SIAM J. NUMER. ANAL.
Vol. 58, No. 1, pp. 410–429

UNIFORM IN TIME ERROR ESTIMATES FOR A FINITE ELEMENT METHOD APPLIED TO A DOWNSCALING DATA ASSIMILATION ALGORITHM FOR THE NAVIER–STOKES EQUATIONS*

BOSCO GARCÍA-ARCHILLA†, JULIA NOVO‡, AND EDRISS S. TITI§

Abstract. In this paper we analyze a finite element method applied to a continuous downscaling data assimilation algorithm for the numerical approximation of the two- and three-dimensional Navier–Stokes equations corresponding to given measurements on a coarse spatial scale. For representing the coarse mesh measurements we consider different types of interpolation operators including a Lagrange interpolant. We obtain uniform-in-time estimates for the error between a finite element approximation and the reference solution corresponding to the coarse mesh measurements. We consider both the case of a plain Galerkin method and a Galerkin method with grad-div stabilization. For the stabilized method we prove error bounds in which the constants do not depend on inverse powers of the viscosity. Some numerical experiments illustrate the theoretical results.

Key words. data assimilation, downscaling, Navier–Stokes equations, uniform-in-time error estimates, mixed finite elements

AMS subject classifications. 35Q30, 65M12, 65M15, 65M20, 65M60, 65M70, 76B75

DOI. 10.1137/19M1246845

1. Introduction.

Data assimilation refers to a class of techniques that combine experimental data and simulation in order to obtain better predictions in a physical system. There is a vast literature on data assimilation methods, especially in recent years (see, e.g., [4], [33], [35], [41], and the references therein). One of these techniques is *nudging*, where a penalty term is added in order to drive the approximate solution toward coarse mesh or large scale spatial observations of the data. In a recent work [7], a new approach, known as continuous data assimilation, is introduced for a large class of dissipative partial differential equations, including Rayleigh–Bénard convection [19], the planetary geostrophic ocean dynamics model [20], etc. (see also references therein). Continuous data assimilation has also been used in numerical studies, for example, with the Chafee–Infante reaction-diffusion equation, the Kuramoto–Sivashinsky equation (in the context of feedback control) [36], Rayleigh–Bénard convection equations [3], [18], and the Navier–Stokes equations [25], [28]. However, there is much less numerical analysis of this technique. The present work concerns the numerical analysis of continuous data assimilation

*Received by the editors February 25, 2019; accepted for publication (in revised form) August 30, 2019; published electronically January 21, 2020.
https://doi.org/10.1137/19M1246845
Funding: The first author's research is supported by Spanish MICINyU under grant PGC2018-096265-B-I00. The second author's research is supported by Spanish MINECO under grants MTM2016-78995-P (AEI/FEDER, UE) and VA024P17 (Junta de Castilla y Leon, ES), cofinanced by FEDER funds. The third author's research is supported in part by the ONR grant N00014-15-1-2333, the Einstein Stiftung/Foundation–Berlin through the Einstein Visiting Fellow Program, and by the John Simon Guggenheim Memorial Foundation.
†Departamento de Matemática Aplicada II, Universidad de Sevilla, Sevilla, Spain (bosco@esi.us.es).
‡Departamento de Matemáticas, Universidad Autónoma de Madrid, Spain (julia.novo@uam.es).
§Department of Mathematics, Texas A&M University, College Station, TX 77843, and Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, UK, and Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel (titi@math.tamu.edu, Edriss.Titi@damtp.cam.ac.uk).

for the Navier–Stokes equations when discretized with mixed finite element (MFE) methods.

To be more precise, we consider the Navier–Stokes equations

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

in a bounded domain $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$. In (1.1), $\boldsymbol{u}$ is the velocity field, p the kinematic pressure, and $\nu > 0$ the kinematic viscosity coefficient and $\boldsymbol{f}$ represents the accelerations due to external body forces acting on the fluid. The Navier–Stokes equations (1.1) must be complemented with boundary conditions. For simplicity, we only consider homogeneous Dirichlet boundary conditions $\boldsymbol{u} = \boldsymbol{0}$ on $\partial\Omega$.

Following [37] we consider given coarse spatial scale measurements, corresponding to a solution $\boldsymbol{u}$ of (1.1), observed at a coarse spatial mesh. The measurements are assumed to be continuous in time and error-free. We denote by $I_H(\boldsymbol{u})$ the operator used for interpolating these measurements, where H denotes the resolution of the coarse spatial mesh. Since the initial condition for $\boldsymbol{u}$ is missing one cannot compute $\boldsymbol{u}$ by simulating (1.1) directly. To overcome this difficulty it was suggested in [7] to consider instead a solution $\boldsymbol{v}$ of the following approximating system:

$$\begin{aligned} \partial_t \boldsymbol{v} - \nu\Delta \boldsymbol{v} + (\boldsymbol{v}\cdot\nabla)\boldsymbol{v} + \nabla \tilde{p} &= \boldsymbol{f} - \beta(I_H(\boldsymbol{v}) - I_H(\boldsymbol{u})), \text{ in } (0,T]\times\Omega, \\ \nabla\cdot\boldsymbol{v} &= 0, \text{ in } (0,T]\times\Omega, \end{aligned} \tag{1.2}$$

where β is the relaxation (nudging) parameter.

In the case of the Navier–Stokes equations (and indeed, of many other nonlinear dissipative systems), it is well-known that for relatively not so small Reynolds numbers, solutions are unstable and even chaotic. For this reason, it is expected that any small error in the initial data could lead to exponentially growing error in the solutions. Notably, the instabilities in the Navier–Stokes equations occur at the large spatial scales, while the fine scales are stabilized by the viscosity. For this reason once the large spatial scales are stabilized, as is done in the proposed downscaling data assimilation approximation, (1.2), the corresponding solutions are stable and converge to the same solution $\boldsymbol{u}$ that is corresponding to $I_H(\boldsymbol{u})$. This is the very reason that small errors are not magnified in time, and it allows one to obtain uniform in time error bounds.

In this paper we consider a semidiscretization in space with inf-sup stable mixed finite elements for (1.2) and analyze two different methods: the Galerkin method and the Galerkin method and grad-div stabilization. Grad-div stabilization was originally proposed in [21] to improve the conservation of mass in finite element methods (see also [38], [39]) However, it has been observed in the simulation of turbulent flows, [32], [42], that using only grad-div stabilization produced stable (nonoscillating) simulations. We prove uniform-in-time error estimates for approximating the unknown reference solution, $\boldsymbol{u}$, that corresponds to the coarse spatial scale measurement $I_H(\boldsymbol{u})$. For the Galerkin method without grad-div stabilization, the spatial error bounds we prove are optimal, in the sense that the rate of convergence is that of the best interpolant. In the case in which we add grad-div stabilization, as in [15], [16], we get error bounds in which the error constants do not depend on inverse powers of the viscosity parameter ν. This fact is of importance in many applications where viscosity is orders of magnitude smaller than the velocity (i.e., large Rynolds number). The convergence rates we prove in our error bounds are sharp and confirmed by numerical experiments.

We now comment on the analysis of numerical methods for (1.2). In [37], a semidiscrete postprocessed Galerkin spectral method for the two-dimensional Navier–Stokes equations is studied. Under suitable conditions on the nudging parameter β and the coarse mesh resolution H, uniform-in-time error estimates are obtained for the difference between the numerical approximation to $\boldsymbol{v}$ and $\boldsymbol{u}$. Furthermore, the use of a postprocessing technique introduced in [23], [24] allows for higher convergence rates than a standard spectral Galerkin method. A fully discrete method for the spatial discretization in [37] is analyzed in [31], where the backward Euler method is used for time discretization. Fully implicit and semi-implicit methods are considered, and optimal uniform-in-time error estimates are obtained with the same convergence rate in space as in [37].

More closely related to the present work are [34] and [40]. In [40] they only analyze linear problems and, for the proof of the results on the Navier–Stokes equations they present, they refer to [34] with some differences that they point out. They also present a wide collection of numerical experiments. In [34], the authors consider fully discrete approximations to (1.2), where the spatial discretization is performed with an MFE Galerkin method plus grad-div stabilization. A second order implicit/explicit (IMEX) in time scheme is analyzed in [34], and, as in [31], [37], and the present paper, uniform-in-time error bounds are obtained. Compared with [34], for the same convergence rate, the error bounds in the present paper have constants that do not depend on inverse powers of the viscosity parameter ν (Theorem 3.3) or, for similar error constants, error bounds in the present paper have an order of convergence one unit larger (Theorem 3.2 below). Also, the analysis in [34] is restricted to $I_H\boldsymbol{u}$ being an interpolant for nonsmooth functions (Clément, Scott–Zhang, etc.), since it makes explicit use of bound (2.19), which is not valid for nodal (Lagrange) interpolation (neither is it for (2.20)). In the present paper, we prove error bounds for the case in which (2.19) holds, but also for the case in which $I_H\boldsymbol{u}$ is a standard Lagrange interpolant (Theorem 3.12 below). To our knowledge, this is the first time in the literature where such a kind of bounds are proved. Also, compared with [34] and [40], we remove the upper bound assumed on the nudging parameter β. The authors of [34] had observed (see [34, Remark 3.8]) that the upper bound they required in the analysis does not hold in the numerical experiments and they state that a different approach to the analysis should be used to remove the upper bound on β. An analogous upper bound on β appears also in [31] and [37], where the value of H depends on the inverse of the nudging parameter β, which means that increasing the value of β would require a smaller value of H.

Although the analysis of the present paper could be extended to fully discrete methods following, for example, the techniques in [15], [16], we believe that the new ideas introduced in the present paper are easier to understand in the framework of the semidiscrete methods. The extension of the analysis of the present paper to the fully discrete case will be a subject of future work.

The rest of the paper is as follows. Section 2 is devoted to preliminary material, and in section 3 we introduce and analyze the finite element method for (1.2) with and without grad-div stabilization. In subsection 3.1 we analyze the case in which $I_H\boldsymbol{u}$ is the standard Lagrange interpolant. Finally, in section 4 some numerical experiments are shown to illustrate the theoretical results.

2. Preliminaries and notation. Throughout the paper, $W^{s,p}(D)$ will denote the Sobolev space of real-valued functions defined on the domain $D \subset \mathbb{R}^d$ with distributional derivatives of order up to s in $L^p(D)$. We denote by $|\cdot|_{s,p,D}$ a standard seminorm, and, following [14], for $W^{s,p}(D)$ we will use the norm $\|\cdot\|_{s,p,D}$ defined by

$$\|f\|_{s,p,D}^p = \sum_{j=0}^{s} |D|^{\frac{p(j-s)}{d}} \, |f|_{j,p,D}^p \,,$$

where $|D|$ stands for the Lebesgue measure of D so that $\|f\|_{m,p,D} \, |D|^{\frac{m}{d}-\frac{1}{p}}$ is scale invariant. If s is not a positive integer, $W^{s,p}(D)$ is defined by interpolation [1]. In the case $s = 0$ one has $W^{0,p}(D) = L^p(D)$. As is standard, $W^{s,p}(D)^d$ will be endowed with the product norm and, since no confusion can arise, it will be denoted again by $\|\cdot\|_{W^{s,p}(D)}$. The case $p = 2$ will be distinguished by using $H^s(D)$ to denote the space $W^{s,2}(D)$. The space $H_0^1(D)$ is the closure in $H^1(D)$ of the set of infinitely differentiable functions with compact support in D. For simplicity, $\|\cdot\|_s$ (resp., $|\cdot|_s$) is used to denote the norm (resp., seminorm) both in $H^s(\Omega)$ or $H^s(\Omega)^d$. The exact meaning will be clear by the context. The inner product of $L^2(\Omega)$ or $L^2(\Omega)^d$ will be denoted by $(\cdot,\cdot)$ and the corresponding norm by $\|\cdot\|_0$ in general D is skipped in the notation for the norm when $D = \Omega$. For vector-valued functions, the same conventions will be used as before. The norm of the dual space $H^{-1}(\Omega)$ of $H_0^1(\Omega)$ is denoted by $\|\cdot\|_{-1}$. As usual, $L^2(\Omega)$ is always identified with its dual, so one has $H_0^1(\Omega) \subset L^2(\Omega) \subset H^{-1}(\Omega)$ with compact injection. The following Sobolev's embedding [1] will be used in the analysis: For $s > 0$, let $1 \le p < d/s$ and q be such that $\frac{1}{q} = \frac{1}{p} - \frac{s}{d}$. Then, there exists a positive scale invariant constant c_s such that

$$\|v\|_{L^{q'}(\Omega)} \le c_s |\Omega|^{\frac{s}{d}-\frac{1}{p}+\frac{1}{q'}} \|v\|_{W^{s,p}(\Omega)}, \qquad \frac{1}{q'} \ge \frac{1}{q}, \quad \forall v \in W^{s,p}(\Omega). \tag{2.1}$$

If $p > d/s$, the above relation is valid for $q' = \infty$. A similar embedding inequality holds for vector-valued functions.

We will also use the following interpolation inequality (see, e.g., [14, Formula (6.7)] and [22, Exercise II.2.9])

$$\|v\|_{L^{2d/(d-1)}(\Omega)} \le c_1 \, \|v\|_0^{1/2} \, \|v\|_1^{1/2} \qquad \forall v \in H^1(\Omega) \tag{2.2}$$

(where, for simplicity, by enlarging the constants if necessary, we may take the constant c_1 in (2.2) equal to c_s in (2.1) for $s = 1$) and Agmon's inequality

$$\|v\|_\infty \le c_{\mathrm{A}} \, \|v\|_{d-2}^{1/2} \, \|v\|_2^{1/2}, \qquad d = 2, 3, \qquad \forall v \in H^2(\Omega). \tag{2.3}$$

The case $d = 2$ is a direct consequence of [2, Theorem 3.9]. For $d = 3$, a proof for domains of class C^2 can be found in [14, Lemma 4.10]. By means of the Calderón extension theorem (see, e.g., [1, Theorem 4.32] the proof is also valid for bounded Lipschitz domains. Finally, we will use Poincaré's inequality,

$$\|v\|_0 \le c_P |\Omega|^{1/d} \|\nabla v\|_0 \qquad \forall v \in H_0^1(\Omega), \tag{2.4}$$

where the constant c_P can be taken $c_P \le \sqrt{2}/2$. Denoting by

$$\hat{c}_P = 1 + c_P^2, \tag{2.5}$$

observe that from (2.4) it follows that

$$\|v\|_1 \le (\hat{c}_P)^{1/2} \|\nabla v\|_0 \qquad \forall v \in H_0^1(\Omega). \tag{2.6}$$

In all previous inequalities, the constants c_s, c_1, c_A, and c_P are scale-invariant, as will be the case for all constants in the present paper unless explicitly stated otherwise.

Let $\mathcal{H}$ and V be the Hilbert spaces $\mathcal{H} = \{\boldsymbol{u} \in \left(L^2(\Omega)\right)^d \,|\, \mathrm{div}(\boldsymbol{u}) = 0,\, \boldsymbol{u}\cdot n_{|\partial\Omega} = 0\}$, $V = \{\boldsymbol{u} \in \left(H_0^1(\Omega)\right)^d \,|\, \mathrm{div}(\boldsymbol{u}) = 0\}$, endowed with the inner product of $L^2(\Omega)^d$ and $H_0^1(\Omega)^d$, respectively.

Let $\mathcal{T}_h = (\tau_j^h, \phi_j^h)_{j\in J_h}$, $h > 0$, be a family of partitions of suitable domains Ω_h, where h is the maximum diameter of the elements $\tau_j^h \in \mathcal{T}_h$, and ϕ_j^h are the mappings from the reference simplex τ_0 onto τ_j^h. We shall assume that the partitions are shape-regular and quasi-uniform. Letting $r \geq 2$, we consider the finite-element spaces

$$
\begin{aligned}
S_{h,r} &= \left\{\chi_h \in \mathcal{C}\left(\overline{\Omega}_h\right) \,\big|\, \chi_{h|\tau_j^h} \circ \phi_j^h \,\in\, P^{r-1}(\tau_0)\right\} \subset H^1(\Omega_h),\\
S_{h,r}^0 &= S_{h,r} \cap H_0^1(\Omega_h),
\end{aligned}
$$

where $P^{r-1}(\tau_0)$ denotes the space of polynomials of degree at most $r-1$ on τ_0. For $r = 1$, $S_{h,1}$ stands for the space of piecewise constants.

When Ω has polygonal or polyhedral boundary $\Omega_h = \Omega$ and mappings ϕ_j^h from the reference simplex are affine. When Ω has a smooth boundary, for the purpose of analysis we will assume that Ω_h exactly matches Ω, as is done, for example, in [12], [43], although at a price of a more complex analysis discrepancies between Ω_h and Ω can also be taken into account (see, e.g., [6], [44]).

We shall denote by $(X_{h,r}, Q_{h,r-1})$ the MFE pair known as Hood–Taylor elements [10, 46], when $r \geq 3$, where

$$
X_{h,r} = \left(S_{h,r}^0\right)^d, \quad Q_{h,r-1} = S_{h,r-1} \cap L^2(\Omega_h)/\mathbb{R}, \quad r \geq 3,
$$

and, when $r = 2$, the MFE pair known as the minielement [11], where $Q_{h,1} = S_{h,2} \cap L^2(\Omega_h)/\mathbb{R}$, and $X_{h,2} = (S_{h,2}^0)^d \oplus \mathbb{B}_h$. Here, $\mathbb{B}_h$ is spanned by the bubble functions $\boldsymbol{b}_\tau$, $\tau \in \mathcal{T}_h$, defined by $\boldsymbol{b}_\tau(x) = (d+1)^{d+1}\lambda_1(x)\cdots\lambda_{d+1}(x)$, if $x \in \tau$ and 0 elsewhere, where $\lambda_1(x), \ldots, \lambda_{d+1}(x)$ denote the barycentric coordinates of x. For these elements a uniform inf-sup condition is satisfied (see [10]), that is, there exists a constant $\beta_{\mathrm{is}} > 0$ independent of the mesh grid size h such that

$$
\inf_{q_h \in Q_{h,r-1}} \sup_{v_h \in X_{h,r}} \frac{(q_h, \nabla\cdot v_h)}{\|v_h\|_1 \|q_h\|_{L^2/\mathbb{R}}} \geq \beta_{\mathrm{is}}. \tag{2.7}
$$

The velocity will be approximated by elements of the discrete divergence-free space

$$
V_{h,r} = X_{h,r} \cap \left\{\chi_h \in H_0^1(\Omega_h)^d \mid (q_h, \nabla\cdot\chi_h) = 0 \quad \forall q_h \in Q_{h,r-1}\right\}.
$$

For each fixed time $t \in [0,T]$ the solution (u,p) of (1.1) is also the solution of a Stokes problem with right-hand side $\boldsymbol{f} - \boldsymbol{u}_t - (\boldsymbol{u}\cdot\nabla)\boldsymbol{u}$. We will denote by $(\boldsymbol{s}_h, q_h) \in (X_{h,r}, Q_{h,r-1})$ its MFE approximation satisfying

$$
\begin{aligned}
\nu(\nabla\boldsymbol{s}_h, \nabla\boldsymbol{\varphi}_h) - (q_h, \nabla\cdot\boldsymbol{\varphi}_h) &= \nu(\nabla u, \nabla\boldsymbol{\varphi}_h) - (p, \nabla\cdot\boldsymbol{\varphi}_h)\\
&= (\boldsymbol{f} - \boldsymbol{u}_t - (\boldsymbol{u}\cdot\nabla\boldsymbol{u}), \boldsymbol{\varphi}_h) \quad \forall\boldsymbol{\varphi}_h \in X_{h,r},\\
(\nabla\cdot\boldsymbol{s}_h, \psi_h) = 0 \quad \forall\psi_h \in Q_{h,r-1}.&
\end{aligned} \tag{2.8}
$$

We observe that $\boldsymbol{s}_h = S_h(\boldsymbol{u}) : V \to V_{h,r}$ is the discrete Stokes projection of the solution $(\boldsymbol{u}, p)$ of (1.1) (see [29]) and satisfies

$$
\nu(\nabla S_h(\boldsymbol{u}), \nabla\boldsymbol{\varphi}_h) = \nu(\nabla\boldsymbol{u}, \nabla\boldsymbol{\varphi}_h) - (p, \nabla\cdot\boldsymbol{\varphi}_h) = (\boldsymbol{f} - \boldsymbol{u}_t - (\boldsymbol{u}\cdot\nabla)\boldsymbol{u}, \boldsymbol{\varphi}_h) \;\; \forall\, \boldsymbol{\varphi}_h \in V_{h,r}.
$$

The following bound holds:

$$\|\boldsymbol{u}-\boldsymbol{s}_h\|_0 + h\|\boldsymbol{u}-\boldsymbol{s}_h\|_1 \le C N_j(\boldsymbol{u},p)h^j, \qquad 1 \le j \le r, \tag{2.9}$$

where here and in what follows, for $\boldsymbol{v} \in V \cap H^j(\Omega)^d$ and $q \in L_0^2(\Omega) \cap H^{j-1}(\Omega)$ we denote

$$N_j(\boldsymbol{v},q) = \|\boldsymbol{v}\|_j + \nu^{-1}\|q\|_{H^{j-1}/\mathbb{R}}, \qquad j \ge 1. \tag{2.10}$$

The proof of (2.9) when $\Omega = \Omega_h$ can be found in [30, Lemma 5.3]. For the general case, superparametric approximation at the boundary is assumed (see [5]). Under the same conditions, the bound for the pressure is (cf. [27])

$$\|p - q_h\|_{L^2/\mathbb{R}} \le C_{\beta_{\rm is}} \nu N_j(\boldsymbol{u},p)h^{j-1}, \qquad 1 \le j \le r, \tag{2.11}$$

where the constant $C_{\beta_{\rm is}}$ depends on the constant $\beta_{\rm is}$ in (2.7). Assuming that Ω is of class $\mathcal{C}^m$, with $m \ge 3$, and using standard duality arguments and (2.9), one obtains

$$\|\boldsymbol{u}-\boldsymbol{s}_h\|_{-s} \le C N_r(\boldsymbol{u},p)h^{r+s}, \qquad 0 \le s \le \min(r-2,1). \tag{2.12}$$

We also consider a modified Stokes projection that was introduced in [15] and that we denote by $\boldsymbol{s}_h^m : V \to V_{h,r}$ satisfying

$$\nu(\nabla \boldsymbol{s}_h^m, \nabla \boldsymbol{\varphi}_h) = (\boldsymbol{f} - \boldsymbol{u}_t - (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} - \nabla p, \boldsymbol{\varphi}_h) \quad \forall\, \boldsymbol{\varphi}_h \in V_{h,r}. \tag{2.13}$$

The following bound holds (see [15]):

$$\|\boldsymbol{u}-\boldsymbol{s}_h^m\|_0 + h\|\boldsymbol{u}-\boldsymbol{s}_h^m\|_1 \le C\|\boldsymbol{u}\|_j h^j, \qquad 1 \le j \le r. \tag{2.14}$$

Following [12], one can also obtain the following bound:

$$\|\nabla(\boldsymbol{u}-\boldsymbol{s}_h^m)\|_\infty \le C\|\nabla \boldsymbol{u}\|_\infty, \tag{2.15}$$

where C does not depend on ν. We will denote by $\pi_h p$ the L^2 projection of the pressure p onto $Q_{h,r-1}$. It holds that

$$\|p - \pi_h p\|_0 \le C h^{j-1}\|p\|_{H^{j-1}/\mathbb{R}}, \qquad 1 \le j \le r. \tag{2.16}$$

If the family of meshes is quasi-uniform, then the following inverse inequality holds for each $\boldsymbol{v}_h \in S_{h,r}$ (see, e.g., [13, Theorem 3.2.6]):

$$\|\boldsymbol{v}_h\|_{W^{m,p}(K)} \le c_{\rm inv} h_K^{n-m-d\left(\frac{1}{q}-\frac{1}{p}\right)}\|\boldsymbol{v}_h\|_{W^{n,q}(K)}, \tag{2.17}$$

where $0 \le n \le m \le 1$, $1 \le q \le p \le \infty$, and h_K is the diameter of $K \in \mathcal{T}_h$.

In what follows $I_h^{La}\boldsymbol{u} \in X_{h,r}$ will denote the Lagrange interpolant of a continuous function $\boldsymbol{u}$. The following bound can be found in [9, Theorem 4.4.4]:

$$|\boldsymbol{u} - I_h^{La}\boldsymbol{u}|_{W^{m,p}(K)} \le c_{\rm int} h^{n-m}|\boldsymbol{u}|_{W^{n,p}(K)}, \quad 0 \le m \le n \le k+1, \tag{2.18}$$

where $n > d/p$ when $1 < p \le \infty$ and $n \ge d$ when $p = 1$.

We will assume that the interpolation operator I_H is stable in L^2, that is,

$$\|I_H \boldsymbol{u}\|_0 \le c_0\|\boldsymbol{u}\|_0 \quad \forall \boldsymbol{u} \in L^2(\Omega)^d, \tag{2.19}$$

and that it satisfies the following approximation property:

$$\|\boldsymbol{u} - I_H\boldsymbol{u}\|_0 \le c_I H\|\nabla \boldsymbol{u}\|_0 \quad \forall \boldsymbol{u} \in H_0^1(\Omega)^d. \tag{2.20}$$

The Bernardi–Girault [8], Girault–Lions [26], or Scott–Zhang [45] interpolation operators satisfy (2.20) and (2.19). Notice that the interpolation can be on piecewise constants, as we use in the numerical experiments in section 4.

We remark that, for the error analysis, we do not need condition (3.105) in [37], i.e., we do not assume that $\|\boldsymbol{u} - I_H(\boldsymbol{u})\|_{-1} \le c_{-1}H\|\boldsymbol{u}\|_0$ for $\boldsymbol{u} \in L^2(\Omega)^d$.

3. The finite element method. We consider the following method to approximate (1.2). Find $(\boldsymbol{u}_h, p_h) \in X_{h,r} \times Q_{h,r-1}$ satisfying for all $(\boldsymbol{\varphi}_h, \psi_h) \in X_{h,r} \times Q_{h,r-1}$

$$\begin{aligned}(\dot{\boldsymbol{u}}_h, \boldsymbol{\varphi}_h) + \nu(\nabla \boldsymbol{u}_h, \nabla \boldsymbol{\varphi}_h) + b_h(\boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{\varphi}_h) + \mu(\nabla \cdot \boldsymbol{u}_h, \nabla \cdot \boldsymbol{\varphi}_h) + (\nabla p_h, \boldsymbol{\varphi}_h) \\ = (\boldsymbol{f}, \boldsymbol{\varphi}_h) - \beta(I_H(\boldsymbol{u}_h) - I_H(\boldsymbol{u}), I_H \boldsymbol{\varphi}_h), \\ (\nabla \cdot \boldsymbol{u}_h, \psi_h) = 0,\end{aligned} \tag{3.1}$$

where μ is a stabilization parameter that can be zero in case we do not stabilize the divergence or different from zero in case we add grad-div stabilization and $b_h(\cdot,\cdot,\cdot)$ is defined in the following way:

$$b_h(\boldsymbol{u}_h, \boldsymbol{v}_h, \boldsymbol{\varphi}_h) = ((\boldsymbol{u}_h \cdot \nabla)\boldsymbol{v}_h, \boldsymbol{\varphi}_h) + \frac{1}{2}(\nabla \cdot (\boldsymbol{u}_h)\boldsymbol{v}_h, \boldsymbol{\varphi}_h) \quad \forall\, \boldsymbol{u}_h, \boldsymbol{v}_h, \boldsymbol{\varphi}_h \in X_{h,r}.$$

Hereafter, we denote by $(\cdot,\cdot)$ both the inner product in L^2 and the duality action between H^{-1} and H_0^1, depending on the context. It is straightforward to verify that b_h enjoys the skew-symmetry property

$$b_h(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = -b_h(\boldsymbol{u}, \boldsymbol{w}, \boldsymbol{v}) \qquad \forall\, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in H_0^1(\Omega)^d. \tag{3.2}$$

Let us observe that taking $\boldsymbol{\varphi}_h \in V_{h,r}$ from (3.1) we get

$$\begin{aligned}(\dot{\boldsymbol{u}}_h, \boldsymbol{\varphi}_h) + \nu(\nabla \boldsymbol{u}_h, \nabla \boldsymbol{\varphi}_h) + b_h(\boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{\varphi}_h) + \mu(\nabla \cdot \boldsymbol{u}_h, \nabla \cdot \boldsymbol{\varphi}_h) \\ = (\boldsymbol{f}, \boldsymbol{\varphi}_h) - \beta(I_H(\boldsymbol{u}_h) - I_H(\boldsymbol{u}), I_H \boldsymbol{\varphi}_h).\end{aligned} \tag{3.3}$$

For the analysis below, we need to introduce the values $\overline{\mu}$ and $\overline{k}$, defined as follows:

$$\overline{\mu} = \begin{cases} 0 & \text{if } \mu = 0, \\ 1 & \text{otherwise,} \end{cases} \qquad \overline{k} = \begin{cases} 0 & \text{if } \mu = 0, \\ 1/\mu & \text{otherwise.} \end{cases} \tag{3.4}$$

The following lemma will be used for proving the main results of the section.

LEMMA 3.1. *Let $\boldsymbol{u}_h$ be the finite element approximation defined in* (3.3) *and let $\boldsymbol{w}_h, \boldsymbol{\tau}_h^1, \boldsymbol{\tau}_h^2 : [0,T] \to V_{h,r}$ be functions satisying*

$$\begin{aligned}(\dot{\boldsymbol{w}}_h, \boldsymbol{\varphi}_h) + \nu(\nabla \boldsymbol{w}_h, \nabla \boldsymbol{\varphi}_h) + b_h(\boldsymbol{w}_h, \boldsymbol{w}_h, \boldsymbol{\varphi}_h) + \mu(\nabla \cdot \boldsymbol{w}_h, \nabla \cdot \boldsymbol{\varphi}_h) \\ = (\boldsymbol{f}, \boldsymbol{\varphi}_h) + (\boldsymbol{\tau}_h^1, \boldsymbol{\varphi}_h) + \overline{\mu}(\boldsymbol{\tau}_h^2, \nabla \cdot \boldsymbol{\varphi}_h).\end{aligned} \tag{3.5}$$

Assume the quantity L defined in (3.14), *below, when $\mu = 0$, and in* (3.15), *below, when $\mu > 0$ is bounded. Then, if $\beta \geq 8L$ and H satisfies condition* (3.21), *below, the following bounds hold for $\boldsymbol{e}_h = \boldsymbol{u}_h - \boldsymbol{w}_h$:*

$$\begin{aligned}\|\boldsymbol{e}_h(t)\|_0^2 \leq e^{-\gamma t/2}\|\boldsymbol{e}_h(0)\|_0^2 + \int_0^t e^{-\gamma(t-s)/2}\left((1-\overline{\mu})\frac{2\hat{c}_P}{\nu} + \frac{\overline{\mu}}{L}\right)\|\boldsymbol{\tau}_h^1\|_{-1+\overline{\mu}}^2\, ds \\ + \int_0^t e^{-\gamma(t-s)/2}\left(\beta c_0^2\|\boldsymbol{u}(s) - \boldsymbol{w}_h(s)\|_0^2 + 2\overline{k}\|\tau_h^2\|_0^2\right)\, ds,\end{aligned} \tag{3.6}$$

where $\overline{\mu}$ and $\overline{k}$ are defined in (3.4), *and γ is defined in* (3.24) *below.*

Proof. Subtracting (3.5) from (3.3) we get the error equation

$$\begin{aligned}(\dot{\boldsymbol{e}}_h, \boldsymbol{\varphi}_h) + \nu(\nabla \boldsymbol{e}_h, \nabla \boldsymbol{\varphi}_h) + \beta(I_H \boldsymbol{e}_h, I_H \boldsymbol{\varphi}_h) + b_h(\boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{\varphi}_h) - b_h(\boldsymbol{w}_h, \boldsymbol{w}_h, \boldsymbol{\varphi}_h) \\ + \mu(\nabla \cdot \boldsymbol{e}_h, \nabla \cdot \boldsymbol{\varphi}_h) = \beta(I_H \boldsymbol{u} - I_H \boldsymbol{w}_h, I_H \boldsymbol{\varphi}_h) + (\boldsymbol{\tau}_h^1, \boldsymbol{\varphi}_h) + \overline{\mu}(\boldsymbol{\tau}_h^2, \nabla \cdot \boldsymbol{\varphi}_h) \quad \forall \boldsymbol{\varphi}_h \in V_{h,r}.\end{aligned} \tag{3.7}$$

Taking $\boldsymbol{\varphi}_h = \boldsymbol{e}_h$ in (3.7) we get

$$\frac{1}{2}\frac{d}{dt}\|\boldsymbol{e}_h\|_0^2 + \nu\|\nabla \boldsymbol{e}_h\|_0^2 + \beta\|I_H\boldsymbol{e}_h\|_0^2 + \mu\|\nabla\cdot\boldsymbol{e}_h\|_0^2 \le |b_h(\boldsymbol{u}_h,\boldsymbol{u}_h,\boldsymbol{e}_h) - b_h(\boldsymbol{w}_h,\boldsymbol{w}_h,\boldsymbol{e}_h)| \\ + \beta|(I_H\boldsymbol{u} - I_H\boldsymbol{w}_h, I_H\boldsymbol{e}_h)| + |(\boldsymbol{\tau}_h^1,\boldsymbol{e}_h)| + |\overline{\mu}(\boldsymbol{\tau}_h^2,\nabla\cdot\boldsymbol{e}_h)|. \tag{3.8}$$

We will bound the terms on the right-hand side of (3.8). For the nonlinear term and the truncation errors we argue differently depending on whether $\mu = 0$ or $\mu > 0$.

If $\mu = 0$, using the skew-symmetry property (3.2), (2.1), and (2.6), and when $d = 3$, we have

$$|b_h(\boldsymbol{u}_h,\boldsymbol{u}_h,\boldsymbol{e}_h) - b_h(\boldsymbol{w}_h,\boldsymbol{w}_h,\boldsymbol{e}_h)| = |b_h(\boldsymbol{e}_h,\boldsymbol{w}_h,\boldsymbol{e}_h)| \le \|\nabla\boldsymbol{w}_h\|_{L^{2d/(d-1)}}\|\boldsymbol{e}_h\|_{L^{2d}}\|\boldsymbol{e}_h\|_0 \\ + \frac{1}{2}|(\nabla\cdot\boldsymbol{e}_h)\boldsymbol{w}_h,\boldsymbol{e}_h)| \le \hat{c}_1\|\nabla\boldsymbol{w}_h\|_{L^{2d/(d-1)}}\|\nabla\boldsymbol{e}_h\|_0\|\boldsymbol{e}_h\|_0 + \frac{1}{2}\|\nabla\boldsymbol{e}_h\|_0\|\boldsymbol{w}_h\|_\infty\|\boldsymbol{e}_h\|_0 \\ \le \left(2\hat{c}_1^2\frac{\|\nabla\boldsymbol{w}_h\|_{L^{2d/(d-1)}}^2}{\nu} + \frac{\|\boldsymbol{w}_h\|_\infty^2}{\nu}\right)\|\boldsymbol{e}_h\|_0^2 + \frac{\nu}{4}\|\nabla\boldsymbol{e}_h\|_0^2, \tag{3.9}$$

where

$$\hat{c}_1 = (\hat{c}_P)^{1/2}c_1, \tag{3.10}$$

c_1 being the constant in (2.1) for $s = 1$. In the case $d = 2$ and noticing that $2d = 2d/(d-1)$, the first term on the right-hand side above, using (2.2), (2.6), and Young's inequality is bounded as follows:

$$\|\nabla\boldsymbol{w}_h\|_{L^{2d/(d-1)}}\|\boldsymbol{e}_h\|_{L^{2d}}\|\boldsymbol{e}_h\|_0 \le (\hat{c}_P)^{1/4}c_1\|\nabla\boldsymbol{w}_h\|_{L^{2d/(d-1)}}(\|\nabla\boldsymbol{e}_h\|_0\|\boldsymbol{e}_h\|_0)^{1/2}\|\boldsymbol{e}_h\|_0 \\ \le 3(\hat{c}_P)^{1/3}c_1^{4/3}\frac{\|\nabla\boldsymbol{w}_h\|_{L^{2d/(d-1)}}^{4/3}}{(4\nu)^{1/3}}\|\boldsymbol{e}_h\|_0^2 + \frac{\nu}{4}\|\nabla\boldsymbol{e}_h\|_0^2. \tag{3.11}$$

For the truncation error when $\mu = 0$, using (2.6) we get

$$|(\boldsymbol{\tau}_h^1,\boldsymbol{e}_h)| \le \|\boldsymbol{\tau}_h^1\|_{-1}\|\boldsymbol{e}_h\|_1 \le (\hat{c}_P)^{1/2}\|\boldsymbol{\tau}_h^1\|_{-1}\|\nabla\boldsymbol{e}_h\|_0 \le \frac{\hat{c}_P}{\nu}\|\boldsymbol{\tau}_h^1\|_{-1}^2 + \frac{\nu}{4}\|\nabla\boldsymbol{e}_h\|_0^2. \tag{3.12}$$

When $\mu \ne 0$, we bound the nonlinear term in the following way. Using again the skew-symmetry property (3.2) we get

$$|b_h(\boldsymbol{u}_h,\boldsymbol{u}_h,\boldsymbol{e}_h) - b_h(\boldsymbol{w}_h,\boldsymbol{w}_h,\boldsymbol{e}_h)| = |b_h(\boldsymbol{e}_h,\boldsymbol{w}_h,\boldsymbol{e}_h)| \le \|\nabla\boldsymbol{w}_h\|_\infty\|\boldsymbol{e}_h\|_0^2 \\ + \frac{1}{2}\|\nabla\cdot\boldsymbol{e}_h\|_0\|\boldsymbol{w}_h\|_\infty\|\boldsymbol{e}_h\|_0 \le \|\nabla\boldsymbol{w}_h\|_\infty\|\boldsymbol{e}_h\|_0^2 + \frac{\mu}{4}\|\nabla\cdot\boldsymbol{e}_h\|_0^2 + \frac{\|\boldsymbol{w}_h\|_\infty^2}{4\mu}\|\boldsymbol{e}_h\|_0^2. \tag{3.13}$$

In what follows we denote

$$L = \max_{t\ge 0}\left(2\frac{\hat{c}_1^{2d/3}\|\nabla\boldsymbol{w}_h(t)\|_{L^{2d/(d-1)}}^{2d/3}}{\nu^{(2d-3)/3}} + \frac{\|\boldsymbol{w}_h(t)\|_\infty^2}{\nu}\right) \quad \text{if} \quad \mu = 0, \tag{3.14}$$

$$L = 2\max_{t\ge 0}\left(\|\nabla\boldsymbol{w}_h(t)\|_\infty + \frac{\|\boldsymbol{w}_h(t)\|_\infty^2}{4\mu}\right) \quad \text{if} \quad \mu > 0. \tag{3.15}$$

Observe that in the case $\mu = 0$, bounding the factor $3(\hat{c}_P)^{1/3}/4^{1/3}$ in (3.11) by $2(\hat{c}_P)^{2/3}$ we have that the left-hand side of (3.9) can be bounded by $L\|\boldsymbol{e}_h\|_0^2 +$

$(\nu/2)\|\nabla \boldsymbol{e}_h\|_0^2$, and, in the case $\mu > 0$, the left-hand side of (3.13) is bounded by $(L/2)\|\boldsymbol{e}_h\|_0^2 + (\mu/4)\|\nabla\cdot \boldsymbol{e}_h\|_0^2$.

Next, we bound the truncation error when $\mu > 0$,

$$|(\boldsymbol{\tau}_h^1, \boldsymbol{e}_h)| + |\overline{\mu}(\boldsymbol{\tau}_h^2, \nabla\cdot \boldsymbol{e}_h)| \le \frac{1}{2L}\|\boldsymbol{\tau}_h^1\|_0^2 + \frac{L}{2}\|\boldsymbol{e}_h\|_0^2 + \overline{k}\|\tau_h^2\|_0^2 + \frac{\mu}{4}\|\nabla\cdot \boldsymbol{e}_h\|_0^2, \tag{3.16}$$

where $\overline{k}$ is defined in (3.4).

For the second term on the right-hand side of (3.8), applying (2.19) we get

$$\begin{aligned} \beta|(I_H\boldsymbol{u} - I_H\boldsymbol{w}_h, I_H\boldsymbol{e}_h)| &\le \beta c_0\|\boldsymbol{u} - \boldsymbol{w}_h\|_0\|I_H\boldsymbol{e}_h\|_0 \\ &\le \frac{\beta}{2}c_0^2\|\boldsymbol{u} - \boldsymbol{w}_h\|_0^2 + \frac{\beta}{2}\|I_H\boldsymbol{e}_h\|_0^2. \end{aligned} \tag{3.17}$$

Inserting (3.9), (3.11), (3.12), (3.13), (3.16), and (3.17) into (3.8) we get

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|\boldsymbol{e}_h\|_0^2 + (1+\overline{\mu})\frac{\nu}{2}\|\nabla \boldsymbol{e}_h\|_0^2 + \frac{\beta}{2}\|I_H\boldsymbol{e}_h\|_0^2 + \frac{\mu}{2}\|\nabla\cdot \boldsymbol{e}_h\|_0^2 \le L\|\boldsymbol{e}_h\|_0^2 \\ + \overline{k}\|\tau_h^2\|_0^2 + \frac{\beta}{2}c_0^2\|\boldsymbol{u} - \boldsymbol{w}_h\|_0^2 + \left((1-\overline{\mu})\frac{\hat{c}_P}{\nu} + \frac{\overline{\mu}}{2L}\right)\|\boldsymbol{\tau}_h^1\|_{-1+\overline{\mu}}^2. \end{aligned} \tag{3.18}$$

Now we bound

$$L\|\boldsymbