t_n)\bar{\partial}^2 U^n - \Delta(U-\mathcal{R}_u)(t) + \nabla(P-\mathcal{R}_p)(t) - \Delta\mathcal{R}_u^n + \nabla\mathcal{R}_p^n \\
&\quad + (t-t_n)\left(-\Delta\bar{\partial}\mathcal{R}_u^n + \nabla\bar{\partial}\mathcal{R}_p^n\right) \\
&= \bar{\partial}_n^B U^n - \Delta_n U^n + B_n^\top P^n - F^n + f\big(t_n\big) + (t-t_n)\left(\bar{\partial}^2 U^n - \Delta\bar{\partial}\mathcal{R}_u^n + \nabla\bar{\partial}\mathcal{R}_p^n\right) \\
&\quad - \Delta(U-\mathcal{R}_u)(t) + \nabla(P-\mathcal{R}_p)(t) + \left(\bar{\partial}^B - \bar{\partial}_n^B\right)U^n \\
&= f(t_n) + (t-t_n)\left(\bar{\partial}^2 U^n - \Delta\bar{\partial}\mathcal{R}_u^n + \nabla\bar{\partial}\mathcal{R}_p^n\right) - \Delta(U-\mathcal{R}_u)(t) \\
&\quad + \nabla(P-\mathcal{R}_p)(t) + \left(\bar{\partial}^B - \bar{\partial}_n^B\right)U^n.
\end{aligned}
$$

Now subtract equation (2.1) to finish the proof. □

4.1 *Estimating initial errors*

Next we present upper bounds for the initial errors.

LEMMA 4.2 (Initial *a posteriori* error estimates). Let (u,p) be the exact solution of equation (2.1), $X_2 = X_1 = X_0$ and $(U^0, P^0) := (\mathcal{S}_n^u(u_0,p_0), \mathcal{S}_n^p(u_0,p_0))$ the solution of equation (2.4). Let (U^1,P^1) be given by equation (2.10) and (U^2,P^2) by equation (2.9). Then

$$
\begin{aligned}
&\|\hat{U}-u\|^2_{L^\infty((0,t_1);J^*)} + \|U-u\|^2_{L^2((0,t_1);L^2)} \\
&\quad \lesssim \eta^0(J^*) + \int_0^{t_1}\left(\left\|(\hat{U}-U)(t)\right\|^2 + \|(U-\mathcal{R}_u)(t)\|^2 + \|\Upsilon(t)\|^2_{J^*}\right)\mathrm{d}t,
\end{aligned}
\tag{4.3}
$$

$$
\begin{aligned}
&\|U-u\|^2_{L^\infty((0,t_1);L^2)} \\
&\quad \lesssim \eta^0(L^2) + \max_{t\in[0,t_1]}\left\|(U-\hat{U})(t)\right\|^2 + \max_{t\in[0,t_1]}\left\|(\hat{\mathcal{R}}_u-\hat{U})(t)\right\|^2 \\
&\qquad + \int_0^{t_1}\left(\left\|\partial_t(\hat{\mathcal{R}}_u-\hat{U})(t)\right\|^2_{J^*} + \left\|\nabla(\mathcal{R}_u-\hat{\mathcal{R}}_u)(t)\right\|^2 + \|\Upsilon(t)\|^2_{J^*}\right)\mathrm{d}t,
\end{aligned}
\tag{4.4}
$$

$$
\begin{aligned}
&\|\nabla(U-u)\|^2_{L^2((0,t_1);L^2)} \\
&\quad \lesssim \eta^0(L^2) + \int_0^{t_1}\|\nabla(U-\mathcal{R}_u)(t)\|^2\,\mathrm{d}t \\
&\qquad + \int_0^{t_1}\left(\left\|\partial_t(\hat{\mathcal{R}}_u-\hat{U})(t)\right\|^2_{J^*} + \left\|\nabla(\mathcal{R}_u-\hat{\mathcal{R}}_u)(t)\right\|^2 + \|\Upsilon(t)\|^2_{J^*}\right)\mathrm{d}t.
\end{aligned}
\tag{4.5}
$$

$$
\begin{aligned}
&\|\nabla(U-u)\|^2_{L^\infty((0,t_1);L^2)}\\
&\quad\lesssim \eta^0\left(H_0^1\right)+\max_{t\in[0,t_1]}\left\|\nabla(U-\hat U)(t)\right\|^2+\max_{t\in[0,t_1]}\left\|\nabla(\hat{\mathcal{R}}_u-\hat U)(t)\right\|^2\\
&\qquad+\int_0^{t_1}\left(\left\|\partial_t(\hat{\mathcal{R}}_u-\hat U)(t)\right\|^2+\left\|\Delta(\mathcal{R}_u-\hat{\mathcal{R}}_u)(t)\right\|^2+\|\Upsilon(t)\|^2\right)\mathrm{d}t, \qquad (4.6)
\end{aligned}
$$

$$
\begin{aligned}
&\|P-p\|^2_{L^2((0,t_1);L^2)}\\
&\lesssim \eta^0\left(H_0^1\right)+\int_0^{t_1}\left\|(P-\mathcal{R}_p)(t)\right\|^2\mathrm{d}t\\
&\qquad+\int_0^{t_1}\left(\left\|\partial_t(\hat{\mathcal{R}}_u-\hat U)(t)\right\|^2+\left\|\Delta(\mathcal{R}_u-\hat{\mathcal{R}}_u)(t)\right\|^2+\|\Upsilon(t)\|^2\right)\mathrm{d}t, \qquad (4.7)
\end{aligned}
$$

where η^0 is defined in Assumption 3.2 and

$$
\Upsilon := (t-t_{1/2})\left(\bar\partial^2U^2-\Delta_1\bar\partial U^1+B_1^\top\bar\partial P^1-\bar\partial F^1\right)+f(t_1)+(t-t_1)\bar\partial f(t_1)-f(t), \qquad t\in I_1. \quad (4.8)
$$

Proof. For $t\in I_1=(0,t_1]$ define the three-point reconstruction

$$
\hat U(t)=U(t)+\frac{1}{2}(t-t_0)(t-t_1)\bar\partial^2U^2, \qquad \partial_t\hat U(t)=\bar\partial U^1+(t-t_{1/2})\bar\partial^2U^2, \quad (4.9)
$$

where the identity $t-t_{1/2}=\frac{\tau}{2}+(t-t_1)$ was used. The interpolant $\hat U(t)$ fulfills (in H^{-1})

$$
\begin{aligned}
&\partial_t\hat U(t)-\Delta U(t)+\nabla P(t)\\
&\quad=\bar\partial U^1+(t-t_{1/2})\bar\partial^2U^2-\Delta(U-\mathcal{R}_u)(t)+\nabla(P-\mathcal{R}_p)(t)\\
&\qquad-\Delta\mathcal{R}_u^1+\nabla\mathcal{R}_p^1+(t-t_1)\left(-\Delta\bar\partial\mathcal{R}_u^1+\nabla\bar\partial\mathcal{R}_p^1\right)\\
&\quad=\Delta_1U^{1/2}-B_1^\top P^{1/2}+F^{1/2}+(t-t_{1/2})\bar\partial^2U^2-\Delta(U-\mathcal{R}_u)(t)+\nabla(P-\mathcal{R}_p)(t)\\
&\qquad-\Delta_1U^1+B_1^\top P^1-F^1+f(t_1)+(t-t_1)\bar\partial\left(-\Delta_1U^1+B_1^\top P^1-F^1+f(t_1)\right)\\
&\quad=(t-t_{1/2})\left(\bar\partial^2U^2-\Delta_1\bar\partial U^1+B_1^\top\bar\partial P^1-\bar\partial F^1\right)+f(t_1)+(t-t_1)\bar\partial f(t_1)\\
&\qquad-\Delta(U-\mathcal{R}_u)(t)+\nabla(P-\mathcal{R}_p)(t).
\end{aligned}
$$

Note we have used the fact that $X_2=X_1=X_0$. Subtract equation (2.1) from the above identity to get the error equation for $t\in I_1$

$$
\partial_t(\hat U-u)(t)-\Delta(U-u)(t)+\nabla(P-p)(t)=\Upsilon(t)-\Delta(U-\mathcal{R}_u)(t)+\nabla(P-\mathcal{R}_p)(t), \quad (4.10)
$$

where Υ is defined inequation (4.8).

To prove equation (4.3) test equation (4.10) by $w(t) := \mathcal{A}^{-1}(\hat{U}-u)(t)$, i.e. $(w(t), q(t))$ is solution of

$$-\Delta w(t) + \nabla q(t) = (\hat{U}-u)(t), \quad \operatorname{div} w(t) = 0, \quad w(t)|_{\Gamma} = 0.$$

For the first term on the left-hand side, one gets

$$\left\langle \partial_t(\hat{U}-u)(t), w(t) \right\rangle = \frac{1}{2}\frac{d}{\mathrm{d}t}\,\|\nabla w(t)\|^2.$$

The second term can be treated with the help of the identity

$$2\langle a-b, c-b\rangle = \|a-b\|^2 + \|c-b\|^2 - \|a-c\|^2 \tag{4.11}$$

and the fact that the interpolant of the Stokes reconstruction $\mathcal{R}_u(t)$ is divergence free:

$$\begin{aligned}
\langle \nabla(U-u)(t), \nabla w(t)\rangle &= \langle (U-u)(t), -\Delta w(t)\rangle = \left\langle (U-u)(t), (\hat{U}-u)(t)\right\rangle - \langle (U-u)(t), \nabla q(t)\rangle \\
&= \frac{1}{2}\,\|(U-u)(t)\|^2 + \frac{1}{2}\left\|(\hat{U}-u)(t)\right\|^2 - \frac{1}{2}\left\|(\hat{U}-U)(t)\right\|^2 - \langle (U-\mathcal{R}_u)(t), \nabla q(t)\rangle \\
&\geq \frac{1}{2}\,\|(U-u)(t)\|^2 + \frac{1}{2}\left\|(\hat{U}-u)(t)\right\|^2 - \frac{1}{2}\left\|(\hat{U}-U)(t)\right\|^2 \\
&\quad - C\,\|(U-\mathcal{R}_u)(t)\|\,\left\|(\hat{U}-u)(t)\right\|.
\end{aligned}$$

For the last inequality the regularity estimate (2.7) for $\nabla q(t)$ has been used.

For the terms on the right-hand side of equation (4.10), we obtain

$$\begin{aligned}
\langle -\Delta(U-\mathcal{R}_u)(t) + \Upsilon(t), w(t)\rangle &= \langle (U-\mathcal{R}_u)(t), -\Delta w(t)\rangle + \langle \Upsilon(t), w(t)\rangle \\
&= \left\langle (U-\mathcal{R}_u)(t), -\nabla q(t) + (\hat{U}-u)(t)\right\rangle + \langle \Upsilon(t), w(t)\rangle \\
&\lesssim \|(U-\mathcal{R}_u)(t)\|^2 + \left\|(\hat{U}-u)(t)\right\| + \|\Upsilon(t)\|_{J^*}\,\|\nabla w(t)\|.
\end{aligned}$$

Now, integrating from $t = 0$ to t_1, yields

$$\begin{aligned}
&\max_{t\in[0,t_1]}\left\|(\hat{U}-u)(t)\right\|_{J^*}^2 + \int_0^{t_1}\left(\|(U-u)(t)\|^2 + \left\|(\hat{U}-u)(t)\right\|^2\right)\mathrm{d}t \\
&\quad\lesssim \left\|(\hat{U}-u)(0)\right\|_{J^*}^2 + \int_0^{t_1}\left(\left\|(\hat{U}-U)(t)\right\|^2 + \|(U-\mathcal{R}_u)(t)\|^2 + \|\Upsilon(t)\|_{J^*}^2\right)\mathrm{d}t,
\end{aligned} \tag{4.12}$$

with $\|(\hat{U}-u)(0)\|_{J^*}^2 = \|U^0 - u_0\|_{J^*}^2 = \|\mathcal{S}_n^u(u_0, p_0) - u_0\|_{J^*}^2 \leq \eta^0(J^*)$.

Rearrange equation (4.10) to get for $t \in I_1$

$$\partial_t(\hat{\mathcal{R}}_u - u)(t) - \Delta\,(\mathcal{R}_u - u)(t) + \nabla(\mathcal{R}_p - p)(t) = \partial_t(\hat{\mathcal{R}}_u - \hat{U})(t) + \Upsilon(t). \tag{4.13}$$

Test equation (4.13) with $(\mathcal{R}_u - u)(t) \in J$, use identity (4.11) and integrate from $t = 0$ to t_1 to obtain

$$\max_{t\in[0,t_1]} \left\|(\hat{\mathcal{R}}_u - u)(t)\right\|^2 + \int_0^{t_1} \left(\left\|\nabla(\hat{\mathcal{R}}_u - u)(t)\right\|^2 + \|\nabla(\mathcal{R}_u - u)(t)\|^2\right) \mathrm{d}t$$
$$\lesssim \left\|(\hat{\mathcal{R}}_u - u)(0)\right\|^2 + \int_0^{t_1} \left(\left\|\partial_t(\hat{\mathcal{R}}_u - \hat{U})(t)\right\|_{J^*}^2 + \left\|\nabla(\mathcal{R}_u - \hat{\mathcal{R}}_u)(t)\right\|^2 + \|\Upsilon(t)\|_{J^*}^2\right) \mathrm{d}t. \quad (4.14)$$

Finally, the triangle inequality yields

$$\begin{aligned}\max_{t\in[0,t_1]} \|(U-u)(t)\|^2 &\le \max_{t\in[0,t_1]} \left\|(U-\hat{U})(t)\right\|^2 + \max_{t\in[0,t_1]} \left\|(\hat{\mathcal{R}}_u - \hat{U})(t)\right\|^2 + \max_{t\in[0,t_1]} \left\|(\hat{\mathcal{R}}_u - u)(t)\right\|^2 \\ &\lesssim \eta^0(L^2) + \max_{t\in[0,t_1]} \left\|(U-\hat{U})(t)\right\|^2 + \max_{t\in[0,t_1]} \left\|(\hat{\mathcal{R}}_u - \hat{U})(t)\right\|^2 \\ &\quad + \int_0^{t_1} \left(\left\|\partial_t(\hat{\mathcal{R}}_u - \hat{U})(t)\right\|_{J^*}^2 + \left\|\nabla(\mathcal{R}_u - \hat{\mathcal{R}}_u)(t)\right\|^2 + \|\Upsilon(t)\|_{J^*}^2\right) \mathrm{d}t.\end{aligned}$$

To prove equation (4.5) the error is split into

$$\int_0^{t_1} \|\nabla(U-u)(t)\|^2 \,\mathrm{d}t \le \int_0^{t_1} \|\nabla(U-\mathcal{R}_u)(t)\|^2 \,\mathrm{d}t + \int_0^{t_1} \|\nabla(\mathcal{R}_u - u)(t)\|^2$$

and then equation (4.14) is used.

The proofs of equations (4.6) and (4.7) are omitted, since they can be achieved in exactly the same way as the proofs of equations (3.10) and (3.11) in Theorem 3.4 using the error equation (4.13). $\square$

4.2 *Error estimator*

In the following lemma we study how to handle mesh changes within the *a posteriori* framework. One fundamental technical difficulty arises from the fact that because of mesh changes $U^n \in X_n \neq X_{n-1} \ni U^{n-1}$ in general, and thus terms like $\bar{\partial}U^n$ are not well defined as a finite element function.

LEMMA 4.3 (Changing meshes). For an arbitrary function $W^s \in X_s$, $s = 1, \ldots, N$ and for arbitrary $n = 1, \ldots, N$ as well $m = 1, \ldots, k+1$, we have

$$\left\|(I - \Pi_n)W^s\right\|^2 = \|W^s\|^2 - \|\Pi_n W^s\|^2, \quad (4.15)$$

$$\left\|(I - \Pi_n)W^s\right\|_{-m}^2 \lesssim h_n^{2m}\left(\|W^s\|^2 - \|\Pi_n W^s\|^2\right), \quad (4.16)$$

$$\|\bar{\partial}W^s\|_{-m}^2 \lesssim \|\bar{\partial}_s W^s\|_{-m}^2 + C\frac{h_n^{2m}}{\tau^2}\left(\|W^{s-1}\|^2 - \|\Pi_n W^{s-1}\|^2\right), \quad (4.17)$$

$$\left\|(I - \Pi_n)\bar{\partial}^B W^s\right\|_{-m}^2 \lesssim \frac{h_n^{2m}}{\tau^2}\xi_B^{s,n}(W), \quad (4.18)$$

$$\|\bar{\partial}^2 W^s\|^2 \le \left\|\bar{\partial}_s^2 W^s\right\|^2 + C\frac{1}{\tau^4}\xi_B^{s,s}(W), \quad (4.19)$$

where the local BDF2-changing mesh operator $\xi_B^{s,n}(W)$ is defined in equation (3.6).

Proof. By definition of the L^2 projection

$$\big\langle (I-\Pi_n)W^s, v_n\big\rangle = 0 \qquad \forall\, v_n \in X_n$$

and the identity

$$2\langle a-b, a\rangle = \|a\|^2 - \|b\|^2 + \|a-b\|^2 \tag{4.20}$$

it follows

$$0 = 2\big\langle (I-\Pi_n)W^s, \Pi_n W^s\big\rangle = -\|\Pi_n W^s\|^2 + \|W^s\|^2 - \big\|(I-\Pi_n)W^s\big\|^2,$$

which implies equation (4.15).

