SIAM J. NUMER. ANAL.
Vol. 47, No. 1, pp. 409–439

GALERKIN FINITE ELEMENT METHODS WITH SYMMETRIC PRESSURE STABILIZATION FOR THE TRANSIENT STOKES EQUATIONS: STABILITY AND CONVERGENCE ANALYSIS*

ERIK BURMAN† AND MIGUEL A. FERNÁNDEZ‡

Abstract. We consider the stability and convergence analysis of pressure stabilized finite element approximations of the transient Stokes equation. The analysis is valid for a class of symmetric pressure stabilization operators, but also for standard, inf-sup stable, velocity/pressure spaces without stabilization. Provided the initial data are chosen as a specific (method-dependent) Ritz-projection, we get unconditional stability and optimal convergence for both pressure and velocity approximations, in natural norms. For arbitrary interpolations of the initial data, a condition between the space and time discretization parameters has to be verified in order to guarantee pressure stability.

Key words. transient Stokes equations, finite element methods, symmetric pressure stabilization, time discretization, Ritz-projection

AMS subject classifications. 65N12, 65N30, 76M10, 76D07

DOI. 10.1137/070707403

1. Introduction.

In this paper we consider stabilized finite element methods for the transient Stokes problem. For methods of standard pressure stabilized Petrov–Galerkin (PSPG) or Galerkin least squares (GLS) type, the analysis of time-discretization schemes is a difficult issue, unless a space-time approach is applied with a discontinuous Galerkin discretization in time. Indeed, for standard finite difference type time discretizations, the finite difference term must be included in the stabilization operator to ensure consistency (see, e.g., [11, 23]). It has been shown in [3] that even for first order backward difference (BDF1) schemes this perturbs the stability of the numerical scheme when the time step is small, unless the following condition between the space mesh size and the time step is verified:

$$\delta t \geq Ch^2, \tag{1.1}$$

where δt denotes the time step and h the space discretization parameter. For higher order schemes, such as Crank–Nicholson or second order backward differencing, the strongly consistent scheme appears to be unstable (see, e.g., [1]). Similar initial time-step instabilities were observed in [19] for the algebraic (static) subscale stabilization scheme applied to the Navier–Stokes equations, and they were cured by including time dependent subscales.

Our goal in this work is to consider a fairly large class of pressure stabilization methods and show that convergence of velocities and pressures, for the transient Stokes problem, can be obtained without conditions on the space- and time-discretization parameters (like (1.1)), provided the initial data are chosen as a specific (method-dependent) Ritz-projection (see, e.g., [33, 34]) onto a space of discretely *divergence-free* functions. *Discretely divergence-free* should here be interpreted

*Received by the editors November 6, 2007; accepted for publication (in revised form) June 23, 2008; published electronically December 5, 2008.
http://www.siam.org/journals/sinum/47-1/70740.html
†Department of Mathematics, University of Sussex, Brighton, BN1 9RF, UK (E.N.Burman@sussex.ac.uk).
‡INRIA, Rocquencourt, B.P. 105, F–78153 Le Chesnay Cedex, France (miguel.fernandez@inria.fr).

in the sense of the stabilized method. If, on the other hand, the initial data are chosen as some interpolant that does not conserve the discrete divergence-free character, the condition

$$\delta t \geq \tilde{C} h^{2k}, \tag{1.2}$$

with k the polynomial degree of the velocity approximation space, has to be respected in order to avoid pressure oscillations in the transient solution for small times.

Although the stability conditions (1.1) and (1.2) are similar, their natures are different. As mentioned above, if (1.2) fails to be satisfied, pressure instabilities appear when dealing with nondiscrete divergence-free initial velocity approximation, but they are not related to the structure of the pressure stabilization. For residual-based stabilization methods (PSPG, GLS, etc.) on the other hand, the finite difference/pressure coupling of the stabilization perturbs the coercivity of the discrete pressure operator (see [3]) unless condition (1.1) is satisfied (irrespective of the divergence-free character of the initial velocity approximation).

The analysis carried out in this paper is valid not only for pressure stabilization operators that are symmetric and weakly consistent but also for standard methods using inf-sup stable velocity/pressure pairs, but it does not apply to residual-based pressure stabilizations (PSPG, GLS, etc.). In particular, space and time discretizations commute (i.e., lead to the same fully discrete scheme) for the methods we analyze.

We prove unconditional stability of velocities and pressures and optimal convergence (in natural norms) when the initial data are chosen as a certain Ritz-type projection. In the case when a standard interpolation of the initial data is applied, an *inverse parabolic Courant–Friedrich–Lewy (CFL)-type* condition must be respected in order to maintain pressure stability for small time steps. We give the full analysis only for the backward difference formula of order one, and we indicate how the analysis changes in the case of second order approximations in time. Indeed, any $\mathcal{A}$-stable implicit scheme is expected to yield optimal performance.

The remainder of the paper is organized as follows. In the next section we introduce the problem under consideration and some useful notation. The space- and time-discretized formulations are introduced in section 3. In subsection 3.1, the space discretization is formulated using a general framework; we also discuss how some known pressure stabilized finite element methods enter this setting. The time discretization is performed in subsection 3.2 using the first order backward difference (BDF1), Crank–Nicholson, and second order backward difference (BDF2) schemes. Section 4 is devoted to the stability analysis of the resulting fully discrete formulations. The convergence analysis for the BDF1 scheme is carried out in section 5. We illustrate the theoretical results with some numerical experiments in section 6, using interior penalty stabilization of the gradient jumps. Finally, some conclusions are given in section 7.

2. Problem setting. Let Ω be a domain in $\mathbb{R}^d$ ($d = 2$ or 3) with a polyhedral boundary $\partial\Omega$. For $T > 0$ we consider the problem of solving, for $\boldsymbol{u} : \Omega \times (0,T) \longrightarrow \mathbb{R}^d$ and $p : \Omega \times (0,T) \longrightarrow \mathbb{R}$, the following time-dependent Stokes problem:

$$\begin{cases} \partial_t \boldsymbol{u} - \nu \Delta \boldsymbol{u} + \boldsymbol{\nabla} p = \boldsymbol{f} & \text{in} \quad \Omega \times (0,T), \\ \boldsymbol{\nabla} \cdot \boldsymbol{u} = 0 & \text{in} \quad \Omega \times (0,T), \\ \boldsymbol{u} = \boldsymbol{0} & \text{on} \quad \partial\Omega \times (0,T), \\ \boldsymbol{u}(\cdot, 0) = \boldsymbol{u}_0 & \text{in} \quad \Omega. \end{cases} \tag{2.1}$$

Here, $\boldsymbol{f} : \Omega \times (0,T) \longrightarrow \mathbb{R}$ stands for the source term, $\boldsymbol{u}_0 : \Omega \longrightarrow \mathbb{R}^d$ for the initial velocity, and $\nu > 0$ for a given constant viscosity. In order to introduce a variational setting for (2.1) we consider the following standard velocity and pressure spaces:

$$V \stackrel{\text{def}}{=} [H_0^1(\Omega)]^d, \quad H \stackrel{\text{def}}{=} [L^2(\Omega)]^d, \quad Q \stackrel{\text{def}}{=} L_0^2(\Omega),$$

normed with

$$\|\boldsymbol{v}\|_H \stackrel{\text{def}}{=} (\boldsymbol{v},\boldsymbol{v})^{\frac{1}{2}}, \quad \|\boldsymbol{v}\|_V \stackrel{\text{def}}{=} \|\nu^{\frac{1}{2}}\boldsymbol{\nabla}\boldsymbol{v}\|_H, \quad \|q\|_Q \stackrel{\text{def}}{=} \|\nu^{-\frac{1}{2}}q\|_H,$$

where $(\cdot,\cdot)$ denotes the standard L^2-inner product in Ω.

Problem (2.1) can be formulated in weak form as follows: For all $t > 0$, find $\boldsymbol{u}(t) \in V$ and $p(t) \in Q$ such that

$$\text{(2.2)} \qquad \begin{cases} (\partial_t \boldsymbol{u},\boldsymbol{v}) + a(\boldsymbol{u},\boldsymbol{v}) + b(p,\boldsymbol{v}) = (\boldsymbol{f},\boldsymbol{v}) & \text{a.e. in} \quad (0,T), \\ b(q,\boldsymbol{u}) = 0 & \text{a.e. in} \quad (0,T), \\ \boldsymbol{u}(\cdot,0) = \boldsymbol{u}_0 & \text{a.e. in} \quad \Omega \end{cases}$$

for all $\boldsymbol{v} \in V$, $q \in Q$ and with

$$a(\boldsymbol{u},\boldsymbol{v}) \stackrel{\text{def}}{=} (\nu\boldsymbol{\nabla}\boldsymbol{u},\boldsymbol{\nabla}\boldsymbol{v}), \quad b(p,\boldsymbol{v}) \stackrel{\text{def}}{=} -(p,\boldsymbol{\nabla}\cdot\boldsymbol{v}).$$

From these definitions, the following classical coercivity and continuity estimates hold:

$$\text{(2.3)} \qquad a(\boldsymbol{v},\boldsymbol{v}) \geq \|\boldsymbol{v}\|_V^2, \quad a(\boldsymbol{u},\boldsymbol{v}) \leq \|\boldsymbol{u}\|_V\|\boldsymbol{v}\|_V, \quad b(\boldsymbol{v},q) \leq \|\boldsymbol{v}\|_V\|q\|_Q$$

for all $\boldsymbol{u},\boldsymbol{v} \in V$ and $q \in Q$. It is known (see, e.g., [22]) that if $\boldsymbol{f} \in C^0([0,T];H)$ and that $\boldsymbol{u}_0 \in V \cap H_0(\text{div};\Omega)$, problem (2.2) admits a unique solution $(\boldsymbol{u},p)$ in $L^2(0,T;V) \times L^2(0,T;Q)$ with $\partial_t\boldsymbol{u} \in L^2(0,T;V')$.

Thoroughout this paper, C stands for a generic positive constant independent of the physical and discretization parameters.

3. Space and time discretization. In this section we discretize problem (2.2) with respect to the space and time variables. Symmetric pressure stabilized finite elements are used for the space discretization (subsection 3.1), and some known $\mathcal{A}$-stable schemes are used for the time discretization (subsection 3.2).

3.1. Space semidiscretization: Symmetric pressure stabilized formulations. Let $\{\mathcal{T}_h\}_{0<h\leq 1}$ denote a shape-regular family of triangulations of the domain Ω. For each triangulation $\mathcal{T}_h$, the subscript $h \in (0,1]$ refers to the level of refinement of the triangulation, which is defined by

$$h \stackrel{\text{def}}{=} \max_{K\in\mathcal{T}_h} h_K,$$

with h_K the diameter of K. In order to simplify the analysis, we assume that the family of triangulations $\{\mathcal{T}_h\}_{0<h\leq 1}$ is quasi uniform. For more precise information on the constraint on the mesh, we refer the reader to the analysis of the various finite element methods in the steady case; see subsection 3.1.1.

In this paper, we let X_h^k and M_h^l denote, respectively, the standard spaces of continuous and (possibly) discontinuous piecewise polynomial functions of degree $k \geq 1$ and $l \geq 0$ $(k-1 \leq l \leq k)$,

$$X_h^k \stackrel{\text{def}}{=} \left\{v_h \in C^0(\overline{\Omega}) \,:\, v_{h|K} \in \mathbb{P}_k(K) \quad \forall K \in \mathcal{T}_h\right\},$$

$$M_h^l \stackrel{\text{def}}{=} \left\{q_h \in L^2(\Omega) \,:\, q_{h|K} \in \mathbb{P}_l(K) \quad \forall K \in \mathcal{T}_h\right\}.$$

For the approximated velocities, we will use the space $[V_h^k]^d \stackrel{\text{def}}{=} [X_h^k \cap H_0^1(\Omega)]^d$, and for the pressure, we will use either $Q_h^l \stackrel{\text{def}}{=} M_h^l \cap L_0^2(\Omega)$ or $Q_h^l \stackrel{\text{def}}{=} M_h^l \cap L_0^2(\Omega) \cap C^0(\overline{\Omega})$. In order to stabilize the pressure we introduce a bilinear form $j : Q_h \times Q_h \longrightarrow \mathbb{R}$ satisfying the following properties:

- Symmetry:

$$j(p_h, q_h) = j(q_h, p_h) \quad \forall p_h, q_h \in Q_h^l; \tag{3.1}$$

- continuity:

$$|j(p_h, q_h)| \le j(p_h, p_h)^{\frac{1}{2}} j(q_h, q_h)^{\frac{1}{2}} \le C\|p_h\|_Q \|q_h\|_Q \quad \forall p_h, q_h \in Q_h^l; \tag{3.2}$$

- weak consistency:

$$j(\Pi_h^l q, \Pi_h^l q)^{\frac{1}{2}} \le C\frac{h^{s_p}}{\nu}\|q\|_{s_p,\Omega} \quad \forall q \in H^s(\Omega), \tag{3.3}$$

  with $s_p \stackrel{\text{def}}{=} \min\{s, \tilde{l}, l+1\}$, $\tilde{l} \ge 1$, denoting the order of weak consistency of the stabilization operator, and $\Pi_h^l : Q \longrightarrow Q_h^l$ a given projection operator such that

$$\|q - \Pi_h^l q\|_Q \le \frac{C}{\nu^{\frac{1}{2}}} h^{l+1}\|q\|_{l+1,\Omega} \tag{3.4}$$

  for all $q \in H^{l+1}(\Omega)$.

Finally, we assume that there exists a projection operator $\mathcal{I}_h^k : V \longrightarrow V_h^k$ satisfying the following approximation properties:

$$\|\boldsymbol{v} - \mathcal{I}_h^k \boldsymbol{v}\|_H + h\nu^{-\frac{1}{2}}\|\boldsymbol{v} - \mathcal{I}_h^k \boldsymbol{v}\|_V \le C_{\mathcal{I}} h^{r_{\boldsymbol{u}}}\|\boldsymbol{v}\|_{r_{\boldsymbol{u}},\Omega}, \tag{3.5}$$

$$|b(q_h, \boldsymbol{v} - \mathcal{I}_h^k \boldsymbol{v})| \le C j(q_h, q_h)^{\frac{1}{2}} \left(\nu^{\frac{1}{2}}\|h^{-1}(\boldsymbol{v} - \mathcal{I}_h^k \boldsymbol{v})\|_H + \|\boldsymbol{v} - \mathcal{I}_h^k \boldsymbol{v}\|_V\right) \tag{3.6}$$

for all $\boldsymbol{v} \in [H^r(\Omega)]^d$, $r_{\boldsymbol{u}} = \min\{r, k+1\}$, and $(q_h, \boldsymbol{v}_h) \in Q_h^l \times [V_h^k]^d$.

Our space semidiscretized scheme reads as follows: For all $t \in (0, T)$, find $(\boldsymbol{u}_h(t), p_h(t)) \in [V_h^k]^d \times Q_h^l$ such that

$$\begin{aligned} (\partial_t \boldsymbol{u}_h, \boldsymbol{v}_h) + a(\boldsymbol{u}_h, \boldsymbol{v}_h) + b(p_h, \boldsymbol{v}_h) - b(q_h, \boldsymbol{u}_h) + j(p_h, q_h) &= (\boldsymbol{f}, \boldsymbol{v}_h), \\ \boldsymbol{u}_h(0) &= \boldsymbol{u}_h^0, \end{aligned} \tag{3.7}$$

for all $(\boldsymbol{v}_h, q_h) \in [V_h^k]^d \times Q_h^l$ and with $\boldsymbol{u}_h^0$ a suitable approximation of $\boldsymbol{u}_0$ in $[V_h^k]^d$.

The following modified inf-sup condition states the stability of the discrete pressures in (3.7).

Lemma 3.1. *There exists two constants $C, \beta > 0$, independent of h and ν, such that*

$$\sup_{\boldsymbol{v}_h \in [V_h^k]^d} \frac{|b(q_h, \boldsymbol{v}_h)|}{\|\boldsymbol{v}_h\|_V} + C j(q_h, q_h)^{\frac{1}{2}} \ge \beta\|q_h\|_Q \tag{3.8}$$

for all $q_h \in Q_h^l$.

Proof. Let $q_h \in Q_h^l$; from [25, Corollary 2.4] and (3.5) there exists $\boldsymbol{v}_q \in H_0^1(\Omega)$ such that $\boldsymbol{\nabla} \cdot \boldsymbol{v}_q = \nu^{-1} q_h$ and

$$\|\mathcal{I}_h^k \boldsymbol{v}_q\|_V \le C\|\boldsymbol{v}_q\|_V \le C\|q_h\|_Q. \tag{3.9}$$

On the other hand, using (3.6), we have

$$\begin{aligned}\|q_h\|_Q^2 &= b(q_h, \boldsymbol{v}_q)\\ &= b(q_h, \boldsymbol{v}_q - \mathcal{I}_h^k \boldsymbol{v}_q) + b(q_h, \mathcal{I}_h^k \boldsymbol{v}_q)\\ &\leq C j(q_h, q_h)^{\frac{1}{2}} \left(\|\nu^{\frac{1}{2}} h^{-1} (\boldsymbol{v} - \mathcal{I}_h^k \boldsymbol{v})\|_H + \|\boldsymbol{v} - \mathcal{I}_h^k \boldsymbol{v}\|_V \right) + b(q_h, \mathcal{I}_h^k \boldsymbol{v}_q)\\ &\leq C j(q_h, q_h)^{\frac{1}{2}} \|q_h\|_Q + b(q_h, \mathcal{I}_h^k \boldsymbol{v}_q).\end{aligned}$$

We conclude the proof by dividing this last inequality by $\|\mathcal{I}_h^k \boldsymbol{v}_q\|_V$ and using (3.9). ∎

The above lemma ensures the well-posedness of problem (3.7). This is stated in the following theorem.

THEOREM 3.2. *The discrete problem* (3.7) *with* $\boldsymbol{u}_h^0 \in V_{h,k}^{\mathrm{div}} \stackrel{\mathrm{def}}{=} \{\boldsymbol{v}_h \in V_j^k : b(q_h, \boldsymbol{v}_h) = 0$ *for all* $q_h \in Q_h \cap \operatorname{Ker} j\}$ *has a unique solution* $(\boldsymbol{u}_h, p_h) \in C^1\big((0,T]; [V_h^k]^d\big) \times C^0\big((0,T]; Q_h^k\big)$.

To facilitate the analysis we introduce the following (mesh-dependent) seminorm, which is a norm for the velocity and a seminorm for the pressure:

$$|||(\boldsymbol{v}_h, q_h)|||_h^2 \stackrel{\mathrm{def}}{=} \|\boldsymbol{v}_h\|_V^2 + j(q_h, q_h). \tag{3.10}$$

Remark 3.3. If the velocity/pressure finite element pair V_h^k/Q_h^k is inf-sup stable, we can take $j(\cdot,\cdot) \stackrel{\mathrm{def}}{=} 0$ in (3.7), as usual. Obviously, this choice is compatible with hypothesis (3.1)–(3.3) so that the results of this paper still apply. In particular, the relation (3.8) becomes the standard inf-sup condition between V_h^k and Q_h^k.

3.1.1. Examples. In this section we will review some of the most well-known pressure projection stabilization methods and discuss how they enter the abstract framework of the previous subsection. For detailed results on analysis for the respective methods, we refer the reader to the references considering the stationary case.

Recently, several different weakly consistent symmetric pressure stabilized finite element methods have been proposed. These methods take their origin from the works of Silvester [32] and Codina and Blasco [17]. Further developments include the work by Becker and Braack [2] on local projection schemes; the extension of the interior penalty method, using penalization of gradient jumps, to the case of pressure stabilization by Burman and Hansbo [14]; and the interpretation of these methods as minimal stabilization procedures by Brezzi and Fortin [9]. Similar approaches have been advocated in Dohrmann and Bochev in [21], and a review of the analysis (with special focus on discontinuous pressure spaces and the Darcy problem) is given in [12].

The main idea underpinning all these methods is that, when using a velocity-pressure space pair $V_h \times Q_h$, the inf-sup stability constraint on the spaces may be relaxed by the addition of an operator penalizing the difference between the discrete pressure variable and its projection onto a subspace $\tilde{Q}_h \subset Q_h$, such that $V_h \times \tilde{Q}_h$ is inf-sup stable. The penalization may either act directly on the pressure, as in [21, 12], or on the gradient of the pressure, as in [2, 18, 14]. Generally speaking, the pressure approximation properties of the numerical scheme will be given by $\tilde{Q}_h$, expressed in the weak consistency satisfied by the penalty operator. For the Oseen's problem, some of these methods may be extended to include high Reynolds number effects (see, e.g., [13, 6, 16]). The advantages and disadvantages of symmetric weakly consistent pressure stabilization methods compared to GLS or PSPG approaches is discussed in a recent review paper [7].

The methods of Brezzi and Pitkäranta, Silvester, and Dohrmann and Bochev. The original pressure stabilized finite element method was proposed by Brezzi and Pitkäranta in [10]. Here, the velocity and pressure discrete spaces are chosen as the standard finite element space of piecewise affine continuous functions, $[V_h^1]^d \times Q_h^1$. The operator $j(\cdot,\cdot)$ is given by

$$j(p_h, q_h) = \left(\frac{h^2}{\nu}\boldsymbol{\nabla} p_h, \boldsymbol{\nabla} q_h\right). \tag{3.11}$$

A variant of this method was recently proposed by Dohrmann and Bochev in [21], using an equivalent stabilization operator, namely,

$$j(p_h, q_h) = \left(\frac{1}{\nu}(I-\pi_0)p_h, (I-\pi_0)q_h\right), \tag{3.12}$$

where $\pi_0 : Q \longrightarrow Q_h^0$ denotes the (elementwise) projection onto piecewise constants. Property (3.6) is verified after an integration by parts, with $\mathcal{I}_h^1$ simply the Scott–Zhang interpolant onto $[V_h^1]^d$ (see, e.g., [31, 22]),

$$\begin{aligned} b(q_h, \boldsymbol{v} - \mathcal{I}_h^1\boldsymbol{v}) &= (\boldsymbol{\nabla} q_h, \boldsymbol{v} - \mathcal{I}_h^1\boldsymbol{v}) \\ &\leq j(q_h,q_h)^{\frac12}\left(h^{-1}\nu^{\frac12}\|\boldsymbol{v} - \mathcal{I}_h^1\boldsymbol{v}\|_H + \|\boldsymbol{v} - \mathcal{I}_h^1\boldsymbol{v}\|_V\right). \end{aligned}$$

One readily verifies that (3.2) and (3.3) hold. Moreover, in both cases (3.11) and (3.12), the weak consistency property holds (with $\tilde{l} = 1$),

$$j(\Pi_h^1 p, \Pi_h^1 p)^{\frac12} \leq \frac{C}{\nu^{\frac12}} h \|p\|_{1,\Omega},$$

with Π_h^1 being, for instance, the L^2-projection onto Q_h^1 (we could use instead the Clément [15] or Scott–Zhang interpolants). Indeed, for (3.11) we apply the H^1-stability of the L^2-projection (see, e.g., [22, 20, 8, 5]), whereas for (3.12) we add and subtract suitable terms (p and $\pi_0 p$) and use the approximation properties of π_0 and Π_h^1 (see, e.g., [22]). As a result, our analysis for the time discretization is valid.

Another low order scheme, covered by the analysis, is the method which consists of using piecewise affine continuous velocities and elementwise constants pressures, $[V_h^1]^d \times Q_h^0$; see, e.g., [27]. Stability is obtained by the addition of the jump over element faces of the discontinuous pressure, namely,

$$j(p_h, q_h) = \sum_{K\in\mathcal{T}_h} \int_{\partial K \setminus \partial\Omega} \frac{h}{\nu} [\![p_h]\!][\![q_h]\!].$$

Here, $[\![q_h]\!]$ denotes the jump of q_h over the interelement boundary, defined by

$$[\![q_h]\!](\boldsymbol{x}) \stackrel{\text{def}}{=} \lim_{\epsilon\longrightarrow 0}\left(q_h(\boldsymbol{x}+\epsilon\boldsymbol{n}) - q_h(\boldsymbol{x}-\epsilon\boldsymbol{n})\right) \quad \forall \boldsymbol{x} \in F,$$

with $\boldsymbol{n}$ standing for a fixed, but arbitrary, normal to the internal face F. In this case, (3.6) is obtained after an integration by parts in the term $b(q_h, \boldsymbol{v} - \mathcal{I}_h^1\boldsymbol{v})$ and an elementwise trace inequality (see, e.g., [22]),

$$\begin{aligned} b(q_h, \boldsymbol{v} - \mathcal{I}_h^1\boldsymbol{v}) &= -\sum_K \int_{\partial K\setminus\partial\Omega} [\![q_h]\!](\boldsymbol{v} - \mathcal{I}_h^1\boldsymbol{v})\cdot\boldsymbol{n} \\ &\leq j(q_h,q_h)^{\frac12}\left(h^{-1}\nu^{\frac12}\|\boldsymbol{v} - \mathcal{I}_h^1\boldsymbol{v}\|_H + \|\boldsymbol{v} - \mathcal{I}_h^1\boldsymbol{v}\|_V\right). \end{aligned}$$

In addition, by taking, for instance, Π_h^0 as the L^2-projection onto Q_h^0, using an elementwise trace inequality and the approximation properties of Π_h^0 (see, e.g., [22]), one also easily shows that the weak consistency property holds,

$$j(\Pi_h^0 p, \Pi_h^0 p)^{\frac{1}{2}} = j\big((I - \Pi_h^0)p, (I - \Pi_h^0)p\big)^{\frac{1}{2}} \leq \frac{C}{\nu^{\frac{1}{2}}} h \|p\|_{1,\Omega},$$

and hence $\tilde{l} = 1$.

For details on the cases of stabilization of the pressure jumps only in macroelements, or the generalization to higher order finite element spaces of the Taylor–Hood family with discontinuous pressures, we refer the reader to [12].

Orthogonal subscale stabilization. The orthogonal subscale stabilization was proposed by Codina and Blasco in [17]. Equal order ($k = l \geq 1$) continuous approximation spaces are used for the velocities and the pressures.

Here the main idea is to penalize the difference between the pressure gradient and its projection onto the finite element space. This imposes the introduction of an auxiliary variable for the projection since it may not be localized and is given only implicitly. Hence, the stabilization operator is given by

$$j(p_h, q_h) = \left(\frac{h^2}{\nu} (\boldsymbol{\nabla} p_h - \pi_h^k \boldsymbol{\nabla} p_h), \boldsymbol{\nabla} q_h \right),$$

where $\pi_h^k : [L^2(\Omega)]^d \longrightarrow [V_h^k]^d$ stands for the L^2-projection onto $[V_h^k]^d$, which is given as the solution of the (global) problem

$$(\pi_h^k \boldsymbol{\nabla} p_h, \boldsymbol{\xi}_h) = (\boldsymbol{\nabla} p_h, \boldsymbol{\xi}_h) \quad \forall \boldsymbol{\xi}_h \in [V_h^k]^d.$$

One may readily show that (3.2) and (3.3) hold. Disregarding for simplicity the boundary conditions, the projection operator $\mathcal{I}_h^k = \pi_h^k$ of (3.6) is here chosen also as the L^2-projection onto $[V_h^k]^d$. This can be justified if boundary conditions are imposed weakly, for instance, using Nitsche's method (see [29, 24]), and V_h^k includes the degrees of freedom on the boundary. Indeed, then we have

$$\begin{aligned} b(q_h, \boldsymbol{v} - \mathcal{I}_h^k \boldsymbol{v}) &= (\boldsymbol{\nabla} q_h - \pi_h^k \boldsymbol{\nabla} q_h, \boldsymbol{v} - \mathcal{I}_h^k \boldsymbol{v}) \\ &\leq j(q_h, q_h)^{\frac{1}{2}} \left(h^{-1} \nu^{\frac{1}{2}} \|\boldsymbol{v} - \mathcal{I}_h^k \boldsymbol{v}\|_H + \|\boldsymbol{v} - \mathcal{I}_h^k \boldsymbol{v}\|_V \right). \end{aligned} \tag{3.13}$$

Finally, by taking $\Pi_h^k : Q \longrightarrow Q_h^k$ as the L^2-projection onto Q_h^k, adding and subtracting suitable terms ($\boldsymbol{\nabla} p$ and $\pi_h^k \boldsymbol{\nabla} p$), and using the approximation properties of π_h^k and Π_h^k (see, e.g., [22]), one readily verifies the weak consistency

$$j(\Pi_h^k p, \Pi_h^k p)^{\frac{1}{2}} = \frac{h}{\nu^{\frac{1}{2}}} \big\| (I - \pi_h^k) \boldsymbol{\nabla} \Pi_h^k p \big\|_{0,\Omega} \leq \frac{C}{\nu^{\frac{1}{2}}} h^{s_p} \|p\|_{s_p,\Omega},$$

for all $p \in H^s(\Omega)$ and with $s_p = \min\{k+1, s\}$. In particular, $\tilde{l} = l = k$. The above analysis is hence valid also in this case (with some modifications of a technical nature due to the weakly imposed boundary conditions).

Local projection stabilization. In the local projection stabilization proposed in [2], stability is obtained by penalizing the projection of the gradient onto piecewise discontinuous functions defined on patches consisting of several elements, obtained by using hierarchic meshes, or by penalizing the gradient of the difference of the pressure and its projection on polynomials of lower polynomial order. The construction relies on the inf-sup stability of a velocity/pressure pair typically of mini-element character or of the Taylor–Hood family. Similar ideas were advocated in [21]. The stabilization operator is written as

$$j(p_h, q_h) = \sum_{\tilde{K}} \left(\frac{h^2}{\nu} \kappa \boldsymbol{\nabla} p_h, \kappa \boldsymbol{\nabla} q_h \right),$$

where κ is the so-called *fluctuation operator* defined as $\kappa \stackrel{\text{def}}{=} I - \tilde{\pi}_h$, where $\tilde{\pi}_h$ denotes a *local* projection operator onto either a polynomial of order k on a macropatch consisting of three triangles (or four quadrilaterals) or a polynomial of order $k-1$ on the element. One may show that (3.6), (3.2), and (3.3) hold (for details on the construction of $\mathcal{I}_h^k$, see [2, 6], and for general conditions on the finite element spaces and stabilization operators, see [28]). In the case when we consider the projection $\tilde{\pi}_h$ onto polynomials of order $k-1$, the stabilization operator may be written as

$$j(p_h, q_h) = \sum_{\tilde{K}} \left(\frac{h^2}{\nu} \boldsymbol{\nabla}(\kappa p_h), \boldsymbol{\nabla}(\kappa q_h) \right), \tag{3.14}$$

or, equivalently, following [21], as

$$j(p_h, q_h) = \sum_{\tilde{K}} \left(\frac{1}{\nu} \kappa\, p_h, \kappa\, q_h \right).$$

In these latter cases, condition (3.6) is obtained by choosing $\mathcal{I}_h^k$ as the Fortin interpolation operator associated with $[V_h^k]^d \times \tilde{Q}_h$, where $\tilde{Q}_h$ is the space of continuous piecewise polynomial functions of order $k-1$. Clearly, we then have

$$\begin{aligned} b(q_h, \boldsymbol{v} - \mathcal{I}_h^k \boldsymbol{v}) &= b(\kappa q_h, \boldsymbol{v} - \mathcal{I}_h^k \boldsymbol{v}) \\ &\le j(q_h, q_h)^{\frac{1}{2}} \left(h^{-1} \nu^{\frac{1}{2}} \|\boldsymbol{v} - \mathcal{I}_h^k \boldsymbol{v}\|_H + \|\boldsymbol{v} - \mathcal{I}_h^k \boldsymbol{v}\|_V \right), \end{aligned}$$

since $b(\tilde{q}_h, \boldsymbol{v} - \mathcal{I}_h^k \boldsymbol{v}) = 0$ for all $\tilde{q}_h \in \tilde{Q}_h$. The form (3.14) is treated in a similar fashion after an integration by parts. On the other hand, by taking $\Pi_h^l : Q \longrightarrow Q_h^l$ as the L^2-projection operator onto Q_h^l and using approximation properties of Π_h^l and $\tilde{\pi}_h$, we have

$$\begin{aligned} j(\Pi_h^l p, \Pi_h^l p)^{\frac{1}{2}} &\le j\big((I - \Pi_h^l)p, (I - \Pi_h^l)p\big)^{\frac{1}{2}} + j(p, p)^{\frac{1}{2}} \\ &\le \frac{C}{\nu^{\frac{1}{2}}} h^{s_p} \|p\|_{s_p, \Omega} \quad \forall p \in H^s(\Omega), \end{aligned}$$

where $s_p = \min\{\tilde{l}, s, l+1\}$ and $\tilde{l} - 1$ denotes the polynomial order of the space on which the local projection is taken. Clearly, if we project on polynomials of order $k-1$, the stabilization operator loses one order in the weak consistency; however, the estimates remain optimal since we expect the velocities to be one order more regular than the pressure.

Continuous interior penalty (CIP) stabilization. The CIP stabilization for the stationary Stokes problem was proposed in [14] and generalized to Oseen's problem in [13]. It uses equal order continuous approximation spaces for velocities and pressures ($k = l \geq 1$) and relies on the fact that the component of the pressure gradient orthogonal to the finite element space may be controlled by the gradient jumps using an interpolation estimate between discrete spaces. Indeed, it was shown in [13] that the following inequality holds:

$$\|h(\nabla p_h - \tilde{i}\nabla p_h)\|_H^2 \leq \sum_{K\in\mathcal{T}_h} \int_{\partial K\setminus\partial\Omega} h_K^3 [\![\nabla p_h \cdot \boldsymbol{n}]\!]^2 \tag{3.15}$$

for a certain Clément-type quasi-interpolation operator $\tilde{i}$. This motivates the use of the pressure stabilization operator

$$j(p_h, q_h) = \sum_{K\in\mathcal{T}_h} \int_{\partial K\setminus\partial\Omega} \frac{h^3}{\nu} [\![\nabla p_h \cdot \boldsymbol{n}]\!][\![\nabla q_h \cdot \boldsymbol{n}]\!].$$

Clearly (3.2) and (3.3) are verified in this case. Moreover, (3.6) may be shown to hold if $\mathcal{I}_h^k$ is chosen to be the L^2-projection onto $[V_h^k]^d$ and boundary conditions are imposed weakly [13]. To show the inequality we combine (3.13) with (3.15). Finally, by taking Π_h^k as the L^2-projection onto Q_h^k, since $[\![\mathcal{C}_h^k \nabla p]\!] = \mathbf{0}$ (with $\mathcal{C}_h^k$ the Clément interpolant onto $[X_h^k]^d$), using an elementwise trace inequality, adding and subtracting ∇p, and using the approximation properties of $\mathcal{C}_h^k$ and Π_h^k , one readily verifies (see [13, Lemma 4.7]) the weak consistency

$$\begin{aligned} j(\Pi_h^k p, \Pi_h^k p)^{\frac{1}{2}} &\leq C \frac{h}{\nu^{\frac{1}{2}}} \|\nabla \Pi_h^k p - \mathcal{C}_h^k \nabla p\|_{0,\Omega} \\ &\leq \frac{C}{\nu^{\frac{1}{2}}} h^{s_p} \|p\|_{s_p,\Omega} \quad \forall p \in H^s(\Omega), \end{aligned}$$

with $s_p = \min\{k+1, s\}$, so that $\tilde{l} = l = k$.

We refer the reader to [13] for the details on the technical issue related to the weak imposition of the boundary conditions using Nitsche's method.

3.1.2. The Ritz-projection operator. For the purpose of the stability and convergence analysis below we introduce the Ritz-projection operator

$$S_h^{k,l} : [H^1(\Omega)]^d \times L^2(\Omega) \longrightarrow V_h^k \times Q_h^l.$$

For each $(\boldsymbol{u}, p) \in [H^1(\Omega)]^d \times L^2(\Omega)$, the projection $S_h^{k,l}(\boldsymbol{u}, p) \stackrel{\text{def}}{=} \big(P_h^k(\boldsymbol{u}, p), R_h^l(\boldsymbol{u}, p)\big) \in [V_h^k]^d \times Q_h^l$ is defined as the unique solution of

$$\begin{cases} a(P_h^k(\boldsymbol{u}, p), \boldsymbol{v}_h) + b(R_h^l(\boldsymbol{u}, p), \boldsymbol{v}_h) = a(\boldsymbol{u}, \boldsymbol{v}_h) + b(p, \boldsymbol{v}_h), \\ -b(q_h, P_h^k(\boldsymbol{u}, p)) + j(R_h^l(\boldsymbol{u}, p), q_h) = 0 \end{cases} \tag{3.16}$$

for all $(\boldsymbol{v}_h, q_h) \in [V_h^k]^d \times Q_h^l$.

Problem (3.16) is well-posed thanks to the inf-sup condition (3.8); in particular, we have the following a priori stability estimate:

$$|\!|\!| \big(P_h^k(\boldsymbol{u}, p), R_h^l(\boldsymbol{u}, p)\big) |\!|\!|_h^2 \leq C \left(\|\boldsymbol{u}\|_V^2 + \|p\|_Q^2\right), \tag{3.17}$$

with $C > 0$ a constant independent of h and ν.

Finally, we have the following approximation result.

LEMMA 3.4. *Let $(\boldsymbol{u}, p) \in C^1([0,T], [H^r(\Omega) \cap H_0^1(\Omega))]^d \cap H_0(\mathrm{div};\Omega) \times H^s(\Omega))$ with $r \geq 2$ and $s \geq 1$. The following error estimate for the projection $S_h^{k,l}$ holds with $\alpha = 0, 1$:*

$$\left\|\!\left|\left(\partial_t^\alpha(\boldsymbol{u} - P_h^k(\boldsymbol{u},p)), \partial_t^\alpha R_h^l(\boldsymbol{u},p)\right)\right|\!\right\|_h \leq C\left(\nu^{\frac{1}{2}} h^{r_{\boldsymbol{u}}-1}\|\partial_t^\alpha \boldsymbol{u}\|_{r_{\boldsymbol{u}},\Omega} + \nu^{-\frac{1}{2}} h^{s_p}\|\partial_t^\alpha p\|_{s_p,\Omega}\right),$$

$$\|p - R_h^l(\boldsymbol{u},p)\|_Q \leq C\left(\nu^{\frac{1}{2}} h^{r_{\boldsymbol{u}}-1}\|\boldsymbol{u}\|_{r_{\boldsymbol{u}},\Omega} + \nu^{-\frac{1}{2}} h^{s_p}\|p\|_{s_p,\Omega}\right)$$

for all $t \in [0,T]$ and with $r_{\boldsymbol{u}} \stackrel{\text{def}}{=} \min\{r, k+1\}$ and $s_p \stackrel{\text{def}}{=} \min\{s, \tilde{l}, l+1\}$, and $C > 0$ independent of ν and h. Moreover, provided the domain Ω is sufficiently smooth and, if $\tilde{l} \geq 1$, there also holds

$$\|\partial_t^\alpha(\boldsymbol{u} - P_h^k(\boldsymbol{u},p))\|_H \leq Ch\left\|\!\left|(\partial_t^\alpha(\boldsymbol{u} - P_h^k(\boldsymbol{u},p)), \partial_t^\alpha R_h^l p)\right|\!\right\|_h. \tag{3.18}$$

Proof. For simplicity we here use the notation $\boldsymbol{u}_h \stackrel{\text{def}}{=} P_h^k(\boldsymbol{u},p)$ and $p_h \stackrel{\text{def}}{=} R_h^l(\boldsymbol{u},p)$. From (3.10), the V-coercivity of $a(\cdot,\cdot)$ (see (2.3)), and the orthogonality provided by (3.16), we have

$$\begin{aligned}\left\|\!\left|\left(\boldsymbol{u}_h - \mathcal{I}_h^k\boldsymbol{u}, p_h - \Pi_h^l p\right)\right|\!\right\|_h^2 =& a(\boldsymbol{u} - \mathcal{I}_h^k\boldsymbol{u}, \boldsymbol{u}_h - \mathcal{I}_h^k\boldsymbol{u}) + b(p - \Pi_h^l p, \boldsymbol{u}_h - \mathcal{I}_h^k\boldsymbol{u})\\ &+ b(p_h - \Pi_h^l p, \boldsymbol{u} - \mathcal{I}_h^k\boldsymbol{u}) + j(\Pi_h^l p, p_h - \Pi_h^l p).\end{aligned}$$

Finally, using (2.3) and (3.6), we have that

$$\begin{aligned}&\left\|\!\left|\left(\boldsymbol{u}_h - \mathcal{I}_h^k\boldsymbol{u}, p_h - \Pi_h^l p\right)\right|\!\right\|_h^2 \leq \left(\|\boldsymbol{u} - \mathcal{I}_h^k\boldsymbol{u}\|_V + \|p - \Pi_h^l p\|_Q\right)\|\boldsymbol{u}_h - \mathcal{I}_h^k\boldsymbol{u}\|_V\\ &+ C\left(\nu^{\frac{1}{2}}\|h^{-1}(\boldsymbol{u} - \mathcal{I}_h^k\boldsymbol{u})\|_H + \|\boldsymbol{u} - \mathcal{I}_h^k\boldsymbol{u}\|_V + j(\Pi_h^l p, \Pi_h^l p)^{\frac{1}{2}}\right) j(p_h - \Pi_h^l p, p_h - \Pi_l^k p)^{\frac{1}{2}}.\end{aligned}$$

We obtain the estimation for the velocity ($\alpha = 0$) using the approximation properties of $\mathcal{I}_h^k$ and Π_h^l (see (3.5) and (3.4)) and the weak consistency (3.3) of the stabilizing term $j(\cdot,\cdot)$. The convergence for the time derivative ($\alpha = 1$) is obtained in a similar fashion after the time derivation of (3.16).

For the pressure estimate, we use the generalized inf-sup condition (3.8) and the orthogonality provided by (3.16). We then have

$$\begin{aligned}&\beta\|\Pi_h^l p - p_h\|_Q\\ &\quad\leq \sup_{\boldsymbol{v}_h \in [V_h^k]^d} \frac{b(\Pi_h^l p - p_h, \boldsymbol{v}_h)}{\|\boldsymbol{v}_h\|_V} + Cj(\Pi_h^l p - p_h, \Pi_h^l p - p_h)^{\frac{1}{2}}\\ &\quad\leq \sup_{\boldsymbol{v}_h \in [V_h^k]^d} \frac{b(\Pi_h^k p - p, \boldsymbol{v}_h) - a(\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{v}_h)}{\|\boldsymbol{v}_h\|_V} + Cj(\Pi_h^l p - p_h, \Pi_h^l p - p_h)^{\frac{1}{2}}.\end{aligned}$$

We conclude by using the continuity of $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$, approximability, the weak consistency of $j(\cdot,\cdot)$, and the previous error estimate. For a proof of the optimality in the H-norm, see, e.g., [13, Theorem 4.14]. ∎

3.2. Fully discrete formulation: Time discretization. In this subsection we discretize (3.7) with respect to the time variable. To this end, we will use some known $\mathcal{A}$-stable time discretization schemes for ODEs.

Let $N \in \mathbb{N}^*$ be given. We consider a uniform partition $\{[t_n, t_{n+1}]\}_{0 \leq n \leq N-1}$, with $t_n \stackrel{\text{def}}{=} n\delta t$, of the time interval of interest $[0, T]$ with time-step size $\delta t \stackrel{\text{def}}{=} T/N$. The discrete pair $(\boldsymbol{u}_h^n, p_h^n)$ stands for an approximation of $(\boldsymbol{u}(t_n), p(t_n))$ in $[V_h^k]^d \times Q_h^l$.

First order backward difference formula (BDF1). By introducing the first order backward difference quotient

$$\bar{D}u_h^{n+1} \stackrel{\text{def}}{=} \frac{u_h^{n+1} - u_h^n}{\delta t},$$

our first fully discrete scheme reads as follows: For $0 \leq n \leq N-1$, find $(\boldsymbol{u}_h^{n+1}, p_h^{n+1}) \in [V_h^k]^d \times Q_h^l$ such that

$$\begin{aligned}(\bar{D}\boldsymbol{u}_h^{n+1}, \boldsymbol{v}_h) + a(\boldsymbol{u}_h^{n+1}, \boldsymbol{v}_h) + b(p_h^{n+1}, \boldsymbol{v}_h) - b(q_h, \boldsymbol{u}_h^{n+1})& \\ + j(p_h^{n+1}, q_h) = (\boldsymbol{f}(t_{n+1}), \boldsymbol{v}_h)&\end{aligned} \tag{3.19}$$

for all $(\boldsymbol{v}_h, q_h) \in V_h^k \times Q_h^l$ and with $\boldsymbol{u}_h^0$ a suitable approximation of $\boldsymbol{u}_0$ in $[V_h^k]^d$.

Crank–Nicholson scheme. Let us consider now the scheme given by the following: For $0 \leq n \leq N-1$, find $(\boldsymbol{u}_h^{n+1}, p_h^{n+\frac{1}{2}}) \in [V_h^k]^d \times Q_h^l$ such that

$$\begin{aligned}(\bar{D}\boldsymbol{u}_h^{n+1}, \boldsymbol{v}_h) + a(\boldsymbol{u}_h^{n+\frac{1}{2}}, \boldsymbol{v}_h) + b(p_h^{n+\frac{1}{2}}, \boldsymbol{v}_h) - b(q_h, \boldsymbol{u}_h^{n+\frac{1}{2}})& \\ + j(p_h^{n+\frac{1}{2}}, q_h) = (\boldsymbol{f}^{n+\frac{1}{2}}, \boldsymbol{v}_h)&\end{aligned} \tag{3.20}$$

for all $(\boldsymbol{v}_h, q_h) \in [V_h^k]^d \times Q_h^l$, where $\boldsymbol{u}_h^{n+\frac{1}{2}} \stackrel{\text{def}}{=} \frac{1}{2}(\boldsymbol{u}_h^{n+1} + \boldsymbol{u}_h^n)$ and $\boldsymbol{u}_h^0$ is a suitable approximation of $\boldsymbol{u}_0$ in $[V_h^k]^d$.

Remark 3.5. Note that (3.20) uniquely determines $\boldsymbol{u}_h^{n+1}$, since $\boldsymbol{u}_h^0$ is given. For the pressure, however, neither p_h^{n+1} nor p_h^n is used in (3.20). Therefore, by working with $p_h^{n+\frac{1}{2}}$ as the pressure variable, we do not need to provide an initial condition for the pressure. On the other hand, we do not have an approximation of p_h^{n+1} unless one is constructed by extrapolation.

Second order backward difference (BDF2). Finally, by considering the second order backward difference quotient

$$\tilde{D}\boldsymbol{u}^{n+1} \stackrel{\text{def}}{=} \frac{1}{2\delta t}(3\boldsymbol{u}_h^{n+1} - 4\boldsymbol{u}_h^n + \boldsymbol{u}_h^{n-1}),$$

we obtain the following BDF2 scheme: For $1 \leq n \leq N-1$, find $(\boldsymbol{u}_h^{n+1}, p_h^{n+1}) \in [V_h^k]^d \times Q_h^l$ such that

$$\begin{aligned}(\tilde{D}\boldsymbol{u}_h^{n+1}, \boldsymbol{v}_h) + a(\boldsymbol{u}_h^{n+1}, \boldsymbol{v}_h) + b(p_h^{n+1}, \boldsymbol{v}_h) - b(q_h, \boldsymbol{u}_h^{n+1})& \\ + j(p_h^{n+1}, q_h) = (\boldsymbol{f}^{n+1}, \boldsymbol{v}_h)&\end{aligned} \tag{3.21}$$

for all $(\boldsymbol{v}_h, q_h) \in [V_h^k]^d \times Q_h^l$ and $(\boldsymbol{u}_h^1, p_h^1) \in [V_h^k]^d \times Q_h^l$ given by the first step of backward Euler scheme (3.19).

4. Stability. In this section we analyze the stability properties of the fully discrete schemes introduced in subsection 3.2. For the sake of simplicity, full details will be given only for the backward scheme (3.19). Nevertheless, in subsection 4.2, we will discuss how the results extend to the second order time-stepping schemes Crank–Nicholson and BDF2.

4.1. First order $\boldsymbol{\mathcal{A}}$-stable scheme. The next result provides the unconditional stability of the velocity. It also provides a uniform estimate for the pressure, in terms of the discrete velocity time derivative. Theorem 4.2 points out the role of the initial velocity approximation on the stability of the velocity time derivative approximations. Finally, Corollary 4.3 states the (conditional or unconditional) stability of the pressure, depending on the choice of the initial velocity approximation.

THEOREM 4.1. *Let $\boldsymbol{u}_h^0$ be a given H-stable approximation of $\boldsymbol{u}_0$ in $[V_h^k]^d$, and let $\{(\boldsymbol{u}_h^n, p_h^n)\}_{n=1}^N$ be the solution of the fully discrete problem* (3.19). *The following estimate holds for* $1 \le n \le N$*:*

$$\|\boldsymbol{u}_h^n\|_H^2 + \sum_{m=0}^{n-1} \delta t \|\!|\!|(\boldsymbol{u}_h^{m+1}, p_h^{m+1})|\!|\!|_h^2 \le C\|\boldsymbol{u}_0\|_H^2 + \frac{C_{\mathrm{P}}^2}{\nu}\sum_{m=0}^{n-1}\delta t \|\boldsymbol{f}(t_{m+1})\|_H^2,$$

$$(4.1)\quad \sum_{m=0}^{n-1}\delta t\|p_h^{m+1}\|_Q^2 \le \frac{C}{\beta^2}\sum_{m=0}^{n-1}\delta t\left(|\!|\!|(\boldsymbol{u}_h^{m+1}, p_h^{m+1})|\!|\!|_h^2 + \nu^{-1}\left\|\bar{D}\boldsymbol{u}_h^{m+1}\right\|_H^2 + \nu^{-1}\|\boldsymbol{f}(t_{m+1})\|_H^2\right),$$

with $C_{\mathrm{P}} > 0$ the Poincaré constant.

Proof. Taking $\boldsymbol{v}_h = \boldsymbol{u}_h^{n+1}$ and $q_h = p_h^{n+1}$ in (3.19), using the coercivity of the bilinear form, the Cauchy–Schwarz inequality, and the Poincaré inequality, we have

$$(4.2)\qquad (\bar{D}\boldsymbol{u}_h^{n+1}, \boldsymbol{u}_h^{n+1}) + \frac{1}{2}|\!|\!|(\boldsymbol{u}_h^{n+1}, p_h^{n+1})|\!|\!|_h^2 \le \frac{C_{\mathrm{P}}^2}{2\nu}\|\boldsymbol{f}(t_{n+1})\|_H^2.$$

Now, recalling that

$$(4.3)\qquad (\bar{D}\boldsymbol{u}_h^{n+1}, \boldsymbol{u}_h^{n+1}) = \frac{1}{2}\bar{D}\|\boldsymbol{u}_h^{n+1}\|_H^2 + \frac{1}{2\delta t}\|\boldsymbol{u}_h^{n+1} - \boldsymbol{u}_h^n\|_H^2,$$

we have

$$\bar{D}\|\boldsymbol{u}_h^{n+1}\|_H^2 + |\!|\!|(\boldsymbol{u}_h^{n+1}, p_h^{n+1})|\!|\!|_h^2 \le \frac{C_{\mathrm{P}}^2}{\nu}\|\boldsymbol{f}(t_{n+1})\|_H^2,$$

leading to, after summation over $0 \le m \le n-1$,

$$\|\boldsymbol{u}_h^n\|_H^2 + \sum_{m=0}^{n-1}\delta t|\!|\!|(\boldsymbol{u}_h^{m+1}, p_h^{m+1})|\!|\!|_h^2 \le \|\boldsymbol{u}_h^0\|_H^2 + \frac{C_{\mathrm{P}}^2}{\nu}\sum_{m=0}^{n-1}\delta t\|\boldsymbol{f}(t_{m+1})\|_H^2.$$

For the pressure estimate, from (3.8), (3.19) (with $q_h = 0$) and the Poincaré inequality, we have

$$\beta\|p_h^{n+1}\|_Q \le C\left(|\!|\!|(\boldsymbol{u}_h^{n+1}, p_h^{n+1})|\!|\!|_h + \nu^{-\frac{1}{2}}\|\bar{D}\boldsymbol{u}_h^{n+1}\|_H + \nu^{-\frac{1}{2}}\|\boldsymbol{f}(t_{n+1})\|_H\right),$$

which completes the proof. ∎

The next theorem states some a priori estimates of the approximations of the velocity time derivative.

THEOREM 4.2. *Let* $\{(\boldsymbol{u}_h^n, p_h^n)\}_{n=1}^N$ *be the solution of the fully discrete problem* (3.19).

- *If* $\boldsymbol{u}_0 \in [H^1(\Omega)]^d$ *and* $\boldsymbol{u}_h^0 = P_h^k(\boldsymbol{u}_0, 0)$, *the following estimate holds for* $1 \le n \le N$:

$$\sum_{m=0}^{n-1} \delta t \|\bar{D}\boldsymbol{u}_h^{m+1}\|_H^2 + |\!|\!|(\boldsymbol{u}_h^n, p_h^n)|\!|\!|_h^2 \le C\left(\|\boldsymbol{u}_0\|_V^2 + \sum_{m=0}^{n-1} \delta t \|\boldsymbol{f}(t_{m+1})\|_H^2\right). \tag{4.4}$$

- *If* $\boldsymbol{u}_0 \in [H^r(\Omega) \cap H_0^1(\Omega)]^d \cap H_0(\mathrm{div};\Omega)$, $r \ge 2$, *and* $\boldsymbol{u}_h^0 = \mathcal{I}_h^k \boldsymbol{u}_0$, *the following estimate holds for* $1 \le n \le N$:

$$\begin{aligned}\sum_{m=0}^{n-1} \delta t \|\bar{D}\boldsymbol{u}_h^{m+1}\|_H^2 &+ |\!|\!|(\boldsymbol{u}_h^n, p_h^n)|\!|\!|_h^2 \\ &\le C\left(\|\boldsymbol{u}_0\|_V^2 + \nu h^{2(r_{\boldsymbol{u}}-1)}\|p_h^1\|_Q^2 + \|\boldsymbol{u}_0\|_{r_{\boldsymbol{u}},\Omega}^2 + \sum_{m=0}^{n-1} \delta t \|\boldsymbol{f}(t_{m+1})\|_H^2\right),\end{aligned} \tag{4.5}$$

with $r_{\boldsymbol{u}} \stackrel{\text{def}}{=} \min\{k+1, r\}$.

Proof. For $0 \le n \le N-1$, by taking $\boldsymbol{v}_h = \bar{D}\boldsymbol{u}_h^{n+1}$ and $q_h = 0$ in (3.19) and using the Cauchy–Schwarz inequality, we have

$$\frac{1}{2}\|\bar{D}\boldsymbol{u}_h^{n+1}\|_H^2 + a(\boldsymbol{u}_h^{n+1}, \bar{D}\boldsymbol{u}_h^{n+1}) + b(p_h^{n+1}, \bar{D}\boldsymbol{u}_h^{n+1}) = \frac{1}{2}\|\boldsymbol{f}(t_{n+1})\|_H^2. \tag{4.6}$$

On the other hand, for $1 \le n \le N-1$, testing (3.19) at the time levels n and $n+1$ with $\boldsymbol{v}_h = \boldsymbol{0}$ and $q_h = p_h^{n+1}$, we have

$$\begin{aligned} b(p_h^{n+1}, \boldsymbol{u}_h^{n+1}) &= j(p_h^{n+1}, p_h^{n+1}), \\ b(p_h^{n+1}, \boldsymbol{u}_h^{n}) &= j(p_h^{n}, p_h^{n+1}). \end{aligned} \tag{4.7}$$

Therefore, by subtracting these equalities and using the bilinearity of $j(\cdot,\cdot)$, we obtain

$$b(p_h^{n+1}, \bar{D}\boldsymbol{u}_h^{n+1}) = j(\bar{D}p_h^{n+1}, p_h^{n+1}) \tag{4.8}$$

for $1 \le n \le N-1$. It then follows from (4.6) that

$$\frac{1}{2}\|\bar{D}\boldsymbol{u}_h^{n+1}\|_H^2 + a(\boldsymbol{u}_h^{n+1}, \bar{D}\boldsymbol{u}_h^{n+1}) + j(p_h^{n+1}, \bar{D}p_h^{n+1}) \le \frac{1}{2}\|\boldsymbol{f}(t_{n+1})\|_H^2. \tag{4.9}$$

On the other hand, using the symmetry and bilinearity of $a(\cdot,\cdot)$ and $j(\cdot,\cdot)$, we have

$$\begin{aligned} a(\boldsymbol{u}_h^{n+1}, \bar{D}\boldsymbol{u}_h^{n+1}) &= \frac{1}{2}\bar{D}a(\boldsymbol{u}_h^{n+1}, \boldsymbol{u}_h^{n+1}) + \frac{\delta t}{2} a(\bar{D}\boldsymbol{u}_h^{n+1}, \bar{D}\boldsymbol{u}_h^{n+1}), \\ j(p_h^{n+1}, \bar{D}p_h^{n+1}) &= \frac{1}{2}\bar{D}j(p_h^{n+1}, p_h^{n+1}) + \frac{\delta t}{2} j(\bar{D}p_h^{n+1}, \bar{D}p_h^{n+1}). \end{aligned}$$

Hence,

$$\|\bar{D}\boldsymbol{u}_h^{n+1}\|_H^2 + \bar{D}\big(a(\boldsymbol{u}_h^{n+1}, \boldsymbol{u}_h^{n+1}) + j(p_h^{n+1}, p_h^{n+1})\big) \le \|\boldsymbol{f}(t_{n+1})\|_H^2$$

for $1 \le n \le N-1$. After multiplication by δt and summation over $1 \le m \le n-1$, it follows that

$$\sum_{m=1}^{n-1} \delta t \|\bar{D}\boldsymbol{u}_h^{m+1}\|_H^2 + \|\!|(\boldsymbol{u}_h^n, p_h^n)|\!\|_h^2 \le \|\!|(\boldsymbol{u}_h^1, p_h^1)|\!\|_h^2 + \sum_{m=1}^{n-1} \delta t \|\boldsymbol{f}(t_{m+1})\|_H^2. \tag{4.10}$$

In order to highlight the impact of the initial velocity approximation on the stability of the time derivative, we consider now the first time level ($n=0$) of (3.19). By testing with $\boldsymbol{v}_h = \bar{D}\boldsymbol{u}_h^1$, $q_h = 0$, after multiplication by $2\delta t$ and using the symmetry and bilinearity of $a(\cdot,\cdot)$, we get

$$\delta t \|\bar{D}\boldsymbol{u}_h^1\|_H^2 + a(\boldsymbol{u}_h^1, \boldsymbol{u}_h^1) - a(\boldsymbol{u}_h^0, \boldsymbol{u}_h^0) + 2\delta t b(p_h^1, \bar{D}\boldsymbol{u}_h^1) \le \delta t \|\boldsymbol{f}(t_1)\|_H^2. \tag{4.11}$$

If the initial velocity approximation is given in terms of the Ritz-projection, $\boldsymbol{u}_h^0 = P_h^k(\boldsymbol{u}_0, 0)$ with $\boldsymbol{u}_0 \in [H^1(\Omega)]^d$, by setting $p_h^0 \stackrel{\text{def}}{=} R_h^l(\boldsymbol{u}_0, 0)$ it follows that (4.7) also holds for $n=0$. Therefore,

$$b(p_h^1, \bar{D}\boldsymbol{u}_h^1) = j(\bar{D}p_h^1, p_h^1). \tag{4.12}$$

Thus, from the symmetry and bilinearity of $j(\cdot,\cdot)$ and (4.11), we have

$$\delta t \|\bar{D}\boldsymbol{u}_h^1\|_H^2 + \|\!|(\boldsymbol{u}_h^1, p_h^1)|\!\|_h^2 \le \|\!|(\boldsymbol{u}_h^0, p_h^0)|\!\|_h^2 + \delta t \|\boldsymbol{f}(t_1)\|_H^2. \tag{4.13}$$

Estimate (4.4) is obtained by adding this last inequality to (4.10) and using the stability of the Ritz-projection (3.17), $\|\!|(\boldsymbol{u}_h^0, p_h^0)|\!\|_h^2 \le C\|\boldsymbol{u}_0\|_V^2$.

If the initial velocity approximation is given in terms of a general interpolant, $\boldsymbol{u}_h^0 = \mathcal{I}_h^k \boldsymbol{u}_0$ with $\boldsymbol{u}_0 \in [H^r(\Omega) \cap H_0^1(\Omega)]^d \cap H_0(\text{div};\Omega)$, equality (4.12) does not hold in general. Instead, we can use an approximation argument to obtain

$$\begin{aligned} b(p_h^1, \bar{D}\boldsymbol{u}_h^1) &= \frac{1}{\delta t}\left(j(p_h^1, p_h^1) - \left(p_h^1, \boldsymbol{\nabla}\cdot(\mathcal{I}_h^k \boldsymbol{u}_0 - \boldsymbol{u}_0)\right)\right) \\ &\ge \frac{1}{\delta t} j(p_h^1, p_h^1) - \frac{C_{\mathcal{I}}}{\delta t}\left(\nu h^{2(r_{\boldsymbol{u}}-1)}\|p_h^1\|_Q^2 + \|\boldsymbol{u}_0\|_{r_{\boldsymbol{u}},\Omega}^2\right), \end{aligned} \tag{4.14}$$

with $r_{\boldsymbol{u}} \stackrel{\text{def}}{=} \min\{k+1, r\}$. As a result, from (4.11) it follows that

$$\delta t \|\bar{D}\boldsymbol{u}_h^1\|_H^2 + \|\!|(\boldsymbol{u}_h^1, p_h^1)|\!\|_h^2 \le a(\boldsymbol{u}_h^0, \boldsymbol{u}_h^0) + C_{\mathcal{I}}\left(\nu h^{2(r_{\boldsymbol{u}}-1)}\|p_h^1\|_Q^2 + \|\boldsymbol{u}_0\|_{r_{\boldsymbol{u}},\Omega}^2\right) + \delta t \|\boldsymbol{f}(t_1)\|_H^2.$$

We conclude the proof by adding this equality to (4.10) and using the stability of the Ritz-projection. ∎

The next corollary solves the problem of the stability of the pressures by combining the results of Theorems 4.1 and 4.2.

COROLLARY 4.3. *Let $\{(\boldsymbol{u}_h^n, p_h^n)\}_{n=1}^N$ be the solution of the fully discrete problem* (3.19). *Then*

- *if $\boldsymbol{u}_0 \in [H^1(\Omega)]^d$ and $\boldsymbol{u}_h^0 = P_h^k(\boldsymbol{u}_0, 0)$, the following estimate holds for $1 \le n \le N$:*

$$\begin{aligned} \sum_{m=0}^{n-1} \delta t \|p_h^{m+1}\|_Q^2 &\le \frac{C}{\beta^2 \nu}\|\boldsymbol{u}_0\|_V^2 \\ &\quad + \frac{C}{\beta^2}\sum_{m=0}^{n-1} \delta t \left(\|\!|(\boldsymbol{u}_h^{m+1}, p_h^{m+1})|\!\|_h^2 + \nu^{-1}\|\boldsymbol{f}(t_{m+1})\|_H^2\right). \end{aligned} \tag{4.15}$$

- *if* $\boldsymbol{u}_0 \in [H^r(\Omega) \cap H_0^1(\Omega)]^d \cap H_0(\mathrm{div};\Omega)$, $r \geq 2$, $\boldsymbol{u}_h^0 = \mathcal{I}_h^k \boldsymbol{u}_0$, *and*

$$\frac{2C_{\mathcal{I}}}{\beta^2} h^{2(r_{\boldsymbol{u}}-1)} \leq \delta t, \tag{4.16}$$

the following estimate holds for $1 \leq n \leq N$*:*

$$\sum_{m=0}^{n-1} \delta t \|p_h^{m+1}\|_Q^2 \leq \frac{C}{\beta^2 \nu} \left(\|\boldsymbol{u}_0\|_V^2 + \|\boldsymbol{u}_0\|_{r_{\boldsymbol{u}},\Omega}^2 \right) + \frac{C}{\beta^2} \sum_{m=0}^{n-1} \delta t \left(\|\!|(\boldsymbol{u}_h^{m+1}, p_h^{m+1})|\!\|_h^2 + \nu^{-1} \|\boldsymbol{f}(t_{m+1})\|_H^2 \right), \tag{4.17}$$

with $r_{\boldsymbol{u}} \stackrel{\mathrm{def}}{=} \min\{k+1, r\}$.

Proof. Estimate (4.17) is a direct consequence of Theorem 4.1 and estimate (4.4). On the other hand, from Theorem 4.1 and estimate (4.5), we have

$$\left(\beta^2 \delta t - C_{\mathcal{I}} h^{2(r_{\boldsymbol{u}}-1)}\right) \|p_h^1\|_Q^2 + \beta^2 \sum_{m=1}^{n-1} \delta t \|p_h^{m+1}\|_Q^2 \leq \frac{C}{\nu} \left(\|\boldsymbol{u}_0\|_V^2 + \|\boldsymbol{u}_0\|_{r_{\boldsymbol{u}},\Omega}^2 \right) + C \sum_{m=0}^{n-1} \delta t \left(\|\!|(\boldsymbol{u}_h^{m+1}, p_h^{m+1})|\!\|_h^2 + \nu^{-1} \|\boldsymbol{f}(t_{m+1})\|_H^2 \right),$$

which combined with the stability condition (4.16) leads to (4.17). ∎

A few observations are now in order. Corollary 4.3 states the unconditional stability of the pressure provided the initial velocity approximation $\boldsymbol{u}_h^0$ is given in terms of the Ritz-projection operator (3.16). In the general case, i.e., whenever $\boldsymbol{u}_h^0$ does not satisfy a discrete divergence-free condition (as $\boldsymbol{u}_h^1$ does), only conditional stability can be guaranteed. As a matter of fact, from the stability condition (4.16), pressure instabilities are expected for very small time steps. This issue will be illustrated by numerical experiments in section 6.

Finally, let us mention that residual-based stabilization methods, such as PSPG and GLS, combined with finite difference time discretization schemes, are known to give rise to pressure instabilities in the small time-step limit; see [3, 19]. Indeed, it has been shown in [3] that the finite difference/pressure coupling of the stabilization perturbs the coercivity of the discrete pressure operator unless a condition of the type

$$Ch^2 \leq \delta t \tag{4.18}$$

is satisfied. It is worth emphasizing that, although the stability conditions (4.18) and (4.16) are somehow similar, their natures are different. Actually, the instabilities anticipated by Corollary 4.3 are related to the discrete divergence-free character of the initial velocity approximation, but not to the structure of the pressure stabilization $j(\cdot,\cdot)$.

4.2. Second order $\mathcal{A}$-stable schemes. In this subsection we discuss how the results of Theorems 4.1 and 4.2 and Corollary 4.3 extend to the second order time-stepping schemes Crank–Nicholson and BDF2.

Crank–Nicholson. The following theorem summarizes the resulting stability estimates.

THEOREM 4.4. *Let $\boldsymbol{u}_h^0$ be a given H-stable approximation of $\boldsymbol{u}_0$ in $[V_h^k]^d$, and let $\{(\boldsymbol{u}_h^n, p_h^n)\}_{n=1}^N$ be the solution of the discrete scheme* (3.20). *Then the following estimate holds for $1 \leq n \leq N$:*

$$\|\boldsymbol{u}_h^n\|_H^2 + \sum_{m=0}^{n-1} \delta t \|\!|(\boldsymbol{u}_h^{m+\frac{1}{2}}, p_h^{m+\frac{1}{2}})\|\!|_h^2 \leq C\|\boldsymbol{u}_0\|_H^2 + \frac{C_{\rm P}^2}{\nu} \sum_{m=0}^{n-1} \delta t \|\boldsymbol{f}(t_{m+\frac{1}{2}})\|_H^2.$$

Moreover, if $\boldsymbol{u}_0 \in [H^1(\Omega)]^d$ and $\boldsymbol{u}_h^0 = P_h^k(\boldsymbol{u}_0, 0)$, the following estimate holds for $1 \leq n \leq N$:

$$\sum_{m=0}^{n-1} \delta t \|p_h^{m+\frac{1}{2}}\|_Q^2 \leq \frac{C}{\beta^2 \nu}\|\boldsymbol{u}_0\|_V^2 + \frac{C}{\beta^2} \sum_{m=0}^{n-1} \delta t \left(\|\!|(\boldsymbol{u}_h^{m+\frac{1}{2}}, p_h^{m+\frac{1}{2}})\|\!|_h^2 + \nu^{-1}\|\boldsymbol{f}(t_{m+\frac{1}{2}})\|_H^2 \right).$$

On the other hand, if $\boldsymbol{u}_0 \in [H^r(\Omega) \cap H_0^1(\Omega)]^d \cap H_0(\mathrm{div};\Omega)$, $r \geq 2$, $\boldsymbol{u}_h^0 = \mathcal{I}_h^k \boldsymbol{u}_0$, and the stability condition (4.16) *is satisfied, the following estimate holds for $1 \leq n \leq N$:*

$$\sum_{m=0}^{n-1} \delta t \|p_h^{m+\frac{1}{2}}\|_Q^2 \leq \frac{C}{\beta^2 \nu} \left(\|\boldsymbol{u}_0\|_V^2 + \|\boldsymbol{u}_0\|_{r_{\boldsymbol{u}},\Omega}^2\right)$$
$$+ \frac{C}{\beta^2} \sum_{m=0}^{n-1} \delta t \left(\|\!|(\boldsymbol{u}_h^{m+\frac{1}{2}}, p_h^{m+\frac{1}{2}})\|\!|_h^2 + \nu^{-1}\|\boldsymbol{f}(t_{m+\frac{1}{2}})\|_H^2 \right).$$

Proof. The first estimate, corresponding to Theorem 4.1, holds by taking $\boldsymbol{v}_h = \boldsymbol{u}_h^{n+\frac{1}{2}}$ and $q_h = p_h^{n+\frac{1}{2}}$ in (3.20).

The pressure estimate requires an a priori bound of the discrete velocity time derivative. As in Theorem 4.2, such an estimate can be obtained by taking $\boldsymbol{v}_h = \bar{D}\boldsymbol{u}_h^{n+1}$ and $q_h = 0$ in (3.20) for $0 \leq n \leq N-1$. The main difference, with respect to the proof of Theorem 4.2, arises in the treatment of the coupling term $b(p_h^{n+\frac{1}{2}}, \bar{D}\boldsymbol{u}_h^{n+1})$. Indeed, in the Crank–Nicholson scheme incompressibility is enforced on $\boldsymbol{u}_h^{n+\frac{1}{2}}$ instead of $\boldsymbol{u}_h^{n+1}$. We first note that, since

$$\boldsymbol{u}_h^{n+1} - \boldsymbol{u}_h^n = 2\left(\boldsymbol{u}_h^{n+\frac{1}{2}} - \boldsymbol{u}_h^n\right), \quad \boldsymbol{u}_h^n = 2\boldsymbol{u}_h^{n-1+\frac{1}{2}} - \boldsymbol{u}_h^{n-1},$$

we have

$$\boldsymbol{u}_h^{n+1} - \boldsymbol{u}_h^n = 2\boldsymbol{u}_h^{n+\frac{1}{2}} + 4\sum_{l=1}^{n}(-1)^l \boldsymbol{u}_h^{n-l+\frac{1}{2}} - (-1)^n 2\boldsymbol{u}_h^0$$

for $0 \leq n \leq N-1$. Therefore, from (3.20) and using the bilinearity of $j(\cdot,\cdot)$, we get

$$\begin{aligned} b(p_h^{n+\frac{1}{2}}, \boldsymbol{u}_h^{n+1} - \boldsymbol{u}_h^n) =& j\left(2p_h^{n+\frac{1}{2}} + 4\sum_{l=1}^{n}(-1)^l p_h^{n-l+\frac{1}{2}}, p_h^{n+\frac{1}{2}}\right) \\ & - 2(-1)^n b\big(p_h^{n+\frac{1}{2}}, \boldsymbol{u}_h^0\big). \end{aligned} \tag{4.19}$$

On the other hand, for $0 \leq n \leq N-1$, we introduce the following change of variables (or extrapolation):

$$\frac{1}{2}\left(p_h^{n+1} + p_h^n\right) \stackrel{\text{def}}{=} p_h^{n+\frac{1}{2}},$$

with $p_h^0 \in Q_h^l$ to be specified later on. By inserting this expression into (4.19), we obtain

$$\begin{aligned} b(p_h^{n+\frac{1}{2}}, \boldsymbol{u}_h^{n+1} - \boldsymbol{u}_h^n) =& j\left(p_h^{n+1} - p_h^n, p_h^{n+\frac{1}{2}}\right) \\ &+ 2(-1)^n\left[j\left(p_h^0, p_h^{n+\frac{1}{2}}\right) - b\big(p_h^{n+\frac{1}{2}}, \boldsymbol{u}_h^0\big)\right]. \end{aligned} \tag{4.20}$$

If $\boldsymbol{u}_h^0 = P_h^k(\boldsymbol{u}_0, 0)$ and we choose $p_h^0 \stackrel{\text{def}}{=} R_h^l(\boldsymbol{u}_0, 0)$, from $(3.16)_2$ it follows that the last term in (4.20) cancels. Thus, we have

$$\begin{aligned} b(p_h^{n+\frac{1}{2}}, \bar{D}\boldsymbol{u}_h^{n+1}) &= j\left(\bar{D}p_h^{n+1}, p_h^{n+\frac{1}{2}}\right) \\ &= \frac{1}{2}\bar{D}j\left(p_h^{n+1}, p_h^{n+1}\right), \end{aligned}$$

which corresponds to the Crank–Nicholson counterpart of (4.8).

Finally, when the initial velocity approximation is given in terms of a general interpolant, $\boldsymbol{u}_h^0 = \mathcal{I}_h^k \boldsymbol{u}_0$ with $\boldsymbol{u}_0 \in [H^r(\Omega) \cap H_0^1(\Omega)]^d \cap H_0(\mathrm{div};\Omega)$, we take $p_h^0 \stackrel{\text{def}}{=} 0$. Therefore, from (4.20) and using an approximation argument (as in (4.14)), we get

$$\begin{aligned} b(p_h^{n+\frac{1}{2}}, \bar{D}\boldsymbol{u}_h^{n+1}) =& \frac{1}{2}\bar{D}j\left(p_h^{n+1}, p_h^{n+1}\right) - \frac{2}{\delta t}(-1)^n b\big(p_h^{n+\frac{1}{2}}, \boldsymbol{u}_h^0\big) \\ \geq& \frac{1}{2}\bar{D}j\left(p_h^{n+1}, p_h^{n+1}\right) \\ &- \frac{2C_{\mathcal{I}}}{\delta t}\left(\nu h^{2(r_{\boldsymbol{u}}-1)}\|p_h^{n+\frac{1}{2}}\|_Q^2 + \|\boldsymbol{u}_0\|_{r_{\boldsymbol{u}},\Omega}^2\right), \end{aligned} \tag{4.21}$$

which leads to the stability condition (4.16). The rest of the proof follows with minor modifications. ∎

Remark 4.5. By comparing the proofs of Corollary 4.3 and the previous theorem, we can notice that, if the initial velocity approximation is not discretely divergence free, the stability condition (4.16) has to be satisfied at each time level when using the Crank–Nicholson scheme (due to (4.21)), whereas for the backward Euler scheme that condition is needed only at the first time step (thanks to (4.8) and (4.14)).

BDF2. The following theorem summarizes the resulting stability estimates.

THEOREM 4.6. *Let $\boldsymbol{u}_h^0$ be a given H-stable approximation of $\boldsymbol{u}_0$ in $[V_h^k]^d$, let (u_h^1, p_h^1) be the corresponding first time step of the backward Euler scheme* (3.19)*, and let $\{(\boldsymbol{u}_h^n, p_h^n)\}_{n=2}^N$ be the solution of the discrete scheme* (3.21)*. Then, the following estimate holds for $2 \leq n \leq N$:*

$$\|\boldsymbol{u}_h^n\|_H^2 + 2\sum_{m=1}^{n-1}\delta t \|(\boldsymbol{u}_h^{m+1}, p_h^{m+1})\|_h^2 \leq C\left(\|\boldsymbol{u}_0\|_H^2 + \|\boldsymbol{u}_h^1\|_H^2\right) + \frac{2C_{\mathrm{P}}^2}{\nu}\sum_{m=1}^{n-1}\delta t \|\boldsymbol{f}(t_{m+1})\|_H^2.$$

Moreover, if $\boldsymbol{u}_0 \in [H^1(\Omega)]^d$ *and* $\boldsymbol{u}_h^0 = P_h^k(\boldsymbol{u}_0, 0)$*, the following estimate holds for* $2 \leq n \leq N$*:*

$$\sum_{m=1}^{n-1} \delta t \|p_h^{m+1}\|_Q^2 \leq \frac{C}{\beta^2 \nu} \left(\|\boldsymbol{u}_0\|_V^2 + |||(\boldsymbol{u}_h^1, p_h^1)|||_h^2 \right) + \frac{C}{\beta^2} \sum_{m=1}^{n-1} \delta t \left(|||(\boldsymbol{u}_h^{m+1}, p_h^{m+1})|||_h^2 + \nu^{-1} \|\boldsymbol{f}(t_{m+1})\|_H^2 \right).$$

On the other hand, if $\boldsymbol{u}_0 \in [H^r(\Omega) \cap H_0^1(\Omega)]^d \cap H_0(\mathrm{div}; \Omega)$*,* $r \geq 2$*,* $\boldsymbol{u}_h^0 = \mathcal{I}_h^k \boldsymbol{u}_0$*, and the stability condition* (4.16) *is satisfied, the following estimate holds for* $2 \leq n \leq N$*:*

$$\sum_{m=1}^{n-1} \delta t \|p_h^{m+1}\|_Q^2 \leq \frac{C}{\beta^2 \nu} \left(\|\boldsymbol{u}_0\|_V^2 + \|\boldsymbol{u}_0\|_{r_{\boldsymbol{u}},\Omega}^2 + |||(\boldsymbol{u}_h^1, p_h^1)|||_h^2 \right) + \frac{C}{\beta^2} \sum_{m=1}^{n-1} \delta t \left(|||(\boldsymbol{u}_h^{m+1}, p_h^{m+1})|||_h^2 + \nu^{-1} \|\boldsymbol{f}(t_{m+1})\|_H^2 \right).$$

Proof. The first estimate, corresponding to Theorem 4.1, holds by taking $\boldsymbol{v}_h = \boldsymbol{u}_h^{n+1}$ and $q_h = p_h^{n+1}$ in (3.21) and applying the standard identity

$$(3a - 4b + c)a = \frac{1}{2} \left[a^2 - b^2 + (2a - b)^2 - (2b - c)^2 + (a - 2b + c)^2 \right], \tag{4.22}$$

which provides the numerical dissipation of the BDF2 scheme.

Since the pressure estimate is here based on the control of the time derivative, $\tilde{D}\boldsymbol{u}_h^{n+1}$, we take $\boldsymbol{v}_h = \tilde{D}\boldsymbol{u}_h^{n+1}$ and $q_h = 0$ in (3.21). In particular, for the coupling term $b\big(p_h^{n+1}, \tilde{D}\boldsymbol{u}_h^{n+1}\big)$, using (3.21) and (4.22), we have

$$\begin{aligned} b\left(p_h^{n+1}, \tilde{D}\boldsymbol{u}_h^{n+1}\right) &= j\big(\tilde{D}p_h^{n+1}, p_h^{n+1}\big) \\ &\geq \frac{1}{4} \bar{D} \left(j(p_h^{n+1}, p_h^{n+1}) + j(2p_h^{n+1} - p_h^n, 2p_h^{n+1} - p_h^n) \right) \end{aligned} \tag{4.23}$$

for $2 \leq n \leq N - 1$, which corresponds to the BDF2 counterpart of (4.8). On the other hand, for $n = 1$, from (3.21) and (3.19), we obtain

$$b\left(p_h^2, \tilde{D}\boldsymbol{u}_h^2\right) = \frac{1}{2\delta t} \left(3j(p_h^2, p_h^2) - 4j(p_h^1, p_h^2) + b(p_h^2, \boldsymbol{u}_h^0) \right). \tag{4.24}$$

If the initial velocity approximation is given in terms of the Ritz-projection, $\boldsymbol{u}_h^0 = P_h^k(\boldsymbol{u}_0, 0)$, it follows that $b(p_h^2, \boldsymbol{u}_h^0) = j(p_h^0, p_h^2)$, with $p_h^0 \stackrel{\text{def}}{=} R_h^l(\boldsymbol{u}_0, 0)$. Thus, (4.24) reduces to

$$b\left(p_h^2, \tilde{D}\boldsymbol{u}_h^2\right) = j(\tilde{D}p_h^2, p_h^2),$$

so that (4.23) holds true also for $n = 1$.

Finally, if the initial velocity approximation is given in terms of a general interpolant, $\boldsymbol{u}_h^0 = \mathcal{I}_h^k \boldsymbol{u}_0$, we apply an approximation argument (as in (4.14)). Hence, from (4.24)

(4.25)
$$b\left(p_h^2, \tilde{D}\boldsymbol{u}_h^2\right) \geq \frac{1}{2\delta t}\left[3j(p_h^2, p_h^2) - 4j(p_h^1, p_h^1) - C_{\mathcal{I}}\left(\nu h^{2(r_{\boldsymbol{u}}-1)}\|p_h^2\|_Q^2 + \|\boldsymbol{u}_0\|_{r_{\boldsymbol{u}},\Omega}^2\right)\right],$$

which is the BDF2 counterpart of (4.14) and leads to the stability condition (4.16). The rest of the proof follows with minor modifications. ∎

Remark 4.7. A bound for the backward Euler initialization terms $\|\boldsymbol{u}_h^1\|_H$ and $|||(\boldsymbol{u}_h^1, p_h^1)|||_h$, appearing in the above estimates, is provided by Theorems 4.1 and 4.2 with $n = 1$.

Remark 4.8. When the initial velocity approximation is not discretely divergence free, the stability condition (4.16) has to be satisfied twice when using BDF2, at the first time step (according to (4.25)) and at the backward Euler initialization (see Theorem 4.2).

5. Convergence. In this section we provide optimal convergence error estimates for the discrete formulation (3.19), the backward Euler scheme.

Theorem 5.2 concerns the convergence for the velocity and gives an estimate for the pressure in terms of the error in the velocity time derivative. Theorem 5.3 answers the question of optimal convergence of the pressure by providing an optimal error estimate for the time derivative, provided the exact pressure is smooth. Finally, Theorem 5.4 provides an improved $L^\infty((0,T),H)$ estimate that justifies the initialization of the BDF2 scheme with a backward Euler step.

The following result expresses the modified Galerkin orthogonality in terms of the consistency error in space and time.

LEMMA 5.1 (consistency error). *Let* $(\boldsymbol{u}, p)$ *be the solution of* (2.1) *and let* $\{(\boldsymbol{u}_h^n, p_h^n)\}_{0\leq n\leq N}$ *be the solution of* (3.19). *Assume that* $\boldsymbol{u} \in \mathcal{C}^0\big([0,T];V\big) \cap \mathcal{C}^1\big((0,T];H\big)$, *and let* $p \in \mathcal{C}^0\big((0,T];Q\big)$. *Then, for* $0 \leq n \leq N-1$, *there holds*

$$\begin{aligned}\big(\bar{D}\boldsymbol{u}(t_{n+1}) - \bar{D}\boldsymbol{u}_h^{n+1}, \boldsymbol{v}_h\big) + a\big(\boldsymbol{u}(t_{n+1}) - \boldsymbol{u}_h^{n+1}, \boldsymbol{v}_h\big) + b(p(t_{n+1}) - p_h^{n+1}, \boldsymbol{v}_h)\\ - b(q_h, \boldsymbol{u}(t_{n+1}) - \boldsymbol{u}_h^{n+1}) = j\big(p_h^{n+1}, q_h\big) + \big(\bar{D}\boldsymbol{u}(t_{n+1}) - \partial_t\boldsymbol{u}(t_{n+1}), \boldsymbol{v}_h\big)\end{aligned}$$

for all $(\boldsymbol{v}_h, q_h) \in [V_h^k]^d \times Q_h^l$.

THEOREM 5.2. *Assume that* $\boldsymbol{u} \in H^1(0,T;[H^r(\Omega)]^d) \cap H^2(0,T;[L^2(\Omega)]^d)$ *and* $p \in C^0((0,T];H^s(\Omega))$ *with* $r \geq 2$ *and* $s \geq 1$, *and set* $\boldsymbol{u}_h^0 \in [V_h^k]^d$ *as a given approximation of* $\boldsymbol{u}_0$. *Then the following estimate holds for* $1 \leq n \leq N$:

$$\begin{aligned}\|\boldsymbol{u}_h^n - \boldsymbol{u}(t_n)\|_H^2 + \sum_{m=0}^{n-1}\delta t|||(\boldsymbol{u}_h^{m+1} - \boldsymbol{u}(t_{m+1}), p_h^{m+1})|||_h^2 \leq \|\mathcal{I}_h^k\boldsymbol{u}_0 - \boldsymbol{u}_h^0\|_H^2\\ + Ch^{2r_{\boldsymbol{u}}}\left(\|\boldsymbol{u}\|_{C^0([t_1,t_n];H^{r_{\boldsymbol{u}}}(\Omega))}^2 + \nu^{-1}\|\partial_t\boldsymbol{u}\|_{L^2(0,t_n;H^{r_{\boldsymbol{u}}}(\Omega))}^2\right)\\ + C\Big(\frac{\delta t^2}{\nu}\|\partial_{tt}\boldsymbol{u}\|_{L^2(0,t_n;H)}^2 + \frac{h^{2s_p}}{\nu}t_n\|p\|_{C^0([t_1,t_n];H^{s_p}(\Omega))}^2\\ + \nu h^{2(r_{\boldsymbol{u}}-1)}t_n\|\boldsymbol{u}\|_{C^0([t_1,t_n];H^{r_{\boldsymbol{u}}}(\Omega))}^2\Big),\end{aligned}$$

$$\sum_{m=0}^{n-1} \delta t \|p_h^{m+1} - p(t_{m+1})\|_Q^2 \le C \left(1 + \frac{1}{\beta^2}\right) \frac{h^{2s_p}}{\nu} t_n \|p\|^2_{C^0([t_1,t_n];H^{s_p}(\Omega))}$$
$$+ \frac{C}{\beta^2} \sum_{m=0}^{n-1} \delta t \Big(\|\!|(\boldsymbol{u}_h^{m+1} - \boldsymbol{u}(t_{m+1}), p_h^{m+1})|\!\|_h^2 + \nu^{-1} \left\|\partial_t \boldsymbol{u}(t_{m+1}) - \bar{D}\boldsymbol{u}_h^{m+1}\right\|_H^2 \Big),$$

with $C > 0$ a positive constant independent of h, δt, and ν.

Proof. The error estimate for the velocity follows standard energy arguments, and for the pressure we use the modified inf-sup condition (3.8). We start by decomposing the velocity and pressure error using, respectively, the projections $\mathcal{I}_h^k$ and Π_h^l. This yields

(5.1)
$$\boldsymbol{u}(t_{n+1}) - \boldsymbol{u}_h^{n+1} = \underbrace{\boldsymbol{u}(t_{n+1}) - \mathcal{I}_h^k \boldsymbol{u}(t_{n+1})}_{\boldsymbol{\theta}_\pi^{n+1}} + \underbrace{\mathcal{I}_h^k \boldsymbol{u}(t_{n+1}) - \boldsymbol{u}_h^{n+1}}_{\boldsymbol{\theta}_h^{n+1}} = \boldsymbol{\theta}_\pi^{n+1} + \boldsymbol{\theta}_h^{n+1},$$
$$p(t_{n+1}) - p_h^{n+1} = \underbrace{p(t_{n+1}) - \Pi_h^l p(t_{n+1})}_{y_\pi^{n+1}} + \underbrace{\Pi_h^l p(t_{n+1}) - p_h^{n+1}}_{y_h^{n+1}} = y_\pi^{n+1} + y_h^{n+1}.$$

The first term $\boldsymbol{\theta}_\pi^{n+1}$ can bounded using approximation (3.5). In order to estimate $\boldsymbol{\theta}_h^{n+1}$ we first note, using (4.3) and the coercivity of the bilinear form $a(\cdot,\cdot) + j(\cdot,\cdot)$,

(5.2)
$$\frac{1}{2}\bar{D}\|\boldsymbol{\theta}_h^{n+1}\|_H^2 + \|\!|(\boldsymbol{\theta}_h^{n+1}, y_h^{n+1})|\!\|_h^2 \le (\bar{D}\boldsymbol{\theta}_h^{n+1}, \boldsymbol{\theta}_h^{n+1}) + \|\!|(\boldsymbol{\theta}_h^{n+1}, y_h^{n+1})|\!\|_h^2$$
$$\le \underbrace{(\bar{D}\boldsymbol{\theta}_h^{n+1}, \boldsymbol{\theta}_h^{n+1}) + a(\boldsymbol{\theta}_h^{n+1}, \boldsymbol{\theta}_h^{n+1}) + b(y_h^{n+1}, \boldsymbol{\theta}_h^{n+1}) - b(y_h^{n+1}, \boldsymbol{\theta}_h^{n+1}) + j(y_h^{n+1}, y_h^{n+1})}_{T_1^{n+1}}.$$

In addition, using (5.1) we have

$$T_1^{n+1} = -(\bar{D}\boldsymbol{\theta}_\pi^{n+1}, \boldsymbol{\theta}_h^{n+1}) - a(\boldsymbol{\theta}_\pi^{n+1}, \boldsymbol{\theta}_h^{n+1}) + j(\Pi_h^l p(t_{n+1}), y_h^{n+1}) - b(y_\pi^{n+1}, \boldsymbol{\theta}_h^{n+1})$$
$$+ b(y_h^{n+1}, \boldsymbol{\theta}_\pi^{n+1}) + (\bar{D}\boldsymbol{u}(t_{n+1}) - \bar{D}\boldsymbol{u}_h^{n+1}, \boldsymbol{\theta}_h^{n+1}) + a(\boldsymbol{u}(t_{n+1}) - \boldsymbol{u}_h^{n+1}, \boldsymbol{\theta}_h^{n+1})$$
$$+ b(p(t_{n+1}) - p_h^{n+1}, \boldsymbol{\theta}_h^{n+1}) - b(y_h^{n+1}, \boldsymbol{u}(t_{n+1}) - \boldsymbol{u}_h^{n+1}) - j(p_h^{n+1}, y_h^{n+1}).$$

By the modified Galerkin orthogonality (Lemma 5.1), this expression reduces to

(5.3)
$$T_1^{n+1} = -(\bar{D}\boldsymbol{\theta}_\pi^{n+1}, \boldsymbol{\theta}_h^{n+1}) + \big(\bar{D}\boldsymbol{u}(t_{n+1}) - \partial_t \boldsymbol{u}(t_{n+1}), \boldsymbol{\theta}_h^{n+1}\big)$$
$$- a(\boldsymbol{\theta}_\pi^{n+1}, \boldsymbol{\theta}_h^{n+1}) + j(\Pi_h^l p(t_{n+1}), y_h^{n+1}) - b(y_\pi^{n+1}, \boldsymbol{\theta}^{n+1}) + b(y_h^{n+1}, \boldsymbol{\theta}_\pi^{n+1}).$$

Now, using the Cauchy–Schwarz and the Poincaré inequalities and (3.6), we have

(5.4)
$$T_1^{n+1} \le \underbrace{\big(\|\bar{D}\boldsymbol{u}(t_{n+1}) - \partial_t \boldsymbol{u}(t_{n+1})\|_H + \|\bar{D}\boldsymbol{\theta}_\pi^{n+1}\|_H\big)}_{T_2^{n+1}} \frac{C_{\mathrm{P}}}{\nu^{\frac{1}{2}}} \|\!|(\boldsymbol{\theta}_h^{n+1}, y_h^{n+1})|\!\|$$
$$+ \Big(\|\boldsymbol{\theta}_\pi^{n+1}\|_V + \|y_\pi^{n+1}\|_Q + \nu^{\frac{1}{2}} \|h^{-1}\boldsymbol{\theta}_\pi^{n+1}\|_H$$
$$+ j\big(\Pi_h^l p(t_{n+1}), \Pi_h^l p(t_{n+1})\big)^{\frac{1}{2}} \Big) \|\!|(\boldsymbol{\theta}_h^{n+1}, y_h^{n+1})|\!\|.$$

The term T_2^{n+1} can be treated, in a standard way (see, e.g., [30]), using a Taylor expansion and the Cauchy–Schwarz inequality, which yields

$$
\begin{aligned}
T_2^{n+1} &\leq \frac{1}{\delta t}\int_{t_n}^{t_{n+1}} (\delta t\|\partial_{tt}\boldsymbol{u}(s)\|_H + \|\partial_t\boldsymbol{\theta}_\pi(s)\|_H)\,\mathrm{d}s \\
&\leq \delta t^{\frac{1}{2}}\|\partial_{tt}\boldsymbol{u}(s)\|_{L^2((t_n,t_{n+1});H)} + \delta t^{-\frac{1}{2}}\|\partial_t\boldsymbol{\theta}_\pi\|_{L^2((t_n,t_{n+1});H)}.
\end{aligned} \tag{5.5}
$$

Thus, from (5.4), using Young's inequality, it follows that

$$
\begin{aligned}
T_1^{n+1} \leq{}& \frac{1}{2}|\!|\!|(\boldsymbol{\theta}_h^{n+1}, y_h^{n+1})|\!|\!|^2 \\
&+ C\Bigg[\frac{C_{\mathrm{P}}^2}{\nu}\left(\delta t\|\partial_{tt}\boldsymbol{u}\|^2_{L^2((t_n,t_{n+1});H)} + \delta t^{-1}\|\partial_t\boldsymbol{\theta}_\pi\|^2_{L^2((t_n,t_{n+1});H)}\right) \\
&+ \|\boldsymbol{\theta}_\pi^{n+1}\|_V^2 + \|y_\pi^{n+1}\|_Q^2 + \nu\|h^{-1}\boldsymbol{\theta}_\pi^{n+1}\|_H^2 + j\big(\Pi_h^l p(t_{n+1}), \Pi_h^l p(t_{n+1})\big)\Bigg].
\end{aligned}
$$

By inserting this expression into (5.2), multiplying the resulting expression by $2\delta t$, and summing over $0 \leq m \leq n-1$, we obtain

$$
\begin{aligned}
&\|\boldsymbol{\theta}_h^n\|_H^2 + \sum_{m=0}^{n-1}\delta t|\!|\!|(\boldsymbol{\theta}_h^{m+1}, y_h^{m+1})|\!|\!|_h^2 \\
&\leq \|\boldsymbol{\theta}_h^0\|_H^2 + C\Bigg[\delta t^2\nu^{-1}\|\partial_{tt}\boldsymbol{u}\|^2_{L^2(0,t_n;H)} + \nu^{-1}\|\partial_t\boldsymbol{\theta}_\pi\|^2_{L^2(0,t_n;H)} \\
&\quad + \sum_{m=0}^{n-1}\delta t\bigg(\|\boldsymbol{\theta}_\pi^{m+1}\|_V^2 + \|y_\pi^{m+1}\|_Q^2 + \nu\|h^{-1}\boldsymbol{\theta}_\pi^{m+1}\|_H^2 + j\big(\Pi_h^l p(t_{m+1}), \Pi_h^l p(t_{m+1})\big)\bigg)\Bigg].
\end{aligned}
$$

Finally, the velocity error estimate is obtained using approximation (3.5) and the consistency of the pressure stabilization (3.3), which yields

$$
\begin{aligned}
\|\boldsymbol{\theta}_h^n\|_H^2 &+ \sum_{m=0}^{n-1}\delta t|\!|\!|(\boldsymbol{\theta}_h^{m+1}, y_h^{m+1})|\!|\!|_h^2 \leq \|\boldsymbol{\theta}_h^0\|_H^2 \\
&+ C\Bigg[\frac{\delta t^2}{\nu}\|\partial_{tt}\boldsymbol{u}\|^2_{L^2(0,t_n;H)} + \frac{h^{2r_{\boldsymbol{u}}}}{\nu}\|\partial_t\boldsymbol{u}\|^2_{L^2(0,t_n;H^{r_{\boldsymbol{u}}}(\Omega))} \\
&+ \nu h^{2(r_{\boldsymbol{u}}-1)}\sum_{m=0}^{n-1}\delta t\|\boldsymbol{u}(t_{m+1})\|^2_{r_{\boldsymbol{u}},\Omega} + \frac{h^{2s_p}}{\nu}\sum_{m=0}^{n-1}\delta t\|p(t_{m+1})\|^2_{s_p,\Omega}\Bigg].
\end{aligned}
$$

For the pressure error estimate we first note that, from (5.1), it suffices to control $\|y_h^{n+1}\|_{0,\Omega}$. To this end, we use the modified inf-sup condition (3.8):

$$
\beta\|y_h^{n+1}\|_Q \leq \sup_{\boldsymbol{v}_h\in[V_h^k]^d}\frac{|b(y_h^{n+1}, \boldsymbol{v}_h)|}{\|\boldsymbol{v}_h\|_V} + Cj(y_h^{n+1}, y_h^{n+1})^{\frac{1}{2}}. \tag{5.6}
$$

From (5.1) we get

$$
b(y_h^{n+1}, \boldsymbol{v}_h) = -b(y_\pi^{n+1}, \boldsymbol{v}_h) + b(p(t_{n+1}) - p_h^{n+1}, \boldsymbol{v}_h).
$$

The first term can be bounded, using the continuity of $b(\cdot,\cdot)$ (see (2.3)), which yields

$$b(y_\pi^{n+1},\boldsymbol{v}_h) \le \|y_\pi^{n+1}\|_Q \|\boldsymbol{v}_h\|_V.$$

On the other hand, using the modified Galerkin orthogonality (Lemma 5.1 with $q_h = 0$) we have

$$\begin{aligned} b(p(t_{n+1}) - p_h^{n+1},\boldsymbol{v}_h) &\\ &= -a\big(\boldsymbol{u}(t_{n+1}) - \boldsymbol{u}_h^{n+1},\boldsymbol{v}_h\big) - \big(\partial_t \boldsymbol{u}(t_{n+1}) - \bar{D}\boldsymbol{u}_h^{n+1},\boldsymbol{v}_h\big)\\ &\le C\|\!|(\boldsymbol{u}(t_{n+1}) - \boldsymbol{u}_h^{n+1},0)|\!\|_h \|\boldsymbol{v}_h\|_V + \|\partial_t \boldsymbol{u}(t_{n+1}) - \bar{D}\boldsymbol{u}_h^{n+1}\|_H \|\boldsymbol{v}_h\|_H. \end{aligned}$$

As a result, from the above estimations we have

$$\beta\|y_h^{n+1}\|_Q \le C\left(\|y_\pi^{n+1}\|_Q + \|\!|(\boldsymbol{u}(t_{n+1}) - \boldsymbol{u}_h^{n+1}, y_h^{n+1})|\!\|_h\right) + \frac{C_{\rm P}}{\nu^{\frac{1}{2}}}\|\partial_t \boldsymbol{u}(t_{n+1}) - \bar{D}\boldsymbol{u}_h^{n+1}\|_H.$$

Therefore,

$$\begin{aligned} \beta^2 \sum_{m=0}^{n-1}\delta t\|y_h^{m+1}\|_Q^2 \le& C\sum_{m=0}^{n-1}\delta t\Big(\|y_\pi^{m+1}\|_Q^2 + \|\!|(\boldsymbol{u}(t_{m+1}) - \boldsymbol{u}_h^{m+1}, y_h^{m+1})|\!\|_h^2 \\ &+ \nu^{-1}\|\partial_t \boldsymbol{u}(t_{m+1}) - \bar{D}\boldsymbol{u}_h^{m+1}\|_H^2\Big), \end{aligned}$$

and we conclude using approximation and the error estimate for the velocity. ∎

We solve the problem of the pressure convergence by providing an error estimate for the time derivative of the velocity.

THEOREM 5.3. *Under the assumptions of Theorem* 5.2, *assuming that* $p \in C^0([0,T];H^s(\Omega))$, $\boldsymbol{u}_0 \in V \cap H_0(\mathrm{div};\Omega)$, *and* $\boldsymbol{u}_h^0 \stackrel{\mathrm{def}}{=} P_h^k(\boldsymbol{u}_0,0)$, *for* $1 \le n \le N$ *we have*

$$\begin{aligned} &\sum_{m=0}^{n-1}\delta t\left\|\bar{D}\boldsymbol{u}_h^{m+1} - \partial_t \boldsymbol{u}(t_{m+1})\right\|_H^2 + \|\!|\big(P_h^k\big(\boldsymbol{u}(t_n),p(t_n)\big) - \boldsymbol{u}_h^n, R_h^l\big(\boldsymbol{u}(t_n),p(t_n)\big) - p_h^n\big)|\!\|_h^2 \\ &\quad \le C\Big(\delta t^2\|\partial_{tt}\boldsymbol{u}\|_{L^2(0,T;H)}^2 + h^{2r_{\boldsymbol{u}}}\|\partial_t \boldsymbol{u}\|_{L^2(0,T;H^{r_{\boldsymbol{u}}}(\Omega))}^2\Big) + C\frac{h^{2s_p}}{\nu}\|p(0)\|_{s_p,\Omega}^2. \end{aligned}$$

Proof. In order to provide an optimal error estimate, we decompose the error in terms of the Ritz-projection operator (3.16) as follows:

(5.7)

$$\begin{aligned} \boldsymbol{u}(t_{n+1}) - \boldsymbol{u}_h^{n+1} &= \underbrace{\boldsymbol{u}(t_{n+1}) - P_h^k(\boldsymbol{u}(t_{n+1}),p(t_{n+1}))}_{\boldsymbol{\theta}_\pi^{n+1}} + \underbrace{P_h^k(\boldsymbol{u}(t_{n+1}),p(t_{n+1})) - \boldsymbol{u}_h^{n+1}}_{\boldsymbol{\theta}_h^{n+1}}\\ &= \boldsymbol{\theta}_\pi^{n+1} + \boldsymbol{\theta}_h^{n+1},\\ p(t_{n+1}) - p_h^{n+1} &= \underbrace{p(t_{n+1}) - R_h^l(\boldsymbol{u}(t_{n+1}),p(t_{n+1}))}_{y_\pi^{n+1}} + \underbrace{R_h^l(\boldsymbol{u}(t_{n+1}),p(t_{n+1})) - p_h^{n+1}}_{y_h^{n+1}}\\ &= y_\pi^{n+1} + y_h^{n+1}. \end{aligned}$$

Using the triangle inequality, we then have

$$\begin{aligned}(5.8)\quad &\sum_{m=0}^{n-1} \delta t \|\partial_t \boldsymbol{u}(t_{m+1}) - \bar{D}\boldsymbol{u}_h^{m+1}\|_H^2 \\ &\leq C \sum_{m=0}^{n-1} \delta t \left(\|\partial_t \boldsymbol{u}(t_{m+1}) - \bar{D}\boldsymbol{u}(t_{m+1})\|_H^2 + \|\bar{D}\boldsymbol{\theta}_\pi^{m+1}\|_H^2 + \|\bar{D}\boldsymbol{\theta}_h^{m+1}\|_H^2 \right).\end{aligned}$$

For the first term, we proceed as in (5.5) using a Taylor expansion, which yields

$$\|\partial_t \boldsymbol{u}(t_{n+1}) - \bar{D}\boldsymbol{u}(t_{n+1})\|_H \leq \delta t^{\frac{1}{2}} \|\partial_{tt} \boldsymbol{u}(s)\|_{L^2((t_n,t_{n+1});H)}.$$

For the second term, we have

$$(5.9)\qquad \|\bar{D}\boldsymbol{\theta}_\pi^{n+1}\|_H = \frac{1}{\delta t} \int_{t_n}^{t_{n+1}} \|\partial_t \boldsymbol{\theta}_\pi(s)\|_H ds \leq \delta t^{-\frac{1}{2}} \|\partial_t \boldsymbol{\theta}_\pi\|_{L^2((t_n,t_{n+1});H)}.$$

Finally, for the third term we use the modified Galerkin orthogonality (Lemma 5.1 with $q_h = 0$) and the definition of the Ritz-projection (3.16) to obtain

$$\begin{aligned}&\|\bar{D}\boldsymbol{\theta}_h^{n+1}\|_H^2 + a(\boldsymbol{\theta}_h^{n+1}, \bar{D}\boldsymbol{\theta}_h^{n+1}) + b(y_h^{n+1}, \bar{D}\boldsymbol{\theta}_h^{n+1}) \\ &\quad = -(\bar{D}\boldsymbol{\theta}_\pi^{n+1}, \bar{D}\boldsymbol{\theta}_h^{n+1}) - a(\boldsymbol{\theta}_\pi^{n+1}, \bar{D}\boldsymbol{\theta}_h^{n+1}) \\ &\qquad - b(y_\pi^{n+1}, \bar{D}\boldsymbol{\theta}_h^{n+1}) + (\bar{D}\boldsymbol{u}(t_{n+1}) - \partial_t \boldsymbol{u}(t_{n+1}), \bar{D}\boldsymbol{\theta}_h^{n+1}) \\ &\quad = -(\bar{D}\boldsymbol{\theta}_\pi^{n+1}, \bar{D}\boldsymbol{\theta}_h^{n+1}) + (\bar{D}\boldsymbol{u}(t_{n+1}) - \partial_t \boldsymbol{u}(t_{n+1}), \bar{D}\boldsymbol{\theta}_h^{n+1}).\end{aligned}$$

Young's inequality yields

$$\begin{aligned}&\frac{1}{2}\|\bar{D}\boldsymbol{\theta}_h^{n+1}\|_H^2 + a(\boldsymbol{\theta}_h^{n+1}, \bar{D}\boldsymbol{\theta}_h^{n+1}) + b(y_h^{n+1}, \bar{D}\boldsymbol{\theta}_h^{n+1}) \\ &\qquad \leq C \left(\|\bar{D}\boldsymbol{\theta}_\pi^{n+1}\|_H^2 + \|\bar{D}\boldsymbol{u}(t_{n+1}) - \partial_t \boldsymbol{u}(t_{n+1})\|_H^2 \right).\end{aligned}$$

In addition, for $0 \leq n \leq N$, testing (3.16) at the time level n with $\boldsymbol{v}_h = \boldsymbol{0}$, we have

$$(5.10)\qquad b\big(q_h, P_h^k(\boldsymbol{u}(t_n), p(t_n))\big) = j\big(R_h^l(\boldsymbol{u}(t_n), p(t_n)), q_h\big).$$

On the other hand, for $1 \leq n \leq N$, testing (3.19) at the time level n with $\boldsymbol{v}_h = \boldsymbol{0}$ and since, by definition, $\boldsymbol{u}_h^0 \overset{\text{def}}{=} P_h^k(\boldsymbol{u}_0, 0)$, we have

$$(5.11)\qquad b(q_h, \boldsymbol{u}_h^n) = j(p_h^n, q_h)$$

for all $q_h \in Q_h^l$ and $0 \leq n \leq N$ and where we have defined $p_h^0 \overset{\text{def}}{=} R_h^l(\boldsymbol{u}_0, 0)$. As a result, from (5.10)–(5.11), we have

$$b(q_h, \boldsymbol{\theta}_h^n) = j(y_h^n, q_h)$$

for all $q_h \in Q_h^l$ and $0 \leq n \leq N$. We therefore have, for $0 \leq n \leq N-1$,

$$b(y_h^{n+1}, \bar{D}\boldsymbol{\theta}_h^{n+1}) = j(\bar{D}y_h^{n+1}, y_h^{n+1}).$$

On the other hand, using the symmetry of a and j, we have

$$a(\boldsymbol{\theta}_h^{n+1}, \bar{D}\boldsymbol{\theta}_h^{n+1}) = \frac{1}{2}\bar{D}a(\boldsymbol{\theta}_h^{n+1}, \boldsymbol{\theta}_h^{n+1}) + \frac{\delta t}{2}a(\bar{D}\boldsymbol{\theta}_h^{n+1}, \bar{D}\boldsymbol{\theta}_h^{n+1}),$$

$$j(y_h^{n+1}, \bar{D}y_h^{n+1}) = \frac{1}{2}\bar{D}j(y_h^{n+1}, y_h^{n+1}) + \frac{\delta t}{2}j(\bar{D}y_h^{n+1}, \bar{D}y_h^{n+1}),$$

so that

$$\frac{1}{2}\|\bar{D}\boldsymbol{\theta}_h^{n+1}\|_H^2 + \frac{1}{2}\bar{D}\left(a(\boldsymbol{\theta}_h^{n+1}, \boldsymbol{\theta}_h^{n+1}) + j(y_h^{n+1}, y_h^{n+1})\right) \leq \|\bar{D}\boldsymbol{\theta}_\pi^{n+1}\|_H^2 + \|\bar{D}\boldsymbol{u}(t_{n+1}) - \partial_t \boldsymbol{u}(t_{n+1})\|_H^2.$$

Thus, after multiplication by $2\delta t$ and summation over $0 \leq n \leq N-1$, we have

$$\sum_{m=0}^{n-1} \delta t\|\bar{D}\boldsymbol{\theta}_h^{m+1}\|_H^2 + \|\!|(\boldsymbol{\theta}_h^n, y_h^n)|\!\|_h^2 \leq \|\!|(\boldsymbol{\theta}_h^0, y_h^0)|\!\|_h^2 + C\sum_{m=0}^{n-1}\delta t\left(\|\bar{D}\boldsymbol{\theta}_\pi^{m+1}\|_H^2 + \|\bar{D}\boldsymbol{u}(t_{m+1}) - \partial_t\boldsymbol{u}(t_{m+1})\|_H^2\right). \tag{5.12}$$

For the initial terms, we use the linearity of the Ritz-projection and its approximation properties (Lemma 3.4) to obtain

$$\|\!|(\boldsymbol{\theta}_h^0, y_h^0)|\!\|_h^2 = \|\!|(P_h^k(\mathbf{0}, p(0)), R_h^l(\mathbf{0}, p(0)))|\!\|_h^2 \leq \frac{C}{\nu}h^{2s_p}\|p(0)\|_{s_p,\Omega}^2.$$

Therefore, using (5.9) and (5.5), we have

$$\sum_{m=0}^{n-1}\delta t\|\bar{D}\boldsymbol{\theta}_h^{m+1}\|_H^2 + \|\!|(\boldsymbol{\theta}_h^n, y_h^n)|\!\|_h^2 \leq C\left(\frac{h^{2s_p}}{\nu}\|p(0)\|_{s_p,\Omega}^2 + \delta t^2\|\partial_{tt}\boldsymbol{u}\|_{L^2(0,t_n;H)}^2\right)$$

for $1 \leq n \leq N$. ◻

Finally, for completeness, we here give a result of optimal convergence in the $L^\infty((0,T),H)$-norm. For this we assume that the domain Ω is such that the optimal convergence in the H-norm holds for the Ritz-projection (see Lemma 3.4). This result is of importance since it shows that the initialization of the BDF2 method using one BDF1 step is justified (i.e., we keep error optimality in time).

THEOREM 5.4. *Assume that the domain Ω is sufficiently smooth so that the H-estimate* (3.18) *holds. Assume also that $\boldsymbol{u} \in H^1(0,T;[H^r_{\boldsymbol{u}}(\Omega)]^d) \cap H^2(0,T;[L^2(\Omega)]^d)$, $p \in C^0([0,T];H^{s_p}(\Omega))$ with $r_{\boldsymbol{u}} \geq 2$, $s_p \geq 1$, $\boldsymbol{u}_0 \in V \cap H_0(\mathrm{div};\Omega)$, and $\boldsymbol{u}_h^0 \stackrel{\mathrm{def}}{=} P_h^k(\boldsymbol{u}_0, 0)$. Then the following estimate holds for $1 \leq n \leq N$:*

$$\|\boldsymbol{u}(t_n) - \boldsymbol{u}_h^n\|_H \leq \frac{C}{\nu^{\frac{1}{2}}}\Big(h^{r_{\boldsymbol{u}}}\|\boldsymbol{u}_0\|_{r_{\boldsymbol{u}},\Omega} + h^{s_p+1}\|p(0)\|_{s_p,\Omega} + h^{r_{\boldsymbol{u}}}\|\partial_t\boldsymbol{u}\|_{L^1(0,t_n;H^{r_{\boldsymbol{u}}}(\Omega))} + \delta t\|\partial_{tt}\boldsymbol{u}\|_{L^1(0,t_n;H)}\Big).$$

Proof. Since the proof is similar to that of Theorem 5.3 and we will give only the outline. Let $\boldsymbol{\theta}_h^{n+1}$ and y_h^{n+1} be defined as in (5.7). From (5.2) and (5.3), it follows that

$$(\bar{D}\boldsymbol{\theta}_h^{n+1}, \boldsymbol{\theta}_h^{n+1}) \leq -(\bar{D}\boldsymbol{\theta}_\pi^{n+1}, \boldsymbol{\theta}_h^{n+1}) + (\bar{D}\boldsymbol{u}(t_{n+1}) - \partial_t \boldsymbol{u}(t_{n+1}), \boldsymbol{\theta}_h^{n+1}).$$

Applying now the Cauchy–Schwarz inequality, we have

$$\|\boldsymbol{\theta}_h^{n+1}\|_H \leq \|\boldsymbol{\theta}_h^{n}\|_H + \delta t \left(\|\bar{D}\boldsymbol{\theta}_\pi^{n+1}\|_H + \|\bar{D}\boldsymbol{u}(t_{n+1}) - \partial_t \boldsymbol{u}(t_{n+1})\|_H\right),$$

and by summation over n, we get

$$\|\boldsymbol{\theta}_h^{n}\|_H \leq \|\boldsymbol{\theta}_h^{0}\|_H + \sum_{m=0}^{n-1} \delta t \left(\|\bar{D}\boldsymbol{\theta}_\pi^{n+1}\|_H + \|\bar{D}\boldsymbol{u}(t_{n+1}) - \partial_t \boldsymbol{u}(t_{n+1})\|_H\right)$$

for $1 \leq n \leq N$. The first term in the right-hand side can be estimated using Lemma 3.4 since, by definition,

$$\boldsymbol{\theta}_h^0 = P_h^k(\boldsymbol{u}(0), p(0)) - \boldsymbol{u}_h^0 = P_h^k(\boldsymbol{u}_0, p(0)) - P_h^k(\boldsymbol{u}_0, 0) = P_h^k(\boldsymbol{0}, p(0)). \tag{5.13}$$

Finally, for the finite difference consistency terms we use a standard argument (see, e.g., [34, Theorem 1.5, page 14]). ∎

Remark 5.5. From (5.13), one could pretend to initialize the time-stepping procedure with $\boldsymbol{u}_h^0 = P_h^k(\boldsymbol{u}_0, p(0))$ (as in [33], for instance). In practice, however, the initial pressure is unknown, so that the choice $\boldsymbol{u}_h^0 = P_h^k(\boldsymbol{u}_0, 0)$ is more convenient. Lemma 3.4 shows that we can preserve optimality while keeping this choice (see also [4]).

Remark 5.6. Note that the above convergence proofs use only stability, Galerkin orthogonality, and the truncation error of the finite difference time approximation scheme. Hence the extension to the second order Crank–Nicholson or BDF2 scheme is straightforward. In particular we recall that the estimate of Theorem 5.4 shows that the initialization using one BDF1 step does not make the convergence deteriorate, provided the solution is sufficiently smooth under the first time step. Indeed, for smooth solutions we expect $\|\partial_{tt}\boldsymbol{u}\|_{L^1(0,\delta t;H)}$ to be $O(\delta t)$, and hence the global convergence will be second order in spite of the initial low order perturbation.

6. Numerical experiments.

In this section we will consider some numerical examples using the CIP stabilization, described in subsection 3.1.1. We present computations demonstrating the optimal convergence using finite element spaces consisting of quadratic functions, for the space discretization, BDF1, BDF2, and the Crank–Nicholson scheme for the time discretization. We also verify numerically that, for small time steps, the pressure is unstable for initial data that are not discretely divergence free. All computations have been performed using FreeFem++ [26].

6.1. Convergence rate in time.

We consider problem (2.1) in two dimensions, $\Omega = [0,1] \times [0,1]$ and $T = 1$, with nonhomogeneous boundary conditions. The right-hand side $\boldsymbol{f}$ and the boundary and initial data are chosen in order to ensure that the exact solution is given by

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

$$p(x,y,t) = g(t)\big(\sin(x)\cos(y) + (\cos(1) - 1)\sin(1)\big),$$

with $g(t) = 1 + t^5 + e^{-\frac{t}{10}} + \sin(t)$.

In order to illustrate the convergence rate in time of the discrete solution, we have used quadratic approximations in space and a mesh parameter $h = 0.01$. In this case, the stability condition (4.16) is always satisfied for the range of time steps considered. Thus, the choice of the Lagrange interpolant or of the Ritz-projection as approximation of the initial velocity give similar results.

In Figures 1(a)–(c) we report the convergences of the errors for the velocities ($\|\cdot\|_{L^\infty(0,T;L^2(\Omega))}$) and the pressures ($\|\cdot\|_{L^2(0,T;L^2(\Omega))}$) for the BDF1, Crank–Nicholson, and BDF2 schemes. In all the numerical examples, both the velocities and the pressures converge at the optimal rate ($O(\delta t)$ for BDF1 and $O(\delta t^2)$ for Crank–Nicholson and BDF2). The BDF2 scheme was initialized using one step of BDF1.

6.2. Behavior in the small time-step limit. In this subsection we illustrate the impact of the initial velocity approximation on the approximate pressures for small time steps. For nondiscrete divergence-free initial approximations, a pressure instability is predicted by Corollary 4.3 unless condition (4.16) is satisfied. In other words, pressure instabilities are expected for very small time steps.

We consider problem (2.1) in two dimensions and with nonhomogeneous boundary conditions. We set $\Omega = [0, 1] \times [0, 1]$, and the right-hand side $\boldsymbol{f}$ and the boundary data are chosen in order to ensure that the exact (steady) solution is given by

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

$$p(x, y, t) = \sin x \cos y + (\cos(1) - 1)\sin(1).$$

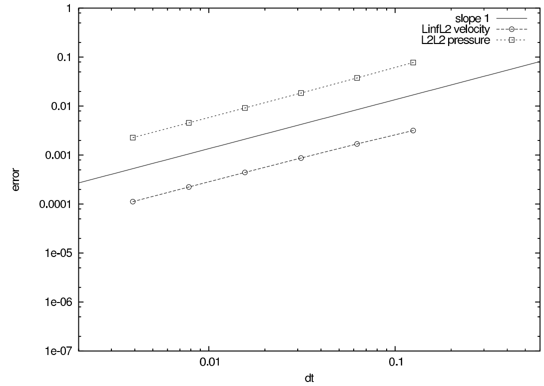
This numerical experiment is, in some degree, motivated by the work reported in [3] (see also [19]), where pressure instabilities, of a different nature, are illustrated for pressure stabilizations involving residuals of the PDEs (e.g., PSPG and GLS). Indeed, the time derivative involved in the residual perturbs the coercivity of the space semidiscrete operator, which leads to pressure instabilities for (sufficiently) small time steps (see [3]). Let us emphasize that, according to section 4, such instabilities do not appear here, in particular since the CIP pressure stabilization (and the other examples of subsection 3.1.1) are consistent without introducing the time derivative.

For different initial velocity approximations, we compare the behavior of the error in the pressure after one time step of the backward Euler scheme, i.e.,

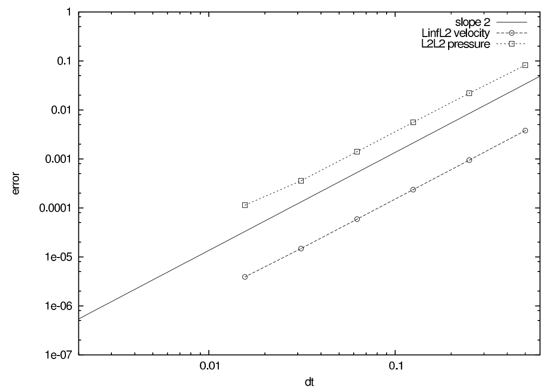
$$\delta t^{\frac{1}{2}} \|p(t_1) - p_h^1\|_Q.$$

We choose the initial data either as the Lagrange interpolant, $\boldsymbol{u}_h^0 = I_h^k \boldsymbol{u}_0$, or as the Ritz-projection, $\boldsymbol{u}_h^0 = P_h^k(\boldsymbol{u}_0, 0)$.

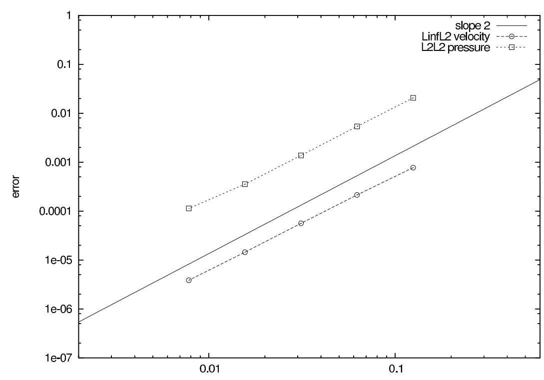
In Figure 2 we have reported the convergence history (in space) of the pressure error, at the first time step, using $\mathbb{P}_1/\mathbb{P}_1$ finite elements for different time step sizes. The pressure instability for small time steps is illustrated in Figure 4(a), where the initial velocity approximation is given in terms of the Lagrange interpolant. Indeed, we can observe that the pressure error has the right convergence rate in space, but it grows when the time step is decreased. On the other hand, as shown in Figure 4(b), the instability is eliminated when the initial velocity approximation is provided by the Ritz-projection, as stated in Corollary 4.3. In this case the error remains bounded (dominated by the space discretization) while reducing the time-step size.



(a) BDF1 scheme



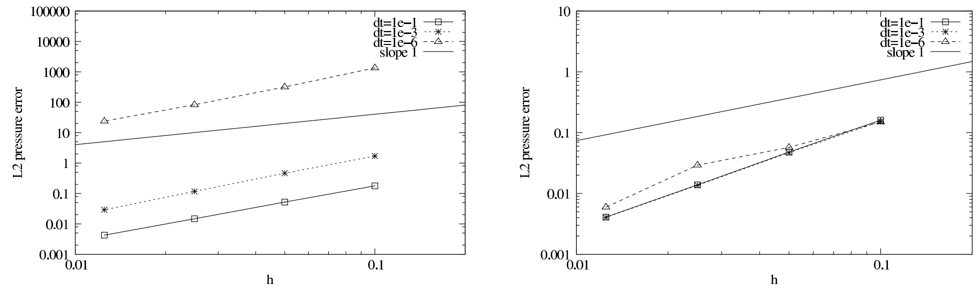
(b) Crank–Nicholson



(c) BDF2 scheme

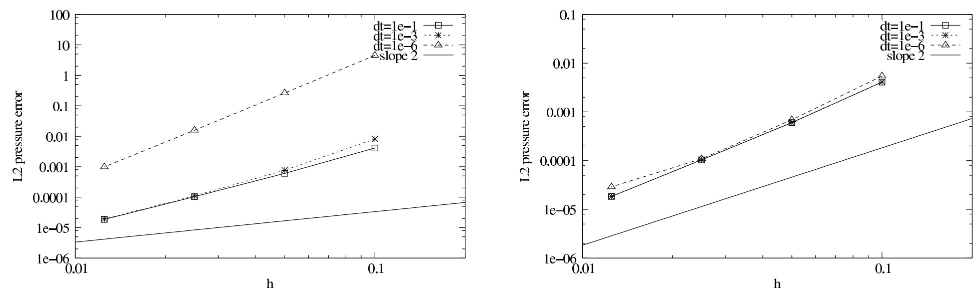
FIG. 1. *Convergence history in time:* $\mathbb{P}_2/\mathbb{P}_2$ *CIP stabilized finite elements.*

Similar results are found with $\mathbb{P}_2/\mathbb{P}_2$ finite elements, as shown in Figure 3. In particular, we can notice, from Figures 2(a) and 3(a), that for quadratic approximations the pressure instability shows up only for very small time steps. As a matter of fact, condition (4.16) is less restrictive for quadratic than for affine velocity approximations of smooth initial data. Finally, some pressure contours are reported in



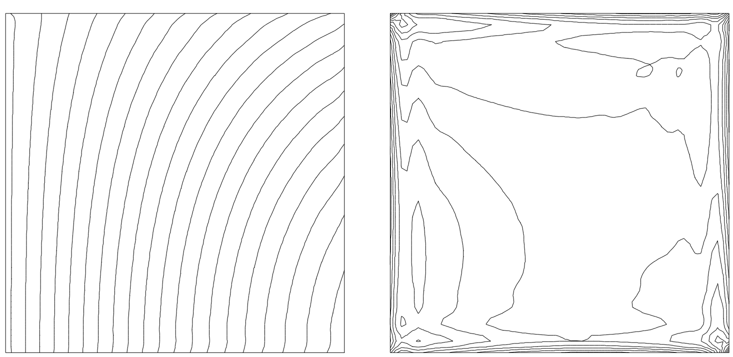
(a) Lagrange interpolant $\boldsymbol{u}_h^0 = I_h^1 \boldsymbol{u}_0$ (b) Ritz-projection $\boldsymbol{u}_h^0 = P_h^1(\boldsymbol{u}_0, 0)$

FIG. 2. *Convergence history:* $\mathbb{P}_1/\mathbb{P}_1$ *finite elements.*



(a) Lagrange interpolant $\boldsymbol{u}_h^0 = I_h^2 \boldsymbol{u}_0$ (b) Ritz-projection $\boldsymbol{u}_h^0 = P_h^2(\boldsymbol{u}_0, 0)$

FIG. 3. *Convergence history:* $\mathbb{P}_2/\mathbb{P}_2$ *finite elements.*

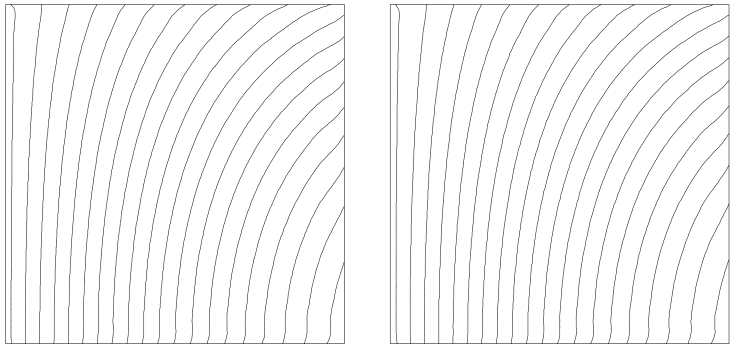


(a) $\delta t = 10^{-1}$ (b) $\delta t = 10^{-6}$

FIG. 4. *Pressure contour lines with* $\mathbb{P}_2/\mathbb{P}_2$ *finite elements in a* 40×40 *mesh:* $\boldsymbol{u}_h^0 = I_h^2 \boldsymbol{u}_0$.

Figure 4 for the Lagrange interpolation, and in Figure 5 for the Ritz-projection. The pressure degradation is clearly visible in Figure 4, whereas with the Ritz-projection initialization (Figure 5) the pressure remains unconditionally stable.

7. Conclusion. In this paper we have proved unconditional stability and optimal error estimates, in natural norms, for pressure stabilized finite element approximations of the transient Stokes problem. It should be noted that the extension of



(a) $\delta t = 10^{-1}$ (b) $\delta t = 10^{-6}$

FIG. 5. *Pressure contour lines with* $\mathbb{P}_2/\mathbb{P}_2$ *finite elements in a* 40×40 *mesh:* $\boldsymbol{u}_h^0 = P_h^2(\boldsymbol{u}_0, 0)$.

the present results to mixed formulations of the Poisson problem is straightforward. We have shown that for small initial time steps the use of a pressure stabilization dependent Ritz-projection, for the initial data, is essential to avoid pressure instabilities, unless a condition between time and space discretization parameters is satisfied. From the analysis, we also conclude that a second order scheme (e.g., BDF2) can be initialized (without optimality loss) using a first step with BDF1, *provided* that the Ritz-projection (3.16) is used for the initial data.

It is interesting to note that for low order elements the weakly consistent stabilization operators still yield optimal convergence in time when used with a second order scheme. However, in the case when streamline upwind Petrov–Galerkin (SUPG)-type stabilization is used for the convective term, the convergence order in time will be lost unless full consistency is guaranteed in the stabilization term. This is why SUPG-type stabilizations prompt space time finite element formulations with discontinuous approximation in time.

Some of the methods described in subsection 3.1.1, on the other hand, may be extended to the case of Oseen's equations, handling all Reynolds numbers, by applying the same type of stabilizing term for the convection (see [13, 6, 16, 7] for details).

REFERENCES

[1] S. Badia and R. Codina, *On a Multiscale Approach to the Transient Stokes Problem. Transient Subscales and Anisotropic Space-Time Discretization*, preprint, 2007.

[2] R. Becker and M. Braack, *A finite element pressure gradient stabilization for the Stokes equations based on local projections*, Calcolo, 38 (2001), pp. 173–199.

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

[4] D. Boffi and L. Gastaldi, *Analysis of finite element approximation of evolution problems in mixed form*, SIAM J. Numer. Anal., 42 (2004), pp. 1502–1526.

[5] M. Boman, *Estimates for the* L_2*-projection onto continuous finite element spaces in a weighted* L_p*-norm*, BIT, 46 (2006), pp. 249–260.

[6] M. Braack and E. Burman, *Local projection stabilization for the Oseen problem and its interpretation as a variational multiscale method*, SIAM J. Numer. Anal., 43 (2006), pp. 2544–2566.

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

[8] J. H. Bramble, J. E. Pasciak, and O. Steinbach, *On the stability of the L^2 projection in $H^1(\Omega)$*, Math. Comp., 71 (2002), pp. 147–156.

[9] F. Brezzi and M. Fortin, *A minimal stabilisation procedure for mixed finite element methods*, Numer. Math., 89 (2001), pp. 457–491.

[10] F. Brezzi and J. Pitkäranta, *On the stabilization of finite element approximations of the Stokes equations*, in Efficient Solutions of Elliptic Systems (Kiel, 1984), Notes Numer. Fluid Mech. 10, Vieweg, Braunschweig, 1984, pp. 11–19.

[11] A. N. Brooks and T. J. R. Hughes, *Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier–Stokes equations*, Comput. Methods Appl. Mech. Engrg., 32 (1982), pp. 199–259.

[12] E. Burman, *Pressure projection stabilizations for Galerkin approximations of Stokes and Darcy's problem*, Numer. Methods Partial Differential Equations, 24 (2008), pp. 127–143.

[13] E. Burman, M. A. Fernández, and P. Hansbo, *Continuous interior penalty finite element method for Oseen's equations*, SIAM J. Numer. Anal., 44 (2006), pp. 1248–1274.

[14] E. Burman and P. Hansbo, *Edge stabilization for the generalized Stokes problem: A continuous interior penalty method*, Comput. Methods Appl. Mech. Engrg., 195 (2006), pp. 2393–2410.

[15] P. Clément, *Approximation by finite element functions using local regularization*, Rev. Française Automat. Informat. Recherche Opérationnelle Sér. RAIRO Analyse Numérique, 9 (1975), pp. 77–84.

[16] R. Codina, *Analysis of a stabilized finite element approximation of the Oseen equations using orthogonal subscales*, Appl. Numer. Math., 58 (2008), pp. 2413–2430.

[17] R. Codina and J. Blasco, *A finite element formulation for the Stokes problem allowing equal velocity-pressure interpolation*, Comput. Methods Appl. Mech. Engrg., 143 (1997), pp. 373–391.

[18] R. Codina and J. Blasco, *Stabilized finite element method for the transient Navier-Stokes equations based on a pressure gradient projection*, Comput. Methods Appl. Mech. Engrg., 182 (2000), pp. 277–300.

[19] R. Codina, J. Principe, O. Guasch, and S. Badia, *Time dependent subscales in the stabilized finite element approximation of incompressible flow problems*, Comput. Methods Appl. Mech. Engrg., 196 (2007), pp. 2413–2430.

[20] M. Crouzeix and V. Thomée, *The stability in L_p and W_p^1 of the L_2-projection onto finite element function spaces*, Math. Comp., 48 (1987), pp. 521–532.

[21] C. R. Dohrmann and P. B. Bochev, *A stabilized finite element method for the Stokes problem based on polynomial pressure projections*, Internat. J. Numer. Methods Fluids, 46 (2004), pp. 183–201.

[22] A. Ern and J.-L. Guermond, *Theory and Practice of Finite Elements*, Appl. Math. Sci. 159, Springer-Verlag, New York, 2004.

[23] L. P. Franca and S. L. Frey, *Stabilized finite element methods.* II. *The incompressible Navier–Stokes equations*, Comput. Methods Appl. Mech. Engrg., 99 (1992), pp. 209–233.

[24] J. Freund and R. Stenberg, *On weakly imposed boundary conditions for second order problems*, in Proceedings of the Ninth International Conference on Finite Elements in Fluids, M. Morandi Cecchi et al., eds., Venice, 1995, pp. 327–336. Available online at http://math.tkk.fi/~rstenber/Publications/Venice95.pdf.

[25] V. Girault and P. A. Raviart, *Finite Element Methods for Navier–Stokes Equations*, Comput. Math. 5, Springer-Verlag, Berlin, 1986.

[26] F. Hecht, O. Pironneau, A. Le Hyaric, and K. Ohtsuka, *FreeFem++ v.* 2.11. *User's Manual*, Laboratoire J. L. Lions, University of Paris 6, Paris, France.

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

[28] G. Matthies, P. Skrzypacz, and L. Tobiska, *A Unified Analysis for Local Projection Stabilisations Applied to the Oseen Problem*, Technical report, Otto-von-Guericke-Universität Magdeburg, Magdeburg, Germany, 2007.

[29] J. Nitsche, *Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind*, Abh. Math. Sem. Univ. Hamburg, 36 (1971), pp. 9–15.

[30] A. Quarteroni and A. Valli, *Numerical Approximation of Partial Differential Equations*, Comput. Math. 23, Springer-Verlag, Berlin, 1994.

[31] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493.

[32] D. SILVESTER, *Optimal low order finite element methods for incompressible flow*, Comput. Methods Appl. Mech. Engrg., 111 (1994), pp. 357–368.

[33] M. SURI, *Mixed Variational Principles for Time-Dependent Problems*, Ph.D. thesis, Carnegie-Mellon University, Pittsburgh, PA, 1983.

[34] V. THOMÉE, *Galerkin Finite Element Methods for Parabolic Problems*, 2nd ed., Comput. Math. 25, Springer-Verlag, Berlin, 2006.