To prove equation (4.16) observe that

$$\big\|(I-\Pi_n)W^s\big\|_{-m} = \sup_{v\in H_0^1\cap H^m} \frac{\langle (I-\Pi_n)W^s, v\rangle}{\|v\|_{H^m}} = \sup_{v\in H_0^1\cap H^m} \frac{\langle (I-\Pi_n)W^s, v-\mathcal{I}_n v\rangle}{\|v\|_{H^m}} \le C\, h_n^m \,\big\|(I-\Pi_n)W^s\big\|.$$

Now using equation (4.15) yields the result. To derive equation (4.17) use the estimate

$$\|\bar\partial W^s\|_{-m}^2 \lesssim \|\bar\partial_s W^s\|_{-m}^2 + \big\|(I-\Pi_s)\bar\partial W^s\big\|_{-m}^2 = \|\bar\partial_s W^s\|_{-m}^2 + \frac{1}{\tau^2}\,\Big\|(I-\Pi_s)W^{s-1}\Big\|_{-m}^2$$

and apply equation (4.16) to the last term on the right-hand side.

Equation (4.18) follows from the definition of the $\bar\partial^B$ operator and equation (4.16):

$$\Big\|(I-\Pi_n)\bar\partial^B W^s\Big\|_{-m}^2 \lesssim \frac{1}{\tau^2}\left(\big\|(I-\Pi_n)W^s\big\|_{-m}^2 + \Big\|(I-\Pi_n)W^{s-1}\Big\|_{-m}^2 + \Big\|(I-\Pi_n)W^{s-2}\Big\|_{-m}^2\right).$$

To prove equation (4.19) rewrite $\bar\partial^2 W^s$ as

$$\bar\partial^2 W^s = \frac{2}{\tau}(\bar\partial^B W^s - \bar\partial W^s).$$

Then one obtains by using equations (4.15) and (4.18)

$$\begin{aligned}
\|\bar\partial^2 W^s\|^2 &= \Big\|\bar\partial_s^2 W^s\Big\|^2 + \Big\|(I-\Pi_s)\bar\partial^2 W^s\Big\|^2 \\
&\le \Big\|\bar\partial_s^2 W^s\Big\|^2 + C\frac{1}{\tau^2}\left(\Big\|(I-\Pi_s)\bar\partial^B W^s\Big\|^2 + \frac{1}{\tau^2}\Big\|(I-\Pi_s)W^{s-1}\Big\|^2\right) \\
&\le \Big\|\bar\partial_s^2 W^s\Big\|^2 + C\frac{1}{\tau^4}\xi_B^{s,s}(W).
\end{aligned}$$

Note that the difference $\|(I-\Pi_s)W^{s-1}\|^2$ is already contained in the definition of $\xi_B^{s,s}$. □

Below we derive the spatial and reconstruction estimators. We start with bounds of the error between the discrete solution and the reconstruction.

LEMMA 4.4 Let (U^n, P^n) be the discrete solution of equation (2.1) and equation (2.10) and $(\mathcal{R}_u^n, \mathcal{R}_p^n)$ be the reconstruction defined in equation (3.2). Then we have the following estimates for the interpolants

$$\int_0^T \|(U - \mathcal{R}_u)(t)\|^2 \lesssim \Lambda_1, \tag{4.21}$$

$$\max_{t\in[0,T]} \left\|(\hat{U} - \hat{\mathcal{R}}_u)(t)\right\|^2 \lesssim \Lambda_2, \tag{4.22}$$

$$\int_0^T \left\|\partial_t(\hat{U} - \hat{\mathcal{R}}_u)(t)\right\|_{J^*}^2 \lesssim \Lambda_3 + \Xi_2, \tag{4.23}$$

$$\int_0^T \|\nabla(U - \mathcal{R}_u)(t)\|^2 \lesssim \Lambda_4, \tag{4.24}$$

$$\max_{t\in[0,T]} \left\|\nabla(\hat{U} - \hat{\mathcal{R}}_u)(t)\right\|^2 \lesssim \Lambda_5, \tag{4.25}$$

$$\int_0^T \left\|\partial_t(\hat{U} - \hat{\mathcal{R}}_u)(t)\right\|^2 \lesssim \Lambda_6 + \Xi_6, \tag{4.26}$$

$$\int_0^T \left\|(P - \mathcal{R}_p)(t)\right\|^2 \lesssim \Lambda_7, \tag{4.27}$$

where the estimators are defined in Definition 3.3.

Proof. Due to the definitions of the interpolant in equation (3.1) and the reconstruction in equation (3.2) and equation (4.9), it holds

$$\begin{aligned}
\|(U - \mathcal{R}_u)(t)\| &\lesssim \left\|U^n - \mathcal{R}_u^n\right\| + \left\|U^{n-1} - \mathcal{R}_u^{n-1}\right\|, && t \in I_n,\ n \geq 1,\\
\left\|(\hat{U} - \hat{\mathcal{R}}_u)(t)\right\| &\lesssim \left\|U^n - \mathcal{R}_u^n\right\| + \left\|U^{n-1} - \mathcal{R}_u^{n-1}\right\| + \left\|U^{n-2} - \mathcal{R}_u^{n-2}\right\|, && t \in I_n,\ n \geq 2,\\
\left\|(\hat{U} - \hat{\mathcal{R}}_u)(t)\right\| &\lesssim \left\|U^2 - \mathcal{R}_u^2\right\| + \left\|U^1 - \mathcal{R}_u^1\right\| + \left\|U^0 - \mathcal{R}_u^0\right\|, && t \in I_1.
\end{aligned}$$

Hence equations (4.21), (4.22), (4.24), (4.25) and (4.27) follow directly by applying the above inequalities and integrating over time. To prove equation (4.23) and equation (4.26), the definition of the three-point reconstruction yields

$$\begin{aligned}
\int_0^T \left\|\partial_t(\hat{U} - \hat{\mathcal{R}}_u)(t)\right\|_*^2 &\lesssim \sum_{n=1}^N \tau \left\|\bar{\partial}\left(U^n - \mathcal{R}_u^n\right)\right\|_*^2\\
&\lesssim \sum_{n=1}^N \tau \left\|\bar{\partial}_n\left(U^n - \bar{\partial}\mathcal{R}_u^n\right)\right\|_*^2 + \sum_{n=2}^N \tau \left\|(\bar{\partial} - \bar{\partial}_n)U^n\right\|_*^2, \quad * \in \{L^2, J^*\},
\end{aligned}$$

respectively. This proves equations (4.23) and (4.26), where the last term on the right-hand side was absorbed in the estimators Ξ_2 and Ξ_6, respectively. □

In the next lemma, upper bounds for the Stokes reconstruction are given.

LEMMA 4.5 Let U^n be the discrete solution of equation (2.1) and equation (2.10) and $\mathcal{R}_u^n$ the corresponding reconstruction defined in equation (3.2). Further let $X_2 = X_1 = X_0$ and set $U^{-1} := U^1 - 2\tau(\Delta_0 U^0 - B_0^\top P^0 + F^0) \in X_0$ and $U^{-2} := -6\tau(\Delta_0 U^0 - B_0^\top P^0 + F^0) - 3U^0 + 4U^1 \in X_0$ to simplify the presentation. Then

$$\int_0^T \left\| \nabla(\mathcal{R}_u - \hat{\mathcal{R}}_u)(t) \right\|^2 \mathrm{d}t \lesssim \tau^4 \mathcal{E}_4 + \Xi_4, \tag{4.28}$$

$$\int_0^T \left\| \Delta(\mathcal{R}_u - \hat{\mathcal{R}}_u)(t) \right\|^2 \mathrm{d}t \lesssim \tau^4 \mathcal{E}_5 + \Xi_5, \tag{4.29}$$

where the estimators are defined in Definition 3.3.

Proof. By definition of the reconstruction in equation (3.2) one has for $n \geq 2$

$$\begin{aligned} -\Delta\bar{\partial}^2\mathcal{R}_u^n + \nabla\bar{\partial}^2\mathcal{R}_p^n = \frac{1}{\tau^2}\Big(\Big(-\Delta_n U^n + B_n^\top P^n - F^n\Big) - 2\Big(-\Delta_{n-1}U^{n-1} + B_{n-1}^\top P^{n-1} - F^{n-1}\Big) \\ + \Big(-\Delta_{n-2}U^{n-2} + B_{n-2}^\top P^{n-2} - F^{n-2}\Big)\Big) + \bar{\partial}^2 f^n. \end{aligned}$$

By applying equation (2.9), we get

$$\begin{aligned} -\Delta\bar{\partial}^2\mathcal{R}_u^n + \nabla\bar{\partial}^2\mathcal{R}_p^n &= \frac{1}{\tau^2}\Big(-\bar{\partial}_n^B U^n + 2\bar{\partial}_{n-1}^B U^{n-1} - \bar{\partial}_{n-2}^B U^{n-2}\Big) + \bar{\partial}^2 f^n \\ &= -\bar{\partial}_n^2\bar{\partial}^B U^n + \bar{\partial}^2 f^n + \frac{1}{\tau^2}\Big(2\Big(\bar{\partial}_{n-1}^B - \bar{\partial}_n^B\Big)U^{n-1} - \Big(\bar{\partial}_{n-2}^B - \bar{\partial}_n^B\Big)U^{n-2}\Big). \end{aligned}$$

Note that in the case $n = 2, 3$ it holds $\bar{\partial}_1^B U^1 = \Delta_1 U^1 + F^1$ and $\bar{\partial}_0^B U^0 = \Delta_0 U^0 + F^0$ due to the definition of U^{-1} and U^{-2}.

Now we use the definition of the three-point reconstruction in equation (3.2) and equation (4.9) to get

$$\begin{aligned} \int_0^T \left\| \nabla(\mathcal{R}_u - \hat{\mathcal{R}}_u)(t) \right\|^2 \mathrm{d}t &\lesssim \sum_{n=2}^N \tau^5 \left\| \nabla\bar{\partial}^2\mathcal{R}_u^n \right\|^2 \\ &\lesssim \sum_{n=2}^N \tau^5 \left\| \bar{\partial}^2 f^n - \bar{\partial}_n^2\bar{\partial}^B U^n \right\|_{J^*}^2 \\ &\quad + \sum_{n=2}^N \tau \left(\underbrace{\left\| \Big(\bar{\partial}_{n-1}^B - \bar{\partial}_n^B\Big)U^{n-1} \right\|_{J^*}^2 + \left\| \Big(\bar{\partial}_{n-2}^B - \bar{\partial}_n^B\Big)U^{n-2} \right\|_{J^*}^2}_{=(*)} \right). \end{aligned}$$

The last term on the right-hand side is handled by using equation (4.18):

$$(*) \lesssim \left\|(I-\Pi_{n-1})\bar{\partial}^B U^{n-1}\right\|_{J^*}^2 + \left\|(I-\Pi_n)\bar{\partial}^B U^{n-1}\right\|_{J^*}^2 + \left\|(I-\Pi_{n-2})\bar{\partial}^B U^{n-2}\right\|_{J^*}^2 + \left\|(I-\Pi_n)\bar{\partial}^B U^{n-2}\right\|_{J^*}^2$$
$$\lesssim \frac{h_{n-1}^2}{\tau^2}\xi_B^{n-1,n-1}(U) + \frac{h_n^2}{\tau^2}\xi_B^{n-1,n}(U) + \frac{h_{n-2}^2}{\tau^2}\xi_B^{n-2,n-2}(U) + \frac{h_n^2}{\tau^2}\xi_B^{n-2,n}(U).$$

Thus, we obtain

$$\int_0^T \left\|\nabla(\mathcal{R}_u - \hat{\mathcal{R}}_u)(t)\right\|^2 \mathrm{d}t \lesssim \mathcal{E}_4 + \Xi_4.$$

In the same way equation (4.29) is shown by using Assumption 2.4. □

Now we bound the residual Υ of the error equation (4.1) and the difference between the linear interpolant (3.1) and three-point reconstruction (3.2).

LEMMA 4.6 Let U^n be the discrete solution of equation (2.1) and equation (2.10), respectively. Then we have

$$\int_0^T \left\|(\hat{U} - U)(t)\right\|^2 \lesssim \tau^4\mathcal{E}_1 + \Xi_1, \tag{4.30}$$

$$\int_0^T \|\Upsilon(t)\|_{J^*}^2 \lesssim \tau^4\mathcal{E}_2 + \Xi_2 + \mathcal{F}_1(t) + I_1, \tag{4.31}$$

$$\max_{t\in[0,T]} \left\|(\hat{U} - U)(t)\right\|^2 \lesssim \tau^4\mathcal{E}_3 + \Xi_3, \tag{4.32}$$

$$\int_0^T \|\Upsilon(t)\|^2 \,\mathrm{d}t \lesssim \tau^4\mathcal{E}_6 + \Xi_6 + \mathcal{F}_2(t) + I_2, \tag{4.33}$$

$$\max_{t\in[0,T]} \left\|\nabla(\hat{U} - U)(t)\right\|^2 \lesssim \tau^4\mathcal{E}_7 + \Xi_7, \tag{4.34}$$

where the estimators are defined in Definition 3.3.

Proof. Equations (4.30), (4.32) and (4.34) are easily shown by the definition of the reconstruction, cf. equations (3.2) and (4.9), and equation (4.19):

$$\left\|(\hat{U} - U)(t)\right\|^2 \lesssim \tau^4\|\bar{\partial}^2 U^n\|^2 \lesssim \tau^4\left\|\bar{\partial}_n^2 U^n\right\|^2 + \xi_B^{n,n}(U).$$

For the proof of equations (4.31) and (4.33) we reformulate $\Upsilon(t)$. In the time interval I_n, $n \geq 2$, we have

$$\Upsilon(t) = f(t_n) - f(t) + (t - t_n)\left(\bar{\partial}^2 U^n - \Delta\bar{\partial}\mathcal{R}_u^n + \nabla\bar{\partial}\mathcal{R}_p^n\right) + \left(\bar{\partial}^B - \bar{\partial}_n^B\right)U^n$$
$$= f(t_n) - f(t) + \left(\bar{\partial}^B - \bar{\partial}_n^B\right)U^n + (t - t_n)\underbrace{\left(\bar{\partial}^2 U^n + \frac{1}{\tau}\left(-\Delta\mathcal{R}_u^n + \nabla\mathcal{R}_p^n + \Delta\mathcal{R}_u^{n-1} - \nabla\mathcal{R}_p^{n-1}\right)\right)}_{=(*)}.$$

Let us simplify the last bracket further by using equation (2.9).

$$
\begin{aligned}
(*) &= \bar{\partial}^2 U^n + \frac{1}{\tau}\left(-\Delta_n U^n + B_n^\top P^n - F^n + f(t_n) - \left(-\Delta_{n-1} U^{n-1} + B_{n-1}^\top P^{n-1} - F^{n-1} + f(t_{n-1})\right)\right) \\
&= \bar{\partial} f(t_n) + \frac{1}{\tau}\left(\bar{\partial} U^n - \bar{\partial} U^{n-1} - \bar{\partial}_n^B U^n + \bar{\partial}_{n-1}^B U^{n-1}\right) \\
&= \bar{\partial} f(t_n) - \left(\bar{\partial}_n^2 U^n - \bar{\partial}_{n-1}^2 U^{n-1}\right) + \frac{1}{\tau}(\bar{\partial} - \bar{\partial}_n) U^n - \frac{1}{\tau}(\bar{\partial} - \bar{\partial}_{n-1}) U^{n-1}.
\end{aligned}
$$

Note that we have used the definition of U^{-1} in Lemma 4.5. Then the triangle inequality yields

$$
\begin{aligned}
\|\Upsilon(t)\|_* \lesssim\ & \left\| f(t_n) + (t - t_n)\bar{\partial} f(t_n) - f(t)\right\|_* + \left\|(\bar{\partial} - \bar{\partial}_n^B) U^n\right\|_* \\
& + \tau \left\|\bar{\partial}_n^2 U^n - \bar{\partial}_{n-1}^2 U^{n-1}\right\|_* + \left\|(\bar{\partial} - \bar{\partial}_n) U^n\right\|_* + \left\|(\bar{\partial} - \bar{\partial}_{n-1}) U^{n-1}\right\|_*, \quad * \in \{L^2, J^*\}.
\end{aligned}
$$

For $t \in I_1$, we have cf. equation (4.8),

$$
\|\Upsilon(t)\|_* \lesssim \left\| f(t_1) + (t - t_1)\bar{\partial} f(t_1) - f(t)\right\|_* + \tau \|\Psi_1\|_*, \quad * \in \{L^2, J^*\}.
$$

Together we obtain

$$
\begin{aligned}
\int_0^T \|\Upsilon(t)\|_*^2 \lesssim\ & \tau^3 \|\Psi_1\|_*^2 + \int_0^T \left\| f(t_n) + (t - t_n)\bar{\partial} f(t_n) - f(t)\right\|_* \mathrm{d}t \\
& + \sum_{n=2}^N \tau^3 \left\|\bar{\partial}_n^2 U^n - \bar{\partial}_{n-1}^2 U^{n-1}\right\|_*^2 + \sum_{n=2}^N \tau \left\|\left(\bar{\partial} - \bar{\partial}_n^B\right) U^n\right\|_*^2 + \sum_{n=1}^N \tau \left\|(\bar{\partial} - \bar{\partial}_n) U^n\right\|_*^2.
\end{aligned}
$$

Now apply equation (4.18) to the fourth term on the right-hand side to prove equations (4.31) and (4.33). Note that the last term on the right-hand side can be absorbed by the fourth term. □

4.3 *Proof of Theorem* 3.4

Proof. **(Theorem 3.4)** First we prove equation (3.7). To this end we test equation (4.1) with $y \in J$ to get

$$
\left\langle \partial_t(\hat{U} - u)(t), y\right\rangle + \langle \nabla(U - u)(t), \nabla y\rangle = \langle -\Delta(U - \mathcal{R}_u)(t), y\rangle + \langle \Upsilon(t), y\rangle, \qquad t \in I_n,\ n \geq 2.
$$

Next, set $y = w(t) := \mathcal{A}^{-1}(\hat{U} - u)(t)$, i.e. $(w(t), q(t))$ is solution of

$$
-\Delta w(t) + \nabla q(t) = (\hat{U} - u)(t), \quad \operatorname{div} w(t) = 0, \quad w(t)|_\Gamma = 0.
$$

For the first term on the left-hand side, one gets

$$\langle \partial_t(\hat{U} - u)(t), w(t)\rangle = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla w(t)\|^2.$$

The second term is handled by the use of regularity estimate (2.7) for $\nabla q(t)$, the identity equation (4.11) and the fact that the interpolant of the Stokes reconstruction $\mathcal{R}_u(t)$ is divergence free:

$$\begin{aligned}
\langle \nabla(U-u)(t), \nabla w(t)\rangle = \langle (U-u)(t), -\Delta w(t)\rangle &= \Big\langle (U-u)(t), (\hat{U}-u)(t)\Big\rangle - \langle (U-u)(t), \nabla q(t)\rangle \\
&= \frac{1}{2}\|(U-u)(t)\|^2 + \frac{1}{2}\Big\|(\hat{U}-u)(t)\Big\|^2 - \frac{1}{2}\Big\|(\hat{U}-U)(t)\Big\|^2 - \langle (U-\mathcal{R}_u)(t), \nabla q(t)\rangle \\
&\geq \frac{1}{2}\|(U-u)(t)\|^2 + \frac{1}{2}\Big\|(\hat{U}-u)(t)\Big\|^2 - \frac{1}{2}\Big\|(\hat{U}-U)(t)\Big\|^2 \\
&\quad - \|(U-\mathcal{R}_u)(t)\| \Big\|(\hat{U}-u)(t)\Big\|.
\end{aligned}$$

The terms on the right-hand side are treated as follows

$$\begin{aligned}
\langle -\Delta(U-\mathcal{R}_u)(t) + \Upsilon(t), w(t)\rangle &= \langle (U-\mathcal{R}_u)(t), -\Delta w(t)\rangle + \langle \Upsilon(t), w(t)\rangle \\
&= \Big\langle (U-\mathcal{R}_u)(t), -\nabla q(t) + (\hat{U}-u)(t)\Big\rangle + \langle \Upsilon(t), w(t)\rangle \\
&\lesssim \|(U-\mathcal{R}_u)(t)\|^2 + \frac{1}{8}\Big\|(\hat{U}-u)(t)\Big\| + \|\Upsilon(t)\|_{J^*}\|\nabla w(t)\|.
\end{aligned}$$

Collecting these results, integrating from $t = t_1$ to T, using Gronwall's Lemma and Young's inequality one arrives at

$$\begin{aligned}
&\max_{t\in[t_1,T]}\|\nabla w(t)\|^2 + \int_{t_1}^{T}\left(\|(U-u)(t)\|^2 + \Big\|(\hat{U}-u)(t)\Big\|^2\right)\mathrm{d}t \\
&\quad\lesssim \|\nabla w(t_1)\|^2 + \int_{t_1}^{T}\left(\Big\|(\hat{U}-U)(t)\Big\|^2 + \|(U-\mathcal{R}_u)(t)\|^2 + \|\Upsilon(t)\|^2_{J^*}\right)\mathrm{d}t.
\end{aligned}$$

This combined with the initial estimate (4.3) yields

$$\int_0^T \|(U-u)(t)\|^2\,\mathrm{d}t \;\lesssim\; \eta^0(J^*) + \int_0^T\left(\Big\|(\hat{U}-U)(t)\Big\|^2 + \|(U-\mathcal{R}_u)(t)\|^2 + \|\Upsilon(t)\|^2_{J^*}\right)\mathrm{d}t.$$

Now use equations (4.21), (4.30) and (4.31) to prove equation (3.7).

To show equations (3.8) and (3.9), rearrange terms in equation (4.1) to get the following error equation for $t \in I_n$, $n \geq 2$,

$$\partial_t(\hat{\mathcal{R}}_u - u)(t) - \Delta(\mathcal{R}_u - u)(t) + \nabla(\mathcal{R}_p - p)(t) = \partial_t(\hat{\mathcal{R}}_u - \hat{U})(t) + \Upsilon(t). \tag{4.35}$$

Test equation (4.35) by $(\hat{\mathcal{R}}_u - u)(t) \in J$, use identity (4.11) and Young's inequality to get

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left\|(\hat{\mathcal{R}}_u - u)(t)\right\|^2 + \left\|\nabla(\hat{\mathcal{R}}_u - u)(t)\right\|^2 + \|\nabla(\mathcal{R}_u - u)(t)\|^2 \\ \lesssim \left\|\partial_t(\hat{\mathcal{R}}_u - \hat{U})(t)\right\|_{J^*}^2 + \left\|\nabla(\mathcal{R}_u - \hat{\mathcal{R}}_u)(t)\right\|^2 + \|\Upsilon\|_{J^*}^2.\end{aligned}$$

Integrating from $t = t_1$ to T yields

$$\begin{aligned}&\max_{t\in[t_1,T]}\left\|(\hat{\mathcal{R}}_u - u)(t)\right\|^2 + \int_{t_1}^T\left(\left\|\nabla(\hat{\mathcal{R}}_u - u)(t)\right\|^2 + \|\nabla(\mathcal{R}_u - u)(t)\|^2\right)\mathrm{d}t \\ &\lesssim \left\|(\hat{\mathcal{R}}_u - u)(t_1)\right\|^2 + \int_{t_1}^T\left(\left\|\partial_t(\hat{\mathcal{R}}_u - \hat{U})(t)\right\|_{J^*}^2 + \left\|\nabla(\mathcal{R}_u - \hat{\mathcal{R}}_u)(t)\right\|^2 + \|\Upsilon(t)\|_{J^*}^2\right)\mathrm{d}t. \qquad (4.36)\end{aligned}$$

Now combine this estimate with equation (4.4) and apply the triangle inequality to get

$$\begin{aligned}\max_{t\in[0,T]}\|(U-u)(t)\|^2 &\lesssim \max_{t\in[0,t_1]}\|(U-u)(t)\|^2 + \max_{t\in[t_1,T]}\left\|(U-\hat{U})(t)\right\|^2 \\ &\quad + \max_{t\in[t_1,T]}\left\|(\hat{\mathcal{R}}_u - \hat{U})(t)\right\|^2 + \max_{t\in[t_1,T]}\left\|(\hat{\mathcal{R}}_u - u)(t)\right\|^2 \\ &\lesssim \eta^0(L^2) + \max_{t\in[0,T]}\left\|(U-\hat{U})(t)\right\|^2 + \max_{t\in[0,T]}\left\|(\hat{\mathcal{R}}_u - \hat{U})(t)\right\|^2 \\ &\quad + \int_0^T\left(\left\|\partial_t(\hat{\mathcal{R}}_u - \hat{U})(t)\right\|_{J^*}^2 + \left\|\nabla(\mathcal{R}_u - \hat{\mathcal{R}}_u)(t)\right\|^2 + \|\Upsilon(t)\|_{J^*}^2\right)\mathrm{d}t.\end{aligned}$$

With equations (4.22), (4.23), (4.28), (4.31) and (4.32), the proof of equation (3.8) is finished. In a similar way equation (3.9) is bounded. Use the triangle inequality and the initial estimate (4.5) to get

$$\begin{aligned}\int_0^T\|\nabla(U-u)(t)\|^2 &\lesssim \int_0^{t_1}\|\nabla(U-u)(t)\|^2 + \int_{t_1}^T\|\nabla(U-\mathcal{R}_u)(t)\|^2 + \int_{t_1}^T\|\nabla(\mathcal{R}_u - u)(t)\|^2 \\ &\lesssim \eta^0(L^2) + \int_0^T\|\nabla(U-\mathcal{R}_u)(t)\|^2 \\ &\quad + \int_0^T\left(\left\|\partial_t(\hat{\mathcal{R}}_u - \hat{U})(t)\right\|_{J^*}^2 + \left\|\nabla(\mathcal{R}_u - \hat{\mathcal{R}}_u)(t)\right\|^2 + \|\Upsilon(t)\|_{J^*}^2\right)\mathrm{d}t,\end{aligned}$$

and then use equations (4.23), (4.24), (4.28) and (4.31).

To prove equations (3.10) and (3.11), test equation (4.35) by $\partial_t(\hat{\mathcal{R}}_u - u)(t) \in J$ to obtain for $t \in I_n$, $n \geq 2$,

$$\left\|\partial_t(\hat{\mathcal{R}}_u - u)(t)\right\|^2 + \left\langle -\Delta(\mathcal{R}_u - u)(t), \partial_t(\hat{\mathcal{R}}_u - u)(t)\right\rangle \lesssim \left\|\partial_t(\hat{\mathcal{R}}_u - \hat{U})(t)\right\|^2 + \|\Upsilon(t)\|^2,$$

where the inner product is treated as follows:

$$\left\langle -\Delta(\mathcal{R}_u - u)(t), \partial_t(\hat{\mathcal{R}}_u - u)(t)\right\rangle = \left\langle \nabla(\hat{\mathcal{R}}_u - u)(t), \nabla\partial_t(\hat{\mathcal{R}}_u - u)(t)\right\rangle + \left\langle -\Delta(\mathcal{R}_u - \hat{\mathcal{R}}_u), \partial_t(\hat{\mathcal{R}}_u - u)\right\rangle.$$

Then, integrate from $t = t_1$ to T to get

$$\begin{aligned}\int_{t_1}^{T} &\left\|\partial_t(\hat{\mathcal{R}}_u - u)(t)\right\|^2 + \max_{t\in[t_1,T]}\left\|\nabla(\hat{\mathcal{R}}_u - u)\right\|^2 \\ &\lesssim \left\|\nabla(\hat{\mathcal{R}}_u - u)(t_1)\right\|^2 + \int_{t_1}^{T}\left(\left\|\partial_t(\hat{\mathcal{R}}_u - \hat{U})(t)\right\|^2 + \left\|\Delta(\mathcal{R}_u - \hat{\mathcal{R}}_u)(t)\right\|^2 + \|\Upsilon(t)\|^2\right)\mathrm{d}t. \qquad (4.37)\end{aligned}$$

This estimate combined with equation (4.4) yields by using the triangle inequality

$$\begin{aligned}\max_{t\in[0,T]}\|\nabla(U-u)(t)\|^2 &\lesssim \max_{t\in[0,t_1]}\|\nabla(U-u)(t)\|^2 + \max_{t\in[t_1,T]}\left\|\nabla(U-\hat{U})(t)\right\|^2 \\ &\quad + \max_{t\in[t_1,T]}\left\|\nabla(\hat{\mathcal{R}}_u - \hat{U})(t)\right\|^2 + \max_{t\in[t_1,T]}\left\|\nabla(\hat{\mathcal{R}}_u - u)(t)\right\|^2 \\ &\lesssim \eta^0\left(H_0^1\right) + \max_{t\in[0,T]}\left\|\nabla(U-\hat{U})(t)\right\|^2 + \max_{t\in[0,T]}\left\|\nabla(\hat{\mathcal{R}}_u - \hat{U})(t)\right\|^2 \\ &\quad + \int_0^T\left(\left\|\partial_t(\hat{\mathcal{R}}_u - \hat{U})(t)\right\|^2 + \left\|\Delta(\mathcal{R}_u - \hat{\mathcal{R}}_u)(t)\right\|^2 + \|\Upsilon(t)\|^2\right)\mathrm{d}t.\end{aligned}$$

Then equation (3.10) is shown by equations (4.25), (4.26), (4.29), (4.33) and (4.34). It remains to prove equation (3.11).

The inf–sup condition and equation (4.37) yield for $t \in I_n$

$$\begin{aligned}\left\|(\mathcal{R}_p - p)(t)\right\| &\le \sup_{v\in X,\|\nabla v\|=1}\left\langle\nabla(\mathcal{R}_p - p)(t), v\right\rangle \\ &= \sup_{v\in X,\|\nabla v\|=1}\left\langle -\partial_t(\hat{\mathcal{R}}_u - u)(t) + \Delta(\mathcal{R}_u - u)(t) + \partial_t(\hat{\mathcal{R}}_u - \hat{U})(t) + \Upsilon(t), v\right\rangle \\ &\lesssim \left\|\partial_t(\hat{\mathcal{R}}_u - u)(t)\right\|_{-1} + \|\nabla(\mathcal{R}_u - u)(t)\| + \left\|\partial_t(\hat{\mathcal{R}}_u - \hat{U})(t)\right\|_{-1} + \|\Upsilon(t)\|_{-1} \\ &\lesssim \left\|\partial_t(\hat{\mathcal{R}}_u - u)(t)\right\| + \|\nabla(\mathcal{R}_u - u)(t)\| + \left\|\partial_t(\hat{\mathcal{R}}_u - \hat{U})(t)\right\|_{-1} + \|\Upsilon(t)\|.\end{aligned}$$

Upon squaring this equation, integrating from $t = t_1$ to T and using equations (4.36) and (4.37), one gets

$$\int_{t_1}^{T}\left\|(\mathcal{R}_p - p)(t)\right\|^2 \lesssim \left\|\nabla(\hat{\mathcal{R}}_u - u)(t_1)\right\|^2 + \int_{t_1}^{T}\left(\left\|\partial_t(\hat{\mathcal{R}}_u - \hat{U})(t)\right\|^2 + \left\|\Delta(\mathcal{R}_u - \hat{\mathcal{R}}_u)(t)\right\|^2 + \|\Upsilon(t)\|^2\right)\mathrm{d}t. \qquad (4.38)$$

Next, one gets by the triangle inequality, equation (4.7) and equation (4.38)

$$
\begin{aligned}
\int_0^T \|(P-p)(t)\|^2 &\lesssim \int_0^{t_1} \|(P-p)(t)\|^2 + \int_{t_1}^T \left\|(P-\mathcal{R}_p)(t)\right\|^2 + \int_{t_1}^T \left\|(\mathcal{R}_p-p)(t)\right\|^2 \\
&\lesssim \eta^0\left(H_0^1\right) + \int_0^T \left\|(P-\mathcal{R}_p)(t)\right\|^2 \\
&\quad + \int_0^T \left(\left\|\partial_t(\hat{\mathcal{R}}_u - \hat{U})(t)\right\|^2 + \left\|\Delta(\mathcal{R}_u - \hat{\mathcal{R}}_u)(t)\right\|^2 + \|\Upsilon(t)\|^2 \right) \mathrm{d}t.
\end{aligned}
$$

Then using equations (4.26), (4.27), (4.29) and (4.33) to complete the proof. □

5. Proof of optimality, Theorem 3.6

In this section we show that the time error estimators $\mathcal{E}_1$–$\mathcal{E}_7$ in Theorem 3.4 are bounded independently of the time step τ and the mesh size h, and in turn the *a posteriori* estimates yield the same asymptotic convergence rates in time as the *a priori* estimates. Due to the complexity we assume constant meshes. The proof is subdivided into four steps. First we bound the estimators by difference quotients of the initial values in Lemma 5.1. Hence, for proving optimality of the estimators, corresponding difference quotients of the initial values have to be bounded. To the best of our knowledge such a result is available only for the semidiscrete case, i.e. continuous in space, see for instance the study by Akrivis & Chatzipantelidis (2010) and Bänsch & Brenner (2016). Let u be the exact solution and w be the corresponding semidiscrete solution. In the study by Akrivis & Chatzipantelidis (2010), error estimates for semidiscrete parabolic problems using BDF2 for time discretization are considered, and the resulting error estimator is stable if the third-order difference quotient $\bar{\partial}^3 w^3$ of initial values can be bounded. This is easily done by inserting the exact solution u and estimating

$$
\|\bar{\partial}^3 w^3\| \leq \left\|\bar{\partial}^3\left(w^3 - u(t_3)\right)\right\| + \left\|\bar{\partial}^3 u(t_3)\right\| \leq \frac{1}{\tau^3}\sum_{i=0}^{3}\left\|w^i - u(t_i)\right\| + \left\|\bar{\partial}^3 u(t_3)\right\|.
$$

Since the initial errors $\|w^i - u(t_i)\|$ are of order $\mathcal{O}(\tau^3)$, the third-order difference quotient $\bar{\partial}^3 w^3$ is bounded independently of τ. However, in our case of the fully discrete solution U^n of equation (2.9) the initial errors $U^i - u(t_i)$, $i = 0, 1, 2, 3$, fulfill

$$
\left\|U^i - u(t_i)\right\| \leq C(\tau^3 + h^{k+1})
$$

and applying the above technique directly would result in a condition on the *minimal* size of the time step depending on h.

Therefore, we introduce as a second step in Lemma 5.2 the corresponding time-dependent semidiscrete problem, and estimate the error between a fully and semidiscrete solution. As a result we are able to bound the initial errors by corresponding time derivatives of this semidiscrete solution. In the third step the discrete regularity results of Appendix B are used in Lemma 5.3. This allows us to bound the semidiscrete values by interpolation estimates, and regularity of the exact solution and given data. Finally, in the fourth step, we combine these results to prove Theorem 3.6.

Let us first state two stability estimates.

LEMMA 5.1 (Stability). Assume $X_n = X_h \ \forall \ n = 0, \ldots, N$. Then the scheme defined in equation (2.9) fulfills the following stability results for all $i \geq 0$:

$$\max_{2+i\leq n\leq N} \|\bar{\partial}^i U^n\|^2 + \sum_{n=2+i}^{N} \tau\|\nabla\bar{\partial}^i U^n\|^2 \lesssim \sum_{n=2+i}^{N} \tau\|\bar{\partial}^i F^n\|^2_{J_h^*} + \|\bar{\partial}^i U^{1+i}\|^2 + \|\bar{\partial}^i U^i\|^2, \tag{5.1}$$

$$\max_{2+i\leq n\leq N} \|\nabla\bar{\partial}^i U^n\|^2 + \sum_{n=2+i}^{N} \tau\|\bar{\partial}^{i+1} U^n\|^2 \lesssim \sum_{n=2+i}^{N} \tau\|\bar{\partial}^i F^n\|^2 + \|\nabla\bar{\partial}^i U^{1+i}\|^2 + \tau\|\bar{\partial}^{i+1} U^{i+1}\|^2. \tag{5.2}$$

Proof. Apply the time difference operator $\bar{\partial}^i$ to equation (2.9) to obtain for $n \geq i+2$, $i \geq 0$,

$$\bar{\partial}^i\bar{\partial}^B U^n - \varDelta_h\bar{\partial}^i U^n + B_h^\top\bar{\partial}^i P^n = \bar{\partial}^i F^n. \tag{5.3}$$

To prove equation (5.1), test equation (5.3) with $4\tau\bar{\partial}^i U^n$. The first term on the left-hand side results in

$$\begin{aligned}
4\tau\langle\bar{\partial}^i\bar{\partial}^B U^n, \bar{\partial}^i U^n\rangle &= 2\langle 3\bar{\partial}^i U^n - 4\bar{\partial}^i U^{n-1} + \bar{\partial}^i U^{n-2}, \bar{\partial}^i U^n\rangle \\
&= \|\bar{\partial}^i U^n\|^2 - \|\bar{\partial}^i U^{n-1}\|^2 + \tau^4\|\bar{\partial}^{i+2} U^n\|^2 + \|2\bar{\partial}^i U^n - \bar{\partial}^i U^{n-1}\|^2 \\
&\quad - \|2\bar{\partial}^i U^{n-1} - \bar{\partial}^i U^{n-2}\|^2,
\end{aligned}$$

where we have used identity (A.8) in the last step. Use Young's inequality on the right-hand side to get

$$\begin{aligned}
&\|\bar{\partial}^i U^n\|^2 - \|\bar{\partial}^i U^{n-1}\|^2 + \tau^4\|\bar{\partial}^{i+2} U^n\|^2 + \|2\bar{\partial}^i U^n - \bar{\partial}^i U^{n-1}\|^2 - \|2\bar{\partial}^i U^{n-1} - \bar{\partial}^i U^{n-2}\|^2 \\
&\quad + 4\tau\|\nabla\bar{\partial}^i U^n\|^2 \leq C\tau\|\bar{\partial}^i F^n\|^2_{J_h^*} + \tau\|\nabla\bar{\partial}^i U^n\|^2.
\end{aligned}$$

Summing up from $n = 2 + i$ to N yields equation (5.1).

For proving equation (5.2) we test equation (5.3) with $4\tau\bar{\partial}^{i+1} U^n$. By using identity equation (4.20), we obtain for the left-hand side

$$\begin{aligned}
4\tau\langle\bar{\partial}^i\bar{\partial}^B U^n, \bar{\partial}^{i+1} U^n\rangle &= 4\tau\left\langle\frac{\tau}{2}\bar{\partial}^{i+2} U^n + \bar{\partial}^{i+1} U^n, \bar{\partial}^{i+1} U^n\right\rangle \\
&= \tau\left(\|\bar{\partial}^{i+1} U^n\|^2 - \|\bar{\partial}^{i+1} U^{n-1}\|^2 + \tau^2\|\bar{\partial}^{i+2} U^n\|^2 + 4\|\bar{\partial}^{i+1} U^n\|^2\right).
\end{aligned}$$

Use again equation (4.11) for the second term on the left-hand side and Young's inequality for the right-hand side of equation (5.3) to obtain the error inequality

$$\tau\left(\|\bar{\partial}^{i+1} U^n\|^2 - \|\bar{\partial}^{i+1} U^{n-1}\|^2\right) + \tau\|\bar{\partial}^{i+1} U^n\|^2 + \|\nabla\bar{\partial}^i U^n\|^2 - \|\nabla\bar{\partial}^i U^{n-1}\|^2 \lesssim \tau\|\bar{\partial}^i F^n\|^2.$$

Summing up from $n = 2 + i$ to N yields equation (5.2). □

In the next step we present *a priori* estimates of the corresponding semidiscrete solution of equations (2.9) and (2.10).

LEMMA 5.2 Let $(W(t), Q(t)) \in X_h \times Y_h$ be solution of the semidiscrete scheme

$$\partial_t W(t) - \Delta_h W(t) + B_h^\top Q(t) = \Pi_h f(t), \qquad B_h W = 0, \qquad W(0) = U^0. \tag{5.4}$$

Then for $i = 1, 2, 3$ one can estimate

$$\left\| U^i - W(t_i) \right\| + \tau^{1/2} \left\| \nabla \left(U^i - W(t_i) \right) \right\| \lesssim \tau^{5/2} \left(\left\| \partial_t^{(3)} W(t) \right\|_{L^2((0,3\tau);L^2(\Omega)^d)} + \left\| \partial_t^{(2)} f(t) \right\|_{L^2((0,\tau);L^2(\Omega)^d)} \right).$$

Proof. Set $\Theta^i := U^i - W(t_i)$ and $\Pi_{J_h} : X \to J_h$ the L^2 projection on J_h. The error equation in J_h^* corresponding to equation (2.10) reads

$$\Theta^1 - \frac{\tau}{2} \Delta_h \Theta^1 = \tau R^1,$$

where $R^1 := (\partial_t - \bar{\partial}) W(t_1) - \Delta_h (W(t_1) - W(t_{1/2})) + \Pi_{J_h} (f(t_1) - f(t_{1/2}))$ fulfilling

$$\|R^1\| \lesssim \tau^{3/2} \left(\left\| \partial_t^{(3)} W(t) \right\|_{L^2((0,\tau);L^2(\Omega)^d)} + \left\| \partial_t^{(2)} f(t) \right\|_{L^2((0,\tau);L^2(\Omega)^d)} \right).$$

Test this equation by Θ^1 to get

$$\|\Theta^1\|^2 + \tau \left\| \nabla e_1^1 \right\|^2 \lesssim \tau^2 \|R^1\|^2.$$

To bound the error Θ^2, set $R^2 := (\partial_t - \bar{\partial}^B) W(t_2)$ fulfilling $\|R^2\| \lesssim \tau^{3/2} \|\partial_t^{(3)} W(t)\|_{L^2((\tau,2\tau);L^2(\Omega)^d)}$. This gives

$$3\Theta^2 - 2\tau \Delta_h \Theta^2 = 4\Theta^1 - 2\tau R^2.$$

The same procedure as for Θ^1 then yields the estimate for Θ^2 and also for Θ^3. □

The above lemma bounds the initial errors $U^i - W(t_i)$ in terms of the third-order time derivative of the semidiscrete velocity W. In the next lemma the results from Appendix B are used to further estimate this term.

In what follows we have to make use of the inverse estimate

$$\|\nabla V\| \lesssim \frac{1}{h} \|V\| \qquad \forall\, V \in X_h. \tag{5.5}$$

This is why a condition of quasi-uniform meshes is required. In view of adaptive meshes we would like to drop this condition and use a local variant of equation (5.5). However, the proof of the next lemma needs a global inverse estimate due to the global definition of the operator $\mathcal{A}_h$ and the Stokes projection $\mathcal{S}_n^u$.

LEMMA 5.3 Assume that the exact solution u of equation (2.1) is sufficiently regular. Let W be as in Lemma 5.2. Then

$$\left\|\partial_t^3 W\right\|_{L^2((0,3\tau);L^2(\Omega)^d)} \lesssim \left\|\partial_t^3 u - \mathcal{S}_n^u\left(\partial_t^3 u, \partial_t^3 p\right)\right\|_{L^2((0,3\tau);L^2(\Omega)^d)} + \left\|\partial_t^3 u\right\|_{L^2((0,3\tau);L^2(\Omega)^d)}$$
$$+ \frac{1}{h^3}\left\|u_0 - \mathcal{S}_n^u\left(u_0, p_0\right)\right\| + \frac{1}{h}\left\|\partial_t u_0 - \mathcal{S}_n^u(\partial_t u_0, \partial_t p_0)\right\|.$$

Proof. To simplify the notation we set

$$\mathcal{S}_n^u(u,p)(t) := \mathcal{S}_n^u\left(u(t), p(t)\right) \qquad \text{and} \qquad \partial_t^l \mathcal{S}_n^u(u,p)(t) := \mathcal{S}_n^u\left(\partial_t^l u(t), \partial_t^l p(t)\right).$$

The triangle inequality yields

$$\left\|\partial_t^3 W\right\| \le \left\|\partial_t^3 (W - \mathcal{S}_n^u(u,p))\right\| + \left\|\partial_t^3 (u - \mathcal{S}_n^u(u,p))\right\| + \left\|\partial_t^3 u\right\|.$$

It remains to bound the term $\|\partial_t^3(W - \mathcal{S}_n^u(u,p))\|_{L^2((0,3\tau);L^2(\Omega)^d)}$. To this end we use the results on discrete regularity provided in Appendix B. Subtract the continuous system (2.1) from the semidiscrete system (5.4) and combine the resulting equation with the Stokes projection (2.4) to get

$$\partial_t\left(W - \mathcal{S}_n^u(u,p)\right)(t) + \mathcal{A}_h\left(W - \mathcal{S}_n^u(u,p)\right)(t) = \Pi_{J_h}\partial_t\left(u(t) - \mathcal{S}_n^u(u,p)\right)(t),$$

where Π_{J_h} is the L^2 projection onto J_h and $\mathcal{A}_h$ the discrete Stokes operator defined in equation (B.3). For the initial error one has

$$W(0) - \mathcal{S}_n^u(u_0,p_0) = U^0 - \mathcal{S}_n^u(u_0,p_0) = \mathcal{S}_n^u(u_0,p_0) - \mathcal{S}_n^u(u_0,p_0) = 0.$$

We use equation (B.7) of Theorem B.2 in the case $k = 2$ applied to

$$Z(t) = W(t) - \mathcal{S}_n^u(u,p)(t),\ G = \Pi_{J_h}\partial_t\left(u - \mathcal{S}_n^u(u,p)\right)(t) \ \text{ and } Z(0) = 0.$$

Hence we get

$$\left\|\partial_t^3\left(W - \mathcal{S}_n^u(u,p)\right)\right\|_{L^2((0,3\tau);L^2(\Omega)^d)} \lesssim \left\|\partial_t^3\left(u - \mathcal{S}_n^u(u,p)\right)\right\|_{L^2((0,3\tau);L^2(\Omega)^d)}$$
$$+ \left\|\nabla \Pi_{J_h}\left(u_0 - \mathcal{S}_n^u(u_0,p_0)\right)\right\| + \left\|\nabla \Pi_{J_h}\partial_t\left(u_0 - \mathcal{S}_n^u(u_0,p_0)\right)\right\|$$
$$+ \left\|\nabla \mathcal{A}_h \Pi_{J_h}\left(u_0 - \mathcal{S}_n^u(u_0,p_0)\right)\right\|.$$

The second and third term on the right-hand side can be bounded with the help of the inverse estimate (5.5) and the stability of the L^2 projection:

$$\left\|\nabla \Pi_{J_h}\left(u_0-\mathcal{S}_n^u(u_0,p_0)\right)\right\| \lesssim \frac{1}{h}\left\|\Pi_{J_h}\left(u_0-\mathcal{S}_n^u(u_0,p_0)\right)\right\| \leq \frac{1}{h}\left\|\left(u_0-\mathcal{S}_n^u(u_0,p_0)\right)\right\|,$$

$$\left\|\nabla \Pi_{J_h}\partial_t\left(u_0-\mathcal{S}_n^u(u_0,p_0)\right)\right\| \lesssim \frac{1}{h}\left\|\partial_t\left(u_0-\mathcal{S}_n^u(u_0,p_0)\right)\right\|.$$

For the last term on the right-hand side we apply the inverse estimate (5.5) once again and use the definition of the discrete Stokes operator, cf. (B.12):

$$\begin{aligned}\left\|\nabla \mathcal{A}_h\Pi_{J_h}\left(u_0-\mathcal{S}_n^u(u_0,p_0)\right)\right\|^2 &\lesssim \frac{1}{h^2}\left\|\mathcal{A}_h\Pi_{J_h}\left(u_0-\mathcal{S}_n^u(u_0,p_0)\right)\right\|^2\\ &= \frac{1}{h^2}\left\langle\nabla \Pi_{J_h}\left(u_0-\mathcal{S}_n^u(u_0,p_0)\right), \nabla \mathcal{A}_h\Pi_{J_h}\left(u_0-\mathcal{S}_n^u(u_0,p_0)\right)\right\rangle\\ &\leq \frac{1}{h^3}\left\|\left(u_0-\mathcal{S}_n^u(u_0,p_0)\right)\right\|\left\|\nabla \mathcal{A}_h\Pi_{J_h}\left(u_0-\mathcal{S}_n^u(u_0,p_0)\right)\right\|.\end{aligned}$$

Dividing the last equation by $\|\nabla \mathcal{A}_h\Pi_{J_h}(u_0-\mathcal{S}_n^u(u_0,p_0))\|$ finally together with the above estimates concludes the lemma. $\square$

Now we have all the ingredients to prove Theorem 3.6.

Proof. **(Theorem 3.6)** Use equation (5.1) for $i=2$ to bound the estimators $\mathcal{E}_1$ and $\mathcal{E}_3$

$$\mathcal{E}_1 \lesssim \mathcal{E}_3 \lesssim \sum_{n=4}^{N}\tau\|\bar\partial^2 F^n\|_{J_h^*}^2 + \|\bar\partial^2 U^2\|^2 + \|\bar\partial^2 U^3\|^2$$

and use equation (5.2) for $i=2$ to bound the estimators $\mathcal{E}_2$ and $\mathcal{E}_6$–$\mathcal{E}_7$:

$$\mathcal{E}_2+\mathcal{E}_7 \leq \mathcal{E}_6+\mathcal{E}_7 \lesssim \sum_{n=4}^{N}\tau\|\bar\partial^2 F^n\|^2 + \|\nabla\bar\partial^2 U^2\|^2 + \|\nabla\bar\partial^2 U^3\|^2 + \tau\|\bar\partial^3 U^2\| + \tau\|\bar\partial^3 U^3\|.$$

The estimators $\mathcal{E}_4$ and $\mathcal{E}_5$ can easily be handled by the triangle inequality:

$$\mathcal{E}_4 \leq \mathcal{E}_5 \lesssim \sum_{n=2}^{N}\tau\left\|\bar\partial^2 f(t_n)\right\|^2 + \tau\|\bar\partial^3 U^1\|^2 + \mathcal{E}_6.$$

It is clear that the term $\sum_{n=4}^{N}\tau\|\bar\partial^2 F^n\|^2$ is bounded by regularity of f. It remains to bound the quantities $\|\bar\partial^2 U^2\|^2$, $\|\bar\partial^2 U^3\|^2$, $\|\nabla\bar\partial^2 U^2\|^2$, $\|\nabla\bar\partial^2 U^3\|^2$ and $\tau\|\bar\partial^3 U^3\|$. Note that the terms $\tau\|\bar\partial^3 U^1\|$ and $\tau\|\bar\partial^3 U^2\|$ can be absorbed into the data approximation error estimators $\mathcal{I}_1$ and $\mathcal{I}_2$ and in lower-order difference quotients.

Let $W(t_i)$ be the semidiscrete solution of equation (5.4). Then for instance the term $\tau\|\bar\partial^3 U^3\|^2$ can be split by the triangle inequality:

$$\sqrt{\tau}\|\bar\partial^3 U^3\| \le \sqrt{\tau}\left\|\bar\partial^3\big(U^3 - W(t_3)\big)\right\| + \sqrt{\tau}\left\|\bar\partial^3 W(t_3)\right\|.$$

The first term on the right-hand side can be estimated with the help of Lemma 5.2:

$$\begin{aligned}\sqrt{\tau}\left\|\bar\partial^3\left(U^3 - W(t_3)\right)\right\| &\lesssim \tau^{-5/2}\left(\left\|U^1 - W(t_1)\right\| + \left\|U^2 - W(t_2)\right\| + \left\|U^3 - W(t_3)\right\|\right)\\ &\lesssim \left\|\partial_t^{(3)} W(t)\right\|_{L^2((0,3\tau);L^2(\Omega)^d)} + \left\|\partial_t^{(2)} f(t)\right\|_{L^2((0,\tau);L^2(\Omega)^d)}.\end{aligned}$$

It is easily shown that the second term fulfills $\sqrt{\tau}\|\bar\partial^3 W(t_3)\| \lesssim \|\partial_t^{(3)} W(t)\|_{L^2((0,3\tau);L^2(\Omega)^d)}$. Collect these results and use Lemma 5.3 to get

$$\begin{aligned}\sqrt{\tau}\|\bar\partial^3 U^3\| \lesssim{}& \left\|\partial_t^{(3)} u - \mathcal{S}_n^u\left(\partial_t^{(3)} u, \partial_t^{(3)} p\right)\right\|_{L^2((0,3\tau);L^2(\Omega)^d)} + \left\|\partial_t^{(3)} u\right\|_{L^2((0,3\tau);L^2(\Omega)^d)}\\ &+ \frac{1}{h^3}\left\|u_0 - \mathcal{S}_n^u(u_0, p_0)\right\| + \frac{1}{h}\left\|\partial_t u_0 - \mathcal{S}_n^u(\partial_t u_0, \partial_t p_0)\right\| + \left\|\partial_t^{(2)} f(t)\right\|_{L^2((0,\tau);L^2(\Omega)^d)}.\end{aligned}$$

Finally use Lemma 2.2 to bound the approximation errors $\|u(\cdot) - \mathcal{S}_n^u(u(\cdot), p(\cdot))\|$ to conclude the proof of the error bound of $\sqrt{\tau}\|\bar\partial^3 U^3\|$. Note that at this point we need at least order $k = 2$ elements for the velocity space X_h to bound the term $\frac{1}{h^3}\|u_0 - \mathcal{S}_n^u(u_0, p_0)\|$.

The other terms $\|\bar\partial^2 U^2\|^2$, $\|\bar\partial^2 U^3\|^2$, $\|\nabla\bar\partial^2 U^2\|^2$ and $\|\nabla\bar\partial^2 U^3\|^2$ can be treated in a similar way. □

6. Computational results

In this section computational examples are presented to numerically verify our analysis. The numerical simulations are calculated by NAVIER (Bänsch, 1998) and FENICS (Alnæs *et al.*, 2015).

6.1 *Asymptotic behavior of the temporal estimators*

In this first example we show the asymptotic behavior of the time and Stokes reconstruction error estimators as stated in Theorem 3.4. We chose $\Omega =]0, 1[^2$ and

$$\begin{aligned}u(t,x) &= \Big(\pi\sin(t)\sin(2\pi x_2)\sin^2(\pi x_1),\ -\pi\sin(t)\sin(2\pi x_1)\sin^2(\pi x_2)\Big)^\top,\\ p(t,x) &= \sin(t)\cos(\pi x_1)\sin(\pi x_2),\end{aligned}$$

with a right-hand side f, such that the above pair (u, p) is a solution to equation (2.1). The time interval was [0, 1]. As we are only interested here in the time estimators, we use a constant finite element $\mathcal{P}_2 - \mathcal{P}_1$ Taylor–Hood discretization to compute the numerical solutions and a fine grid consisting of 16641 vertices. Consequently, the error due to space discretization was very small.

For nonchanging meshes the estimators of Definition 3.3 read

$$\mathcal{E}_1 := \sum_{n=2}^{N} \tau \|\bar\partial^2 U^n\|^2, \qquad \mathcal{E}_4 := \sum_{n=2}^{N} \tau \left\|\bar\partial^2 \left(f(t_n) - \bar\partial^B U^n\right)\right\|_{J^*}^2, \qquad \mathcal{E}_7 := \max_{2\le n\le N} \|\nabla \bar\partial^2 U^n\|^2,$$
$$\mathcal{E}_2 := \sum_{n=3}^{N} \tau \|\bar\partial^3 U^n\|_{J^*}^2, \qquad \mathcal{E}_5 := \sum_{n=2}^{N} \tau \left\|\bar\partial^2 \left(f(t_n) - \bar\partial^B U^n\right)\right\|^2$$
$$\mathcal{E}_3 := \max_{2\le n\le N} \|\bar\partial^2 U^n\|^2, \qquad \mathcal{E}_6 := \sum_{n=3}^{N} \tau \|\bar\partial^3 U^n\|^2,$$

and we set

$$\Theta_1^2 := \tau^4(\mathcal{E}_1 + \mathcal{E}_2), \qquad \Theta_2^2 := \tau^4(\mathcal{E}_2 + \mathcal{E}_3 + \mathcal{E}_4), \qquad \Theta_3^2 := \tau^4(\mathcal{E}_2 + \mathcal{E}_4),$$
$$\Theta_4^2 := \tau^4(\mathcal{E}_5 + \mathcal{E}_6 + \mathcal{E}_7), \qquad \Theta_5^2 := \tau^4(\mathcal{E}_5 + \mathcal{E}_6).$$

With this notation Theorem 3.4 reads

$$\begin{aligned}
\|U-u\|_{L^2(L^2)}^2 &\lesssim \Theta_1^2 + \eta^0(J^*) + \Lambda_1 + \mathcal{F}_1 + \mathcal{I}_1,\\
\|U-u\|_{L^\infty(L^2)}^2 &\lesssim \Theta_2^2 + \eta^0(L^2) + \Lambda_2 + \Lambda_3 + \mathcal{F}_1 + \mathcal{I}_1,\\
\|\nabla(U-u)\|_{L^2(L^2)}^2 &\lesssim \Theta_3^2 + \eta^0(L^2) + \Lambda_3 + \Lambda_4 + \mathcal{F}_1 + \mathcal{I}_1,\\
\|\nabla(U-u)(t)\|_{L^\infty(L^2)}^2 &\lesssim \Theta_4^2 + \eta^0\left(H_0^1\right) + \Lambda_5 + \Lambda_6 + \mathcal{F}_2 + \mathcal{I}_2,\\
\|P-p\|_{L^2(L^2)}^2 &\lesssim \Theta_5^2 + \eta^0\left(H_0^1\right) + \Lambda_6 + \Lambda_7 + \mathcal{F}_2 + \mathcal{I}_2.
\end{aligned}$$

For simplicity only the Θ- terms are presented, since the other terms are less interesting in the present context. Figure 1 shows the asymptotic orders of the estimators for velocity and pressure. As can be seen, the order of estimators and errors are in perfect agreement as long as the time error is dominant. For smaller time step sizes the curves for the errors eventually start to stagnate, since then the spatial discretization error becomes dominant.

6.2 *Robustness of the error estimators*

In this second computation we want to test the estimator of Definition 3.3 for robustness in a more complex example whose solution is nonseparable with respect to the time and space variable. We chose the time interval [0, 1], $\Omega =]0, 1[^2$ and

$$u(t,x) = \left(x_2 \sin\left(2\pi\ t\left(x_1^2 + x_2^2\right)\right),\ -x_1 \sin\left(2\pi\ t\left(x_1^2 + x_2^2\right)\right)\right)^\top,$$

$$p(t,x) = x_1 x_2 \sin\Big(2\pi \; t \left(x_1^2 + x_2^2\right)\Big),$$

with a right-hand side f, such that (u, p) is a solution to equation (2.1).

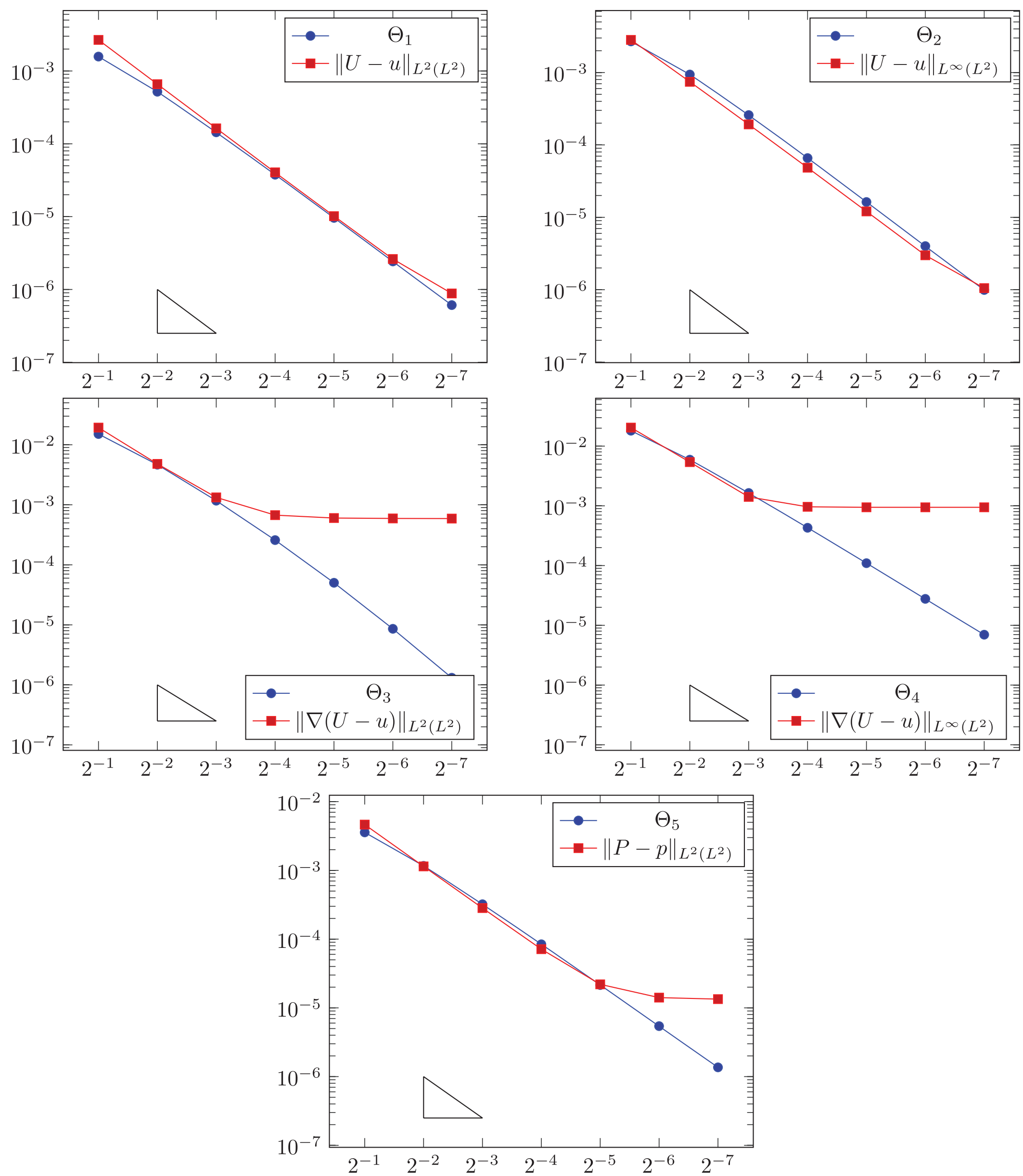

FIG. 1. BDF2: errors and (scaled) estimators vs. time step size. Triangles have slope 2.0.

TABLE 1 *Temporal and space estimators for different time step sizes τ on a constant mesh with $h = 0.011$*

τ	$\mathcal{E}_1^{1/2}$	$\mathcal{E}_2^{1/2}$	$\mathcal{E}_3^{1/2}$	$\mathcal{E}_4^{1/2}$	$\mathcal{E}_5^{1/2}$	$\mathcal{E}_6^{1/2}$	$\mathcal{E}_7^{1/2}$	$\Lambda_1^{1/2}$	$\Lambda_2^{1/2}$	$\Lambda_4^{1/2}$	$\Lambda_5^{1/2}$
0.2	6.49	0.15	8.88	1.68	263	32	54	0.000251	0.000288	0.0227	0.0261
0.1	7.47	0.17	9.60	2.18	374	35	58	0.000244	0.000305	0.0221	0.0276
0.05	7.28	0.16	9.83	2.01	331	34	59	0.000224	0.000313	0.0203	0.0283
0.025	7.35	0.16	9.88	2.05	340	33	59	0.000223	0.000316	0.0202	0.0286
0.0125	7.44	0.16	9.90	2.10	353	33	59	0.000225	0.000316	0.0204	0.0286
0.00625	7.42	0.16	9.90	2.10	351	33	59	0.000224	0.000317	0.0203	0.0287

For the spatial estimators we follow Verführt (2013) and set

$$\eta_u^n\left(U^n; H_0^1\right) := \sum_{T\in\mathcal{T}_n}\left(h_T^2\left\|\Delta_n U^n + F^n - B_n^\top P^n\right\|_T^2 + \|\mathrm{div}\, U^n\|^2\right) + \sum_{e\in\mathcal{E}_\mathcal{T}} h_e\left\|\left[n_e\cdot(\nabla U^n - P^n I)\right]\right\|_e^2,$$

$$\eta_u^n(U^n; L^2) := \sum_{T\in\mathcal{T}_n}\left(h_T^4\left\|\Delta_n U^n + F^n - B_n^\top P^n\right\|_T^2 + h_T^2\|\mathrm{div}\, U^n\|^2\right) + \sum_{e\in\mathcal{E}_\mathcal{T}} h_e^3\left\|\left[n_e\cdot(\nabla U^n - P^n I)\right]\right\|_e^2.$$

To shorten the presentation we present only the spatial estimatorsΛ_1, Λ_2, Λ_4 and Λ_5.

Table 1 shows the values of the estimators on a constant finite element $\mathcal{P}_2 - \mathcal{P}_1$ Taylor–Hood discretization with 16641 vertices ($h = 0.011$). As claimed in Theorem 3.6 the error estimators $\mathcal{E}_1$–$\mathcal{E}_7$ are independent of the time step τ. Also the space estimators do not depend on the time step sizeTable 2 presents the values of the estimators for refining meshes with halving mesh size h and constant time step $\tau = 0.02$. The time estimators $\mathcal{E}_1$–$\mathcal{E}_7$ are independent of the mesh size h. This shows the robustness of the temporal estimators with respect to the spatial discretization. The space estimators behave as expected. We have quadratic convergence for the L^2 estimators $\Lambda_1^{1/2}$ and $\Lambda_2^{1/2}$ and linear convergence for the H^1 estimators $\Lambda_4^{1/2}$ and $\Lambda_5^{1/2}$.

In our next computation we want to show the robustness of our time error estimator with respect to the end time T. To this end we consider a solution of equation (2.1), which is periodic in time:

$$u(t,x) = \Big(\sin(t)\sin(2\pi(x_1+x_2)),\ -\sin(t)\sin(2\pi(x_1+x_2))\Big)^\top, \qquad p(t,x) = 0,$$

TABLE 2 *Temporal and space estimators for different mesh sizes h and fixed $\tau = 0.02$*

h	$\mathcal{E}_1^{1/2}$	$\mathcal{E}_2^{1/2}$	$\mathcal{E}_3^{1/2}$	$\mathcal{E}_4^{1/2}$	$\mathcal{E}_5^{1/2}$	$\mathcal{E}_6^{1/2}$	$\mathcal{E}_7^{1/2}$	$\Lambda_1^{1/2}$	$\Lambda_2^{1/2}$	$\Lambda_4^{1/2}$	$\Lambda_5^{1/2}$
0.3535	7.39	0.16	9.90	2.03	337	34	59	0.85383	1.90111	2.415	5.377
0.1767	7.37	0.16	9.89	2.06	342	33	59	0.12283	0.26697	0.694	1.510
0.0883	7.37	0.16	9.89	2.06	342	33	59	0.01962	0.03554	0.222	0.402
0.0441	7.37	0.16	9.89	2.06	342	33	59	0.00393	0.00541	0.089	0.122
0.0220	7.37	0.16	9.89	2.06	342	33	59	0.00091	0.00126	0.041	0.057
0.0110	7.37	0.16	9.89	2.06	342	33	59	0.00022	0.00031	0.020	0.028

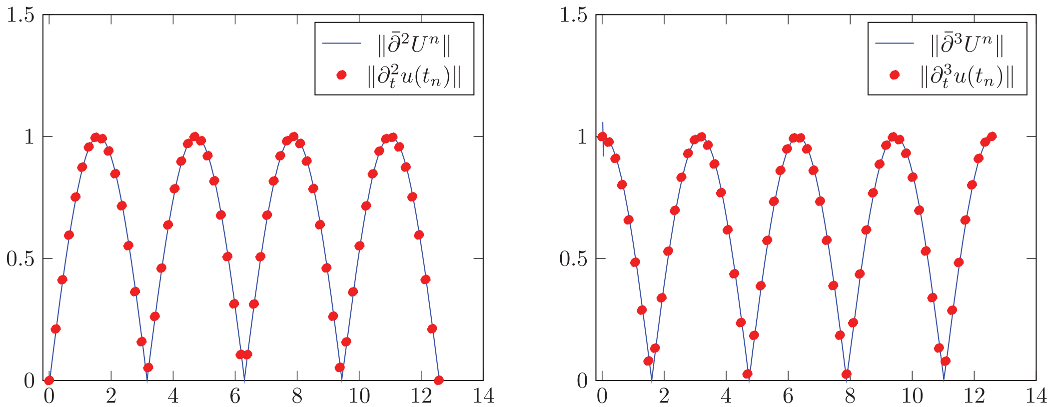

FIG. 2. Time versus local contributions $\|\bar{\partial}^2 U^n\|$ (left), $\|\bar{\partial}^3 U^n\|$ (right) of the time estimator and $\|\partial_t^2 u(t_n)\|$ (left), $\|\partial_t^3 u(t_n)\|$ (right) of the exact solution for end time $T = 4\pi$.

with a corresponding right-hand side f in the time interval $[0, 4\pi]$ on $\Omega =]0, 1[^2$. We set $\tau = 0.05$ and use the same spatial discretization as in the example before.

In Fig. 2 we compare the local contributions $\|\bar{\partial}^2 U^n\|$ and $\|\bar{\partial}^3 U^n\|$ of the time estimators with the time derivatives $\|\partial_t^2 u(t_n)\|$ and $\|\partial_t^3 u(t_n)\|$ of the exact solution in view of a dependence on the final time T. As can be seen the difference quotients of the discrete solution are in perfect agreement with the continuous time derivatives, and we conclude the robustness of our estimators with respect to the end time T.

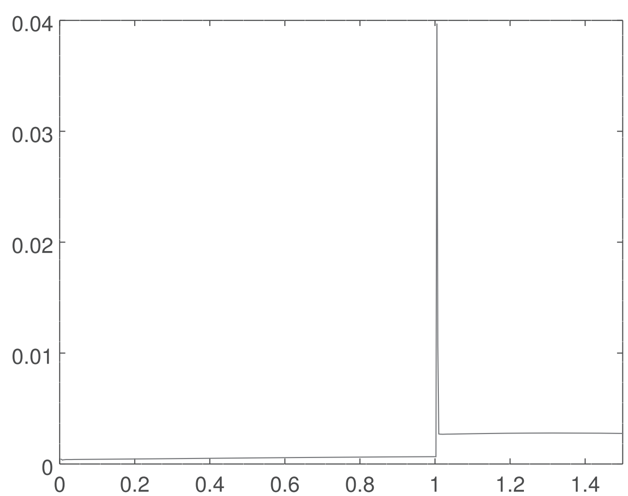

FIG. 3. Example with mesh change. L^2 pressure error versus time (initial mesh with 1024 elements, τ=1.25e-3).

TABLE 3 *Example with mesh change: 'additional' pressure error δE_p equation (6.2) for various meshes (nt = no. of initial elements) and time step sizes τ*

τ	2.$e-2$	1.$e-2$	5.$e-3$	2.5$e-3$	1.25$e-3$
nt= 256	8.3534$e-02$	1.5880$e-01$	3.0628$e-01$	5.9729$e-01$	1.1755$e+00$
nt=1024	5.6903$e-03$	1.0481$e-02$	2.0098$e-02$	3.9038$e-02$	7.6344$e-02$
nt=4096	5.9363$e-04$	8.1377$e-04$	1.3692$e-03$	2.5847$e-03$	5.0541$e-03$
nt=16384	1.2083$e-04$	1.2638$e-04$	1.4346$e-04$	1.9564$e-04$	3.3151$e-04$

6.3 *Mesh changes*

With this example we study the effect of mesh modification. The domain here is $\Omega =]-0.5, 0.5[^2$, $T = 1.5$ and the exact solution is

$$\begin{aligned} u(x,y,t) &= \pi \sin(t+0.25)(\sin(2\pi y)\sin(\pi x)^2, \sin(2\pi x)\sin(\pi y)^2)^T, \\ p(x,y,t) &= \sin(t+0.25)\cos(\pi x)\sin(\pi y). \end{aligned} \tag{6.1}$$

In order to focus on the mesh change, the setting is as follows. The computations are started with different uniform meshes with mesh size h_0 (corresponding to 256, 1024, 4096 and 16384 elements, respectively) and uniform time step sizes τ. The initial mesh is created by uniformly refining a given macro triangulation with the bisection method (Bänsch, 1991b). At the time instant $t_{jump} = 1.0$ the mesh is derefined, resulting in a coarsened mesh with mesh size 2 h_0. This scenario is as in the study by Besier & Wollner (2012). Since the pressure is the unknown most sensitive to mesh changes, we study the effect of mesh modification for this variable. In the study by Besier & Wollner (2012), it was shown that a derefinement of the mesh results in a jump of the pressure error of the order $1/\tau$. This behavior is confirmed by our experiments. In Fig. 3 the L^2 error of the pressure is shown as a function of time. The huge jump in the error at t_{jump} is clearly visible.

More quantitative information is given in Tables 3 and 4. In Table 3 the jumps δE_p in the pressure error are listed for various τ and h_0. Here, δE_p is defined as the 'additional' contribution to the pressure error by the mesh change:

$$\delta E_p := \|P^{n_0} - p(t_{\mathrm{jump}})\| - \|P^{n_0-1} - p(t_{\mathrm{jump}} - \tau)\|, \tag{6.2}$$

where n_0 corresponds to the time instant $t_{jump} = n_0\tau$. Notice that our theory does not provide error control for δE_p. This term is used in our experiments only as a tool to quantify the observed jump in the

TABLE 4 *Example with mesh change: estimator $\sqrt{\tilde{\xi}_B^n}$ for various meshes (nt = no. of initial elements) and time step sizes τ*

τ	2.$e-2$	1.$e-2$	5.$e-3$	2.5$e-3$	1.25$e-3$
nt=256	9.2077$e-01$	1.8415$e+00$	3.6831$e+00$	7.3662$e+00$	1.4732$e+01$
nt=1024	1.3887$e-01$	2.7773$e-01$	5.5546$e-01$	1.1109$e+00$	2.2218$e+00$
nt=4096	1.8840$e-02$	3.7681$e-02$	7.5361$e-02$	1.5072$e-01$	3.0144$e-01$
nt=16384	2.4194$e-03$	4.8389$e-03$	9.6779$e-03$	1.9356$e-02$	3.8712$e-02$

pressure. The comparison then to our coarsening estimator is indicative only in order to access whether or not our estimator can detect the jump in pressure.

Since we perform only one mesh change, a simplified mesh change estimator $\tilde{\xi}_B^n$ is used:

$$\tilde{\xi}_B^n := \|\tau^{-1}(\Pi_n - I)U^{n-1}\|^2.$$

In Table 4, the coarsening estimator terms $\sqrt{\tilde{\xi}_B^n}$ are reported. Note that because of orthogonality the relation

$$\tilde{\xi}_B^n = \|\tau^{-1}U^{n-1}\|^2 - \|\tau^{-1}\Pi^n U^{n-1}\|^2$$

holds, simplifying the computation of this term. Notice that in this example $\tilde{\xi}_B^n = 0$ for $n \neq n_0$. The projection $\Pi_n U^{n-1}$ can easily be computed without the need of (inexact) interpolation or quadrature, see Bänsch, 1991a for details.

From Table 3 one deduces the following behavior of δE_p. If the mesh is fine enough, the 'extra contribution' of the error is negligibly small. However, from a certain small time step size τ (depending on h_0), the extra contribution to the error behaves like $1/\tau$ (reading the rows of Table 3). In contrast, fixing a time step size and reading the columns, δE_p is decreasing.

A similar behavior can be observed for the coarsening estimator $\tilde{\xi}_B^n$. However, the $1/\tau$ increase is valid for all h_0. For fixed time step size, $\sqrt{\tilde{\xi}_B^n}$ decreases approximately like h_0^3, which is the order one would expect for $\mathcal{P}_2 - \mathcal{P}_1$ Taylor–Hood element.

We conclude that, in this experiment, our estimator has the right qualitative behavior and captures the significant increase of the error introduced by the mesh change.

7. Variable time steps

In this section we present a generalization of the analysis of this paper to the case of variable time steps. For the ease of presentation we restrict ourselves to a constant mesh and present only the estimate in the L2–L2 norm. The other *a posteriori* estimates can be obtained in an analogous manner. In contrast, it is an open problem whether the optimality proof in Section 5 carries over to the case of variable time step sizes. At least we expect an assumption on the ratio of successive time steps, see Emmrich (2005) and the literature cited there.

7.1. *A posteriori estimates*

For a partition $0 = t_0 < t_1 < \cdots < t_N = T$ define the time step $\tau_n := t_n - t_{n-1}$ and the time step ratio $\omega_n := \frac{\tau_n}{\tau_{n-1}}$. The finite difference operator $\bar{\partial}$ for variable time steps is defined recursively

$$\bar{\partial} U^n := \bar{\partial}^1 U^n := \frac{U^n - U^{n-1}}{\tau_n}, \qquad \bar{\partial}^k U^n := \frac{\bar{\partial}^{k-1} U^n - \bar{\partial}^{k-1} U^{n-1}}{\tau_n}, \; k = 2, \ldots, n.$$

The *BDF2 operator* is then given by

$$\bar{\partial}^B U^n := \frac{1}{\tau_n}\left(\frac{1+2\omega_n}{1+\omega_n}U^n - (1+\omega_n)U^{n-1} + \frac{\omega_n^2}{1+\omega_n}U^{n-2}\right) = \tau_n \frac{\omega_n}{1+\omega_n}\bar{\partial}^2 U^n + \bar{\partial} U^n.$$

Let X_h, Y_h be a mixed finite element approximation of X and Y. The discrete instationary Stokes system, cf. equations (2.9) and (2.10), with *variable time steps* reads: set $U^0 := u_0$, $P^0 := p_0$ and calculate U^1 from

$$\bar{\partial} U^1 - \Delta_h U^{1/2} + B_h^\top P^{1/2} = F^{1/2}, \qquad B_h U^1 = 0, \qquad U^1|_\Gamma = 0. \tag{7.1}$$

Then for $n \geq 2$ compute U^n from

$$\bar{\partial}^B U^n - \Delta_h U^n + B_h^\top P^n = F^n, \qquad B_h U^n = 0. \tag{7.2}$$

In the next theorem we present the estimators for variable time steps.

THEOREM 7.1 (L2–L2 *a posteriori* error estimates for variable time steps) The discrete solutions (U^n, P^n) of equations (7.1) and (7.2) satisfy the estimates

$$\|U - u\|^2_{L^2(L^2)} \lesssim \check{\mathcal{E}}_1 + \check{\mathcal{E}}_2 + \Lambda_1 + \mathcal{F}_1 + \ell_1,$$

where Λ_1, $\mathcal{F}_1$ and ℓ_1 are given in Definition 3.3 and

$$\check{\mathcal{E}}_1 = \sum_{n=2}^{N} \tau_n^5 \|\bar{\partial}^2 U^n\|^2 \qquad \text{and} \qquad \check{\mathcal{E}}_2 = \sum_{n=2}^{N} \frac{\tau_n^3}{\omega_n^2} \left\| \frac{\omega_n}{\omega_n + 1} \bar{\partial}^2 U^n - \frac{\omega_{n-1}}{\omega_{n-1}+1} \bar{\partial}^2 U^{n-1} \right\|^2_{J^*}. \tag{7.3}$$

Proof. For $n \geq 2$, $t \in I_n$, define the three-point reconstruction for variable time steps

$$\hat{U}(t) := U(t) + \frac{\omega_n}{\omega_n + 1}(t - t_n)(t - t_{n-1})\bar{\partial} U^n.$$

Its time derivative is

$$\partial_t \hat{U}(t) = \bar{\partial}^B U^n + \frac{2\omega_n}{\omega_n + 1}(t - t_n)\bar{\partial}^2 U^n.$$

With this definition, deduce the error equation like for equation (4.1). Now proceed exactly like in the proof of equation (3.7) of Theorem 3.4. □

Note that in case of constant time steps the above theorem reduces to Theorem 3.4 (in the case of constant meshes).

7.2. *Adaptive strategy*

Based on the above result for non constant time steps we devise a simple adaptive algorithm and present some computational results.

For $n \geq 1$ define the local error indicator

$$\vartheta_n{}^2 := \frac{\tau_n^2}{\omega_n^2} \left\| \frac{\omega_n}{\omega_n + 1} \bar{\partial}^2 U^n - \frac{\omega_{n-1}}{\omega_{n-1}+1} \bar{\partial}^2 U^{n-1} \right\|^2_{J^*} + \tau_n^4 \|\bar{\partial}^2 U^n\|^2. \tag{7.4}$$

Note that for $n \leq 3$ the first term is dropped.

From Theorem 7.1 we conclude

$$\|U-u\|^2_{L^2(L^2)} \lesssim \sum_{n=1}^{N} \tau_n {\vartheta_n}^2 \; + \Lambda_1 + \mathcal{F}_1 + \mathfrak{I}_1.$$

The optimal strategy to choose the time step sizes with respect to the L^2–norm in time would consist in equilibrating the local contributions $\tau_n {\vartheta_n}^2$. This, however, is not possible within a time-marching scheme, since it would require to *a priori* know the total number N of time steps. Instead, one contents oneself with equilibrating the local contribution ϑ_n, which is more suited for the L^∞–norm (compare also the discussion in the study by Kreuzer *et al.*, 2012).

Thus we propose the following simple adaptive time-stepping strategy, which is rather common for ordinary differential equations, see Deuflhard & Weiser (2012): in each time step a tentative size of the next time step is chosen according to the extrapolation formula

$$\tau_{n+1} := \tau_n \mathrm{TOL}/{\vartheta_n}^{\frac{1}{\gamma+1}},$$

where γ is the order of the method, $\gamma = 2$ in our case. If the new time step generates too big an error, this time step is rejected and recomputed with a reduced time step. Otherwise the time step is accepted. With some 'safety' parameters ρ_1, ρ_2, ρ_2 the algorithm reads as follows.

Algorithm 1 Let $\rho_1, \rho_2, \rho_3 > 1$ and TOL > 0 be given. Initialize all unknowns. Let τ_1 be given.

---

do

  $n \leftarrow n+1$

  loop forever

    $[U^n, P^n] =$ `solve` (U^{n-1}, U^{n-2})

    compute ϑ_n *from Eq.* (83)

    if $\vartheta_n \rho_1 * TOL$ $*$ *time step rejected*

      $\tau_n \leftarrow \tau_n/\rho_2$

    else $*$ *time step accepted*

      break

    end if

  end loop forever

  $\tau_{n+1} := \min\{\tau_n (TOL/\vartheta_n)^{1/3}, \rho_3 \tau_n\}$

  $t_{n+1} := t_n + \tau_{n+1}$

while $t_{n+1} < T$

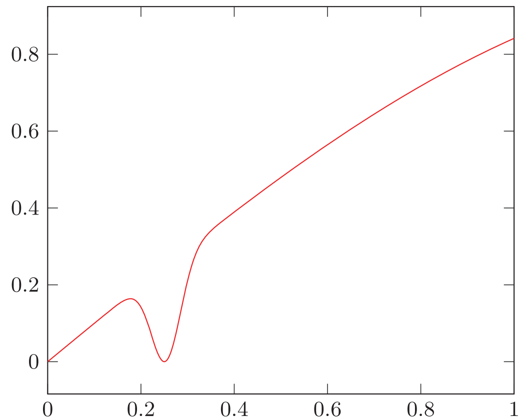

FIG. 4. Graph of function g.

Note that `solve` returns the unknowns for time t_n by solving one time step of equation (2.9). The above adaptive method is tested with the same example as in Section 6, except for taking

$$u(t,x) = g(t)\pi \left[\sin(2\pi x_2)\sin(\pi x_1)^2,\ -\sin(2\pi x_1)\sin(\pi x_2)^2\right]^\top,$$
$$p(t,x) = g(t)\cos(\pi x_1)\sin(\pi x_2)$$

with $g(t) = \sin(t)\left(1-\exp\left(-5.e2(t-0.25)^2\right)\right)$ exhibiting a 'bump' around $t = 0.25$, see Fig. 4. This behavior of g is well suited to study the performance of a time adaptive strategy (compare also Kreuzer *et al.*, 2012).

Figure 5 (left) shows the result for a computation with uniform time step size and with adaptive strategy. For a fair comparison the uniform time step is chosen such that the number of time steps is

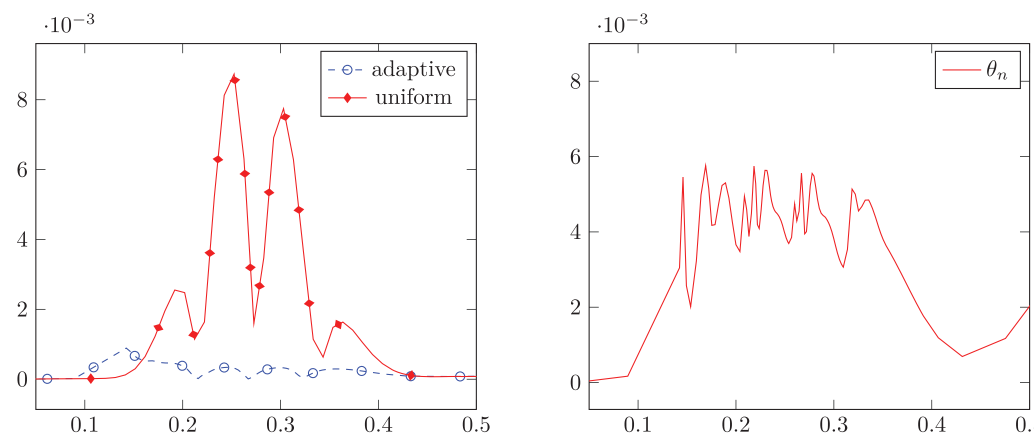

FIG. 5. Instantaneous $L^2(\Omega)$ error versus time with uniform time step and adaptive strategy (left) and local error estimator θ^n versus time (right).

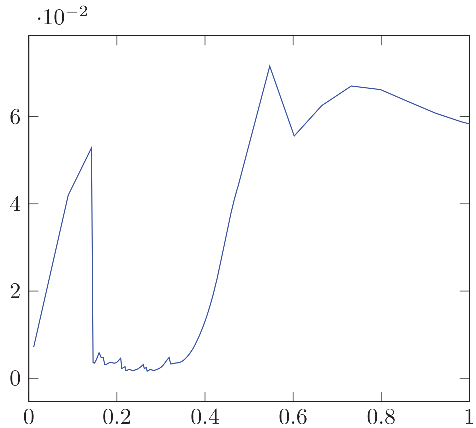

FIG. 6. Time step sizes τ_n versus time.

the same as for the adaptive run. As can be seen, the local error of the computation with uniform time step is rather big around $t = 0.25$, otherwise they are quite small, thus reflecting the character of the function g. The run with the adaptive strategy is done with TOL $= 5.e - 3$, $\rho_1 := 1.2$, $\rho_2 = 1.43$ and $\rho_3 = 1.8$. As one can see the error is fairly well controlled. Figure 5 (right) shows the error estimator θ^n, equilibrated around the value $5.e - 3$, and Fig. 6 reports the time step size as a function of time for the adaptive strategy. We conclude that in principle the derived error estimator is capable of effectively monitoring the time step size.

References

AINSWORTH, M. & ODEN, J. T. (2011) *A Posteriori Error Estimation in Finite element Analysis*, vol. 37. New York: John Wiley.

AKRIVIS, G. & CHATZIPANTELIDIS, P. (2010) A posteriori error estimates for the two-step backward differentiation formula method for parabolic equations. *SIAM J. Numer. Anal.*, **48**, 109–132.

AKRIVIS, G., MAKRIDAKIS, CH. & NOCHETTO, R. H. (2006) A posteriori error estimates for the Crank–Nicolson method for parabolic equations. *Math. Comp.*, **75**, 511–531.

AKRIVIS, G., MAKRIDAKIS, CH. & NOCHETTO, R. H. (2009) Optimal order a posteriori error estimates for a class of Runge–Kutta and Galerkin methods. *Numer. Math.*, **114**, 133–160.

ALNÆS, M., BLECHTA, J., HAKE, J., JOHANSSON, A., KEHLET, B., LOGG, A., RICHARDSON, C., RING, J., ROGNES, M. & WELLS, G. (2015) The FEniCS project version 1.5. *Archive of Num. Software*, **3**, 9–23.

BANK, R. E. & WELFERT, B. D. (1990) A posteriori error estimates for the Stokes equations: a comparison. *Comput. Methods Appl. Mech. Eng.*, **82**, 323–340.

BÄNSCH, E. (1991a) An adaptive finite-element strategy for the three-dimensional time-dependent Navier–Stokes equations. *J. Comput. Appl. Math.*, **36**, 3–28.

BÄNSCH, E. (1991b) Local mesh refinement in 2 and 3 dimensions. *IMPACT Comput. Sci. Engrg.*, **3**, 181–191.

BÄNSCH, E. (1998) Simulation of instationary, incompressible flows. *Acta Math. Univ. Comenianae*, **67**, 101–114.

BÄNSCH, E. & BRENNER, A. (2016) A posteriori error estimates for pressure-correction schemes. *SIAM J. Numer. Anal.*, **54**, 2323–2358.

BÄNSCH, E., KARAKATSANI, F. & MAKRIDAKIS, CH. (2012) A posteriori error control for fully discrete Crank–Nicolson schemes. *SIAM J. Numer. Anal.*, **50**, 2845–2872.

BESIER, M. & WOLLNER, W. (2012) On the pressure approximation in nonstationary incompressible flow simulations on dynamically varying spatial meshes. *Int. J. Numer. Methods Fluids*, **69**, 1045–1064.

Brenner, A., Bänsch, E. & Bause, M. (2014) *A priori* error analysis for finite element approximations of the stokes problem on dynamic meshes. *IMA J. Numer. Anal.*, **34**, 123–146.

Crouzeix, M. & Raviart, P.-A. (1973) Conforming and nonconforming finite element methods for solving the stationary stokes equations i. *Esaim. Math. Model. and Numer. Anal.*, **7**, 33–75.

Demlow, A., Lakkis, O. & Makridakis, Ch. (2009) A posteriori error estimates in the maximum norm for parabolic problems. *SIAM J. Numer. Anal.*, **47**, 2157–2176.

Deuflhard, P. & Weiser, M. (2012) Adaptive numerical solution of PDEs. Berlin: de Gruyter.

Dupont, T. (1982) Mesh modification for evolution equations. *Math. Comp.*, **39**, 85–107.

Emmrich, E. (2005) Stability and error of the variable two-step BDF for semilinear parabolic problems. *JAMS*, **19**, 33–55.

Ern, A., Smears, I. & Vohralík, M. (2016) Guaranteed, locally space-time efficient, and polynomial-degree robust a posteriori error estimates for high-order discretizations of parabolic problems. *Technical report*. arXiv:1610.01804.

Evans, L. C. (2010) *Partial Differential Equations*, 2nd edn. Graduate Studies in Mathematics, vol. 33. Providence, RI: American Mathematical Society.

Georgoulis, E., Lakkis, O. & Virtanen, J. (2011) A posteriori error control for discontinuous Galerkin methods for parabolic problems. *SIAM J. Numer. Anal.*, **49**, 427–458.

Heywood, J. G. (1980) The Navier–Stokes equations: on the existence, regularity and decay of solutions. *Indiana Univ. Math. J.*, **29**, 639–681.

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

Heywood, J. G. & Rannacher, R. (1988) 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**, 489–512.

Karakatsani, F. & Makridakis, Ch. (2007) A posteriori estimates for approximations of time-dependent Stokes equations. *IMA J. Numer. Anal.*, **27**, 741–764.

Kreuzer, C., Schmidt, A., Möller, C. A. & Siebert, K. G. (2012) Design and convergence analysis for an adaptive discretization of the heat equation. *IMA J. Numer. Anal.*, **32**, 1375–1403.

Lakkis, O. & Makridakis, Ch. (2006) Elliptic reconstruction and a posteriori error estimates for fully discrete linear parabolic problems. *Math. Comp.*, **75**, 1627–1658.

Lozinski, A., Picasso, M. & Prachittham, V. (2009) An anisotropic error estimator for the Crank–Nicolson method: application to a parabolic problem. *SIAM J. Sci. Comput.*, **31**, 1757–1783.

Makridakis, Ch. & Nochetto, R. H. (2003) Elliptic reconstruction and a posteriori error estimates for parabolic problems. *SIAM J. Numer. Anal.*, **41**, 1585–1594.

Sohr, H. (2001) *The Navier–Stokes Equations*. Basel: Birkhäuser.

Temam, R. (1977) *Navier–Stokes Equations*, 3rd edn. Amsterdam: North-Holland.

Verfürth, R. (1996) *A Review of A Posteriori Error Estimation and Adaptive Mesh-refinement Techniques*. *Wiley-Teubner Series*. Chichester-Stuttgart.

Verfürth, R. (2003) *A posteriori* error estimates for finite element discretizations of the heat equation. *Calcolo*, **40**, 195–212.

Verfühurt, R. (2013) *A Posteriori Error Estimation Techniques for Finite Element Methods*. Oxford Science Publications.

Appendix A. Proof of Theorem 2.5

In the first step *a priori* estimates of the initial errors are presented:

Lemma A.1 Assume the exact solution of the Stokes system (2.1) is sufficiently regular. Determine p_0 from u_0 and equation (2.1). Set $U^0 = \mathcal{S}_0^u(u_0, p_0)$ and $P^0 := \mathcal{S}_0^p(u_0, p_0)$ of Lemma 2.3. Calculate $U^1 \in X_h$

and $P^1 \in Y_h$ by equation (2.10). Then the approximation estimate of the Stokes operator holds

$$\begin{aligned}\|U^1 - \mathcal{S}_1^u(u(t_1), p(t_1))\|^2 + \tau\|\nabla(U^1 - \mathcal{S}_1^u(u(t_1), p(t_1)))\|^2 + \tau^2\|P^1 - \mathcal{S}_1^p(u(t_1), p(t_1))\|^2 \\ \lesssim \tau^2 h^{2(k+1)} A^1 + \tau^6 \|\partial_t^3 u\|^2_{L^\infty((0,t_1);L^2)} \end{aligned} \tag{A.1}$$

and the fully discrete estimates

$$\|U^1 - u(t_1)\|^2 \lesssim h^{2(k+1)} A^0 + \tau^6 \|\partial_t^3 u\|^2_{L^\infty((0,t_1);L^2)}, \tag{A.2}$$

$$\|\nabla(U^1 - u(t_1))\|^2 \lesssim h^{2k} A^0 + \tau h^{2(k+1)} A^1 + \tau^5 \|\partial_t^3 u\|^2_{L^\infty((0,t_1);L^2)}, \tag{A.3}$$

$$\|P^1 - p(t_1)\|^2 \lesssim h^{2k} A^0 + h^{2(k+1)} A^1 + \tau^4 \|\partial_t^3 u\|^2_{L^\infty((0,t_1);L^2)}, \tag{A.4}$$

where

$$A^i := \|\partial_t^i u\|^2_{L^\infty((0,t_1);H^{k+1})} + \|\partial_t^i p\|^2_{L^\infty((0,t_1);H^k)}.$$

Proof. Set $\theta^i := U^i - \mathcal{S}_i^u(u_i, p_i)$. Subtract the half of equation (2.1) at times $t = 0$ and $t = t_1$ from equation (2.10), use the Stokes projection defined in Lemma 2.3 and note that $\theta^0 = 0$ and $P^0 - \mathcal{S}_0^p(u_0, p_0) = 0$ to deduce the error equation

$$\frac{1}{\tau}\theta^1 - \frac{1}{2}\Delta_h \theta^1 + \frac{1}{2} B_h^\top (P^1 - \mathcal{S}_1^p(u(t_1), p(t_1))) = \bar{\partial}(u(t_1) - \mathcal{S}_n^u(u(t_1), p(t_1))) + (\partial_t u(t_{1/2}) - \bar{\partial} u(t_1)). \tag{A.5}$$

Test this equation by $\tau\theta^1 \in J_h$ to obtain

$$\begin{aligned}\|\theta^1\|^2 + \tau\|\nabla\theta^1\|^2 &\lesssim \tau^2 \|\bar{\partial}(u(t_1) - \mathcal{S}_1^u(u(t_1), p(t_1)))\|^2 + \tau^2\|(\partial_t u(t_{1/2}) - \bar{\partial} u(t_1))\|^2 \\ &\lesssim \tau^2 h^{2(k+1)} A^1 + \tau^6 \|\partial_t^3 u\|^2_{L^\infty((0,t_1);L^2)}, \end{aligned} \tag{A.6}$$

where Lemma 2.2 and Taylor expansion has been used. This proves the estimate for the first and second term in equation (A.1). Equations (A.2) and (A.3) are then easily deduced by the triangle inequality and Lemma 2.2.

To show the pressure estimate (A.4) let us first show the estimate for $\|P^1 - \mathcal{S}_1^p(u(t_1), p(t_1))\|$ in equation (A.1). Use the inf–sup condition and insert the error equation (A.5) to get

$$\begin{aligned}\|P^1 - \mathcal{S}_1^p(u(t_1), p(t_1))\| &\lesssim \sup_{v_h \in X_h} \frac{|\langle P^1 - \mathcal{S}_1^p(u(t_1), p(t_1)), \operatorname{div} v_h\rangle|^2}{\|\nabla v_h\|} \\ &\le \frac{1}{\tau}\|\theta^1\|_{X_h^*} + \|\nabla\theta^1\| + \|\bar{\partial}(u(t_1) - \mathcal{S}_1^u(u(t_1), p(t_1)))\|_{X_h^*} + \|\partial_t u(t_{1/2}) - \bar{\partial} u(t_1)\|_{X^*} \\ &\le h^{2(k+1)} A^1 + \tau^4 \|\partial_t^3 u\|^2_{L^\infty((0,t_1);L^2)}, \end{aligned}$$

where we have bounded $\|\theta^1\|_{X_h^*} \le \|\theta^1\|$ and used equation (A.6) in the last step.

The pressure estimate (A.4) is then obtained by the triangle inequality and Lemma 2.3:

$$\|P^1 - p(t_1)\| \le \|P^1 - \mathcal{S}_1^p(u(t_1), p(t_1))\| + \|p(t_1) - \mathcal{S}_1^p(u(t_1), p(t_1))\|.$$

□

Now we are able to prove Theorem 2.5:

Proof. **(Theorem 2.5)** Set $\theta^n := U^n - \mathcal{S}_n^u(u(t_n), p(t_n))$. Subtract equation (2.1) from equation (2.9) and combine with the Stokes projector $\mathcal{S}_n^u$ to get the error equation

$$\bar{\partial}^B\theta^n - \Delta_h\theta^n + B_h^\top(P^n - \mathcal{S}_n^p(u(t_n), p(t_n))) = \bar{\partial}^B(u(t_n) - \mathcal{S}_n^u(u(t_n), p(t_n))) + (\partial_t - \bar{\partial}^B)u(t_n), \quad n \geq 2. \tag{A.7}$$

Test this equation by $4\tau\theta^n \in J_h$. By the identity

$$2\langle 3a^n - 4a^{n-1} + a^{n-2}, a^n\rangle = \|a^n\|^2 - \|a^{n-1}\|^2 + \tau^4\|\bar{\partial}^2 a^n\|^2 + \|2a^n - a^{n-1}\|^2 - \|2a^{n-1} - a^{n-2}\|^2 \tag{A.8}$$

the first term can be transformed into

$$\begin{aligned} 4\tau\langle\bar{\partial}^B\theta^n, \theta^n\rangle &= 2\langle 3\theta^n - 4\theta^{n-1} + \theta^{n-2}, \theta^n\rangle \\ &= \|\theta^n\|^2 - \|\theta^{n-1}\|^2 + \tau^4\|\bar{\partial}^2\theta^n\|^2 + \|2\theta^n - \theta^{n-1}\|^2 - \|2\theta^{n-1} - \theta^{n-2}\|^2. \end{aligned}$$

Since $\theta^n \in J_h$ the pressure terms drop out and one obtains the error inequality

$$\begin{aligned} &\|\theta^n\|^2 - \|\theta^{n-1}\|^2 + \tau^4\|\bar{\partial}^2\theta^n\|^2 + \|2\theta^n - \theta^{n-1}\|^2 - \|2\theta^{n-1} - \theta^{n-2}\|^2 + \tau\|\nabla\theta^n\|^2 \\ &\qquad \lesssim \tau\|\bar{\partial}^B(u(t_n) - \mathcal{S}_n^u(u(t_n), p(t_n)))\|_{J_h^*}^2 + \tau\|(\partial_t - \bar{\partial}^B)u(t_n)\|_{J_h^*}^2. \end{aligned}$$

Summing up from $n = 2$ to N yields

$$\begin{aligned} \max_{1\leq n\leq N}\|\theta^n\|^2 + \sum_{n=2}^N \tau\|\nabla\theta^n\|^2 &\lesssim \sum_{n=2}^N \tau\|\bar{\partial}^B(u(t_n) - \mathcal{S}_n^u(u(t_n), p(t_n)))\|_{J_h^*}^2 \\ &\quad + \sum_{n=2}^N \tau\|(\partial_t - \bar{\partial}^B)u(t_n)\|_{J_h^*}^2 + \|\theta^1\|^2 + \|\theta^1 - \theta^0\|^2. \end{aligned}$$

For the first term on the right-hand side we have by using Lemma 2.2

$$\sum_{n=2}^N \tau\|\bar{\partial}^B(u(t_n) - \mathcal{S}_n^u(u(t_n), p(t_n)))\|_{J_h^*}^2 \lesssim \sum_{n=1}^N \tau\|\bar{\partial}u(t_n) - \mathcal{S}_n^u(\bar{\partial}u(t_n), \bar{\partial}p(t_n))\|^2 \lesssim h^{2(k+1)}A^1. \tag{A.9}$$

For the second term on the right-hand side use standard estimates for the BDF2 operator $\bar{\partial}^B$ and get

$$\max_{1\leq n\leq N}\|\theta^n\|^2 + \sum_{n=2}^N \tau\|\nabla\theta^n\|^2 \lesssim \|\theta^1\|^2 + h^{2(k+1)}A^1 + \tau^4\|\partial_t^3 u\|_{L^\infty(H^{-1})}^2. \tag{A.10}$$

The velocity estimate (2.11) is finally obtained using the triangle inequality:

$$\max_{1\leq n\leq N}\|U^n - u(t_n)\| \lesssim \max_{1\leq n\leq N}\|\theta^n\| + \max_{1\leq n\leq N}\|u(t_n) - \mathcal{S}_n^u(u(t_n), p(t_n))\|,$$

the initial estimate (A.1), equation (A.10) and Lemma 2.2.

In the same way equation (2.12) is proved.

To derive the pressure estimate (2.13) we first bound the term $\|\bar{\partial}^B U^n - \partial_t u\|$. To this end test the error equation (A.7) by $4\tau\bar{\partial}^B\theta^n$ and use the identity (A.8) to get

$$\tau\|\bar{\partial}^B\theta^n\|^2 + \|\nabla\theta^n\|^2 - \|\nabla\theta^{n-1}\|^2 + \tau^4\|\nabla\bar{\partial}\theta^n\|^2 + \|\nabla(2\theta^n - \theta^{n-1})\|^2 - \|\nabla(2\theta^{n-1} - \theta^{n-2})\|^2$$
$$\lesssim \tau\|\bar{\partial}^B(u(t_n) - \mathcal{S}_n^u(u(t_n), p(t_n)))\|^2 + \tau\|(\partial_t - \bar{\partial}^B)u(t_n)\|^2.$$

Sum up from $n = 2$ to N, observe that $\theta^0 = 0$ to obtain

$$\sum_{n=2}^{N}\tau\|\bar{\partial}^B\theta^n\|^2 + \max_{1\le n\le N}\|\nabla\theta^n\|^2 \lesssim \|\nabla\theta^1\|^2 + \sum_{n=2}^{N}\tau\|\bar{\partial}(u(t_n) - \mathcal{S}_n^u(u(t_n), p(t_n)))\|^2$$
$$+ \sum_{n=2}^{N}\tau\|(\partial_t - \bar{\partial}^B)u(t_n)\|^2. \tag{A.11}$$

Now we are able to prove the pressure estimate (2.13). The inf–sup stability yields

$$\begin{aligned}
\|P^n - p(t_n)\| &\le \|P^n - \mathcal{S}_n^p(u(t_n), p(t_n))\| + \|\mathcal{S}_n^p(u(t_n), p(t_n)) - p(t_n)\| \\
&\lesssim \sup_{v_h\in X_h}\frac{\langle P^n - \mathcal{S}_n^p(u(t_n), p(t_n)), \operatorname{div} v_h\rangle}{\|\nabla v_h\|} + \|\mathcal{S}_n^p(u(t_n), p(t_n)) - p(t_n)\| \\
&= \sup_{v_h\in X_h}\frac{\langle\bar{\partial}^B U^n - \partial_t u(t_n), v_n\rangle + \langle\nabla\theta^n, \nabla v_h\rangle}{\|\nabla v_h\|} + \|\mathcal{S}_n^p(u(t_n), p(t_n)) - p(t_n)\| \\
&\le \|\bar{\partial}^B\theta^n\| + \|\bar{\partial}^B(u(t_n) - \mathcal{S}_n^u(u(t_n), p(t_n)))\| + \|(\bar{\partial}^B - \partial_t)u(t_n)\| \\
&\quad + \|\nabla\theta^n\| + \|\mathcal{S}_n^p(u(t_n), p(t_n)) - p(t_n)\|.
\end{aligned}$$

Square this inequality, multiply by τ and use equation (A.11) to get

$$\sum_{n=1}^{N}\tau\|P^n - p(t_n)\|^2 \lesssim \|\nabla\theta^1\|^2 + \tau\|P^1 - p(t_1)\|^2$$
$$+ \sum_{n=2}^{N}\tau\Big(\|\bar{\partial}^B(u(t_n) - \mathcal{S}_n^u(u(t_n), p(t_n)))\|^2 + \|(\partial_t - \bar{\partial}^B)u(t_n)\|^2$$
$$+\|\nabla\theta^n\|^2 + \|\mathcal{S}_n^p(u(t_n), p(t_n)) - p(t_n)\|^2\Big).$$

Then by using the initial estimate equations (A.1)–(A.4), the velocity estimate (A.10) and Lemma 2.3 the pressure estimate in equation (2.13) is proven.

The $\ell^\infty(L^2)$ estimate for the gradients of the velocity is easily derived by equation (A.11) and Lemma 2.2. □

Appendix B. Discrete regularity

In this section we derive regularity estimates for the solution of the semidiscrete instationary Stokes problem

$$\partial_t Z(t) - \Delta_h Z(t) + B_h^\top Q(t) = \tilde{G}(t), \qquad B_h Z(t) = 0, \qquad Z(t) = V_0, \tag{B.1}$$

for a sufficiently regular (in time) right-hand side $\tilde{G}(t) \in X_h$ and initial value V_0 or, equivalently, by multiplying the above equation by Π_{J_h},

$$\partial_t Z(t) + \mathcal{A}_h Z(t) = F(t) := \Pi_{J_h} \tilde{G}(t), \qquad B_h Z(t) = 0, \qquad Z(0) = V_0. \tag{B.2}$$

Here, $\Pi_{J_h} : X \to J_h$ denotes the L^2 projection onto the discrete solenoidal space J_h and

$$\mathcal{A}_h := \Pi_{J_h}(-\Delta_h) \tag{B.3}$$

the discrete Stokes operator.

Regularity results for the full continuous system equation (2.1) can be found for instance in the study by Temam (1977, Chapter III), Sohr (2001, Chapter IV.2.7) and Heywood (1980).

The following remark seems to be in order. The crucial difference between the continuous and the semidiscrete case is that in the latter case no *compatibility conditions* between the initial data and right-hand side are needed for getting higher regularity, since the domain of the discrete operator $\mathcal{A}_h$ is the whole of J_h.

In contrast, in the continuous case higher regularity in time is always linked to such compatibility conditions, even if data are arbitrarily smooth. This can readily be seen for the heat equation: assume that u, $\triangle u, f$ and $\partial_t u$ are continuous on $[0, T] \times \bar{\Omega}$, satisfying

$$\partial_t u - \Delta u = f, \qquad u_{|\Gamma} = 0, \qquad u(0) = u_0.$$

Since $u(t, x)_{|\Gamma} = 0$ for $t \in [0, T]$, we see that $\partial_t u(0)_{|\Gamma} = 0$ and therefore the additional condition for given data f, u_0 must be satisfied:

$$\Delta u_{0|\Gamma} = f(0)_{|\Gamma}.$$

Further compatibility conditions are required for higher order time derivatives $\partial_t^{(m)} u$, see for instance Evans (2010). For the Stokes problem the situation is even more involved because of the presence of the pressure gradient. Only if the pressure fulfills a Poisson equation with both, a Dirichlet and a Neumann boundary condition (and is thus overdetermined), higher time regularity can hold. This kind of compatibility condition cannot be checked in practice, see also Heywood & Rannacher (1982). To avoid these complicated conditions Sohr presents the following local results, cf. Sohr (2001, Theorem 2.7.2):

Let $m \in \mathbb{N}$, let $\Omega \subset \mathbb{R}^d$, $d \geq 2$, be a uniform C^2-domain, let $0 < T \leq \infty$, $1 < s < \infty$, $u_0 \in D\left(\mathcal{A}^{1-\frac{1}{s}}\right) \cap D\left(\mathcal{A}^{\frac{1}{2}}\right), f \in L^2_{loc}\left([0, T); L^2(\Omega)^d\right)$, and suppose u is a solution of the Stokes system (2.1) with data f, u_0 together with an associated pressure p. Then

$$\partial_t f, \partial_{tt} f, \ldots, \partial_t^{(m)} f \in L^s_{loc}\left((0, T); L^2(\Omega)^d\right)$$

implies

$$\partial_t u, \partial_{tt} u, \ldots, \partial_t^{(m+1)} u,\ \mathcal{A}\partial_t u, \mathcal{A}\partial_{tt} u, \ldots, \mathcal{A}\partial_t^{(m+1)} u \in L^s_{loc}((0, T); H),$$

and

$$\partial_t u, \partial_{tt} u, \ldots, \partial_t^{(m+1)} u \in L^s_{loc}\left((0, T); H^2(\Omega)^d\right).$$

Let us return to the problem of regularity for the semidiscrete problem. Since (B.2) is a linear, finite-dimensional evolution equation, it admits a unique solution for all times $t > 0$. Moreover, differentiability of u with respect to time directly follows from regularity of $G(t)$.

The following lemma gives an equivalent expression for the dual norm of J_h in terms of the discrete Stokes operator.

LEMMA B.1 For $y_h \in J_h$ there is a unique $x_h \in J_h$, solution of the discrete Stokes problem $\mathcal{A}_h x_h = y_h$. Moreover,

$$\|y_h\|_{J_h^*} = \|\nabla x_h\|. \tag{B.4}$$

Proof. One has

$$\langle \mathcal{A}_h v_h, w_h \rangle = \langle \Pi_{J_h}(-\Delta_h v_h), w_h \rangle = \langle -\Delta_h v_h, w_h \rangle = \langle \nabla v_h, \nabla w_h \rangle, \qquad \forall\ v_h,\ w_h \in J_h. \tag{B.5}$$

This shows that $\mathcal{A}_h$ is symmetric and coercive on J_h and in particular bijective. Using the above representation for $\mathcal{A}_h$, the lemma can be easily proven.

First,

$$\|y_h\|_{J_h^*} = \sup_{z_h \in J_h} \frac{\langle y_h, z_h \rangle}{\|\nabla z_h\|} = \sup_{z_h \in J_h} \frac{\langle \mathcal{A}_h x_h, z_h \rangle}{\|\nabla z_h\|} = \sup_{z_h \in J_h} \frac{\langle -\Delta_h x_h, z_h \rangle}{\|\nabla z_h\|} \le \|\nabla x_h\|.$$

Second,

$$\|\nabla x_h\| = \frac{\|\nabla x_h\|^2}{\|\nabla x_h\|} = \frac{\langle -\Delta_h x_h, x_h \rangle}{\|\nabla x_h\|} = \frac{\langle \mathcal{A}_h x_h, x_h \rangle}{\|\nabla x_h\|} = \frac{\langle y_h, x_h \rangle}{\|\nabla x_h\|} \le \sup_{x_h \in J_h} \frac{\langle y_h, x_h \rangle}{\|\nabla x_h\|} = \|y_h\|_{J_h^*}.$$

□

In the following theorem, regularity estimates for the solution of the semidiscrete system (B.2) in terms of the regularity of the right-hand side G and the initial data V_0 are presented.

THEOREM B.2 Suppose that $\partial_t^{(m)} G \in L^2((0,T), J_h)$ for all $1 \le m \le M$. Then the solution $Z(t)$ of equation (B.2) satisfies for all $1 \le m \le M$

$$\|\partial_t^{(m)} Z\|^2_{L^\infty(L^2)} + \|\partial_t^{(m)} \nabla Z\|^2_{L^2(L^2)} \lesssim \|\partial_t^{(m)} G\|^2_{L^2(J_h^*)} + \|Y_0^m\|^2, \tag{B.6}$$

$$\|\partial_t^{(m+1)} Z\|^2_{L^2(L^2)} + \|\partial_t^{(m)} \nabla Z\|^2_{L^\infty(L^2)} \lesssim \|\partial_t^{(m)} G\|^2_{L^2(L^2)} + \|Y_0^m\|^2_1, \tag{B.7}$$

$$\|\partial_t^{(m)} \nabla Z\|^2_{L^\infty(L^2)} + \|\partial_t^{(m)} \mathcal{A}_h Z\|^2_{L^2(L^2)} \lesssim \|\partial_t^{(m)} G\|^2_{L^2(L^2)} + \|Y_0^m\|^2_1, \tag{B.8}$$

$$\|\partial_t^{(m+1)} \nabla Z\|^2_{L^2(L^2)} + \|\partial_t^{(m)} \mathcal{A}_h Z\|^2_{L^\infty(L^2)} \lesssim \|\nabla \partial_t^{(m)} G\|^2_{L^2(L^2)} + \|Y_0^m\|^2_{\mathcal{A}}, \tag{B.9}$$

$$\|\partial_t^{(m)} Z\|^2_{L^\infty(J_h^*)} + \|\partial_t^{(m)} Z\|^2_{L^2(L^2)} \lesssim \|\partial_t^{(m)} G\|^2_{L^1(J_h^*)} + \|Y_0^m\|^2_{J_h^*}, \tag{B.10}$$

$$\|\partial_t^{(m+1)} Z\|^2_{L^2(J_h^*)} + \|\partial_t^{(m)} Z\|^2_{L^\infty(L^2)} \lesssim \|\partial_t^{(m)} G\|^2_{L^2(J_h^*)} + \|Y_0^m\|^2, \tag{B.11}$$

where

$$\|Y_0^m\|_*^2 := \|\mathcal{A}_h^m Z(0)\|_*^2 + \sum_{i=0}^{m-1} \|\partial_t^{(i)} \mathcal{A}_h^{m-1-i} G(0)\|_*^2, \qquad * \in \{1, 0, J_h^*\},$$

$$\|Y_0^m\|_{\mathcal{A}}^2 := \|\mathcal{A}_h^{m+1} Z(0)\|^2 + \sum_{i=0}^{m-1} \|\partial_t^{(i)} \mathcal{A}_h^{m-i} G(0)\|^2.$$

Proof. Let $1 \le m \le M$. Differentiating equation (B.2) m-times with respect to time leads to

$$\partial_t^{(m+1)} Z(t) + \partial_t^{(m)} \mathcal{A}_h Z(t) = \partial_t^{(m)} G(t). \tag{B.12}$$

First, to prove equation (B.6), test equation (B.12) with $2\partial_t^{(m)} Z(t) \in J_h$ to get

$$\frac{\mathrm{d}}{\mathrm{d}t} \|\partial_t^{(m)} Z(t)\|^2 + \|\nabla \partial_t^{(m)} Z(t)\|^2 \lesssim \|\partial_t^{(m)} G(t)\|_{J_h^*}^2,$$

where we have used equation (B.5) for the second term on the left-hand side. Now, integrating the above equation from $t = 0$ to T yields

$$\|\partial_t^{(m)} Z\|_{L^\infty(L^2)}^2 + \|\partial_t^{(m)} \nabla Z\|_{L^2(L^2)}^2 \lesssim \|\partial_t^{(m)} G\|_{L^2(J_h^*)}^2 + \|\partial_t^{(m)} Z(0)\|^2.$$

Next, to prove equations (B.7)–(B.9), test equation (B.12) with $v_h = \partial_t^{(m+1)} Z(t)$, $\partial_t^{(m)} \mathcal{A}_h Z(t)$ and $\partial_t^{(m+1)} \mathcal{A}_h Z(t)$, and proceed exactly as above to get

- $v_h = \partial_t^{(m+1)} Z(t)$: $\|\partial_t^{(m+1)} Z\|_{L^2(L^2)}^2 + \|\partial_t^{(m)} \nabla Z\|_{L^\infty(L^2)}^2 \lesssim \|\partial_t^{(m)} G\|_{L^2(L^2)}^2 + \|\partial_t^{(m)} \nabla Z(0)\|^2$,
- $v_h = \partial_t^{(m)} \mathcal{A}_h Z(t)$: $\|\partial_t^{(m)} \nabla Z\|_{L^\infty(L^2)}^2 + \|\partial_t^{(m)} \mathcal{A}_h Z\|_{L^2(L^2)}^2 \lesssim \|\partial_t^{(m)} G\|_{L^2(L^2)}^2 + \|\partial_t^{(m)} \nabla Z(0)\|^2$,
- $v_h = \partial_t^{(m+1)} \mathcal{A}_h Z(t)$: $\|\partial_t^{(m+1)} \nabla Z\|_{L^2(L^2)}^2 + \|\partial_t^{(m)} \mathcal{A}_h Z\|_{L^\infty(L^2)}^2 \lesssim \|\nabla \partial_t^{(m)} G\|_{L^2(L^2)}^2 + \|\partial_t^{(m)} \mathcal{A}_h Z(0)\|^2$,

respectively.

For proving equations (B.10) and (B.11) we use a duality argument. We test equation (B.12) with v_h, where v_h is solution of the Stokes problem $\mathcal{A}_h v_h = \partial_t^{(m)} Z(t)$. For the left-hand side, one gets

$$\langle \partial_t^{(m+1)} Z(t), v_h(t) \rangle = \langle \partial_t \mathcal{A}_h v_h(t), v_h(t) \rangle = \frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \|\nabla v_h(t)\|^2,$$

$$\langle \partial_t^{(m)} \mathcal{A}_h Z(t), v_h \rangle = \langle \partial_t^{(m)} - \Delta_h Z(t), v_h \rangle = \langle \partial_t^{(m)} Z(t), \mathcal{A}_h v_h \rangle = \|\partial_t^{(m)} Z(t)\|^2,$$

where for the second term equation (B.5) was used. For the right-hand side it follows $\langle \partial_t^{(m)} G(t), v_h \rangle \le \|\partial_t^{(m)} G(t)\|_{J_h^*} \|\nabla v_h\|$ and then

$$\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \|\nabla v_h(t)\|^2 + \|\partial_t^{(m)} Z(t)\|^2 \le \|\partial_t^{(m)} G(t)\|_{J_h^*} \|\nabla v_h\|.$$

Integrate from $t = 0$ to T and use the norm equivalence in Lemma B.1 to get

$$\|\partial_t^{(m)} Z\|_{L^\infty(J_h^*)}^2 + \|\partial_t^{(m)} Z\|_{L^2(L^2)}^2 \lesssim \|\partial_t^{(m)} G\|_{L^1(J_h^*)}^2 + \|\partial_t^{(m)} Z(0)\|_{J_h^*}^2.$$

Now let v_h be solution of the Stokes problem $\mathcal{A}_h v_h = \partial_t^{(m+1)} Z(t)$. The same procedure as above yields

$$\|\partial_t^{(m+1)} Z\|_{L^2(J_h^*)}^2 + \|\partial_t^{(m)} Z\|_{L^\infty(L^2)}^2 \lesssim \|\partial_t^{(m)} G\|_{L^2(J_h^*)}^2 + \|\partial_t^{(m)} Z(0)\|^2.$$

It remains to bound the initial values $\|\partial_t^{(m)} Z(0)\|^2$, $\|\partial_t^{(m)} \nabla Z(0)\|^2$, $\|\partial_t^{(m)} \mathcal{A}_h Z(0)\|^2$ and $\|\partial_t^{(m)} Z(0)\|_{J_h^*}^2$. To this end we prove the following identity for $m \geq 1$

$$\partial_t^{(m)} Z(0) = \sum_{i=0}^{m-1} (-1)^{i-1+m} \partial_t^{(i)} \mathcal{A}_h^{m-1-i} G(0) + (-1)^m \mathcal{A}_h^m Z(0). \tag{B.13}$$

The cases $m = 1$ and $m = 2$ are easily shown by using equation (B.12) and taking the limit $t \searrow 0$.

$$\partial_t Z(0) = G(0) - \mathcal{A}_h Z(0),$$
$$\partial_{tt} Z(0) = \partial_t G(0) - \mathcal{A}_h \partial_t Z(0) = \partial_t G(0) - \mathcal{A}_h G(0) + \mathcal{A}_h^2 Z(0).$$

Now assume equation (B.13) holds for $m \geq 2$. Then we have by another use of equation (B.12)

$$\begin{aligned}
\partial_t^{(m+1)} Z(0) &= \partial_t^{(m)} G(0) - \mathcal{A}_h \partial_t^{(m)} Z(0) \\
&= \partial_t^{(m)} G(0) - \mathcal{A}_h \sum_{i=0}^{m-1} (-1)^{i-1+m} \partial_t^{(i)} \mathcal{A}_h^{m-1-i} G(0) + (-1)^{m+1} \mathcal{A}_h^{m+1} Z(0) \\
&= \sum_{i=0}^{m} (-1)^{i+m} \partial_t^{(i)} \mathcal{A}_h^{m-i} G(0) + (-1)^{m+1} \mathcal{A}_h^{m+1} Z(0)
\end{aligned}$$

and equation (B.13) is shown. Hence, the initial values can be easily bounded by equation (B.13) using the triangle inequality. $\square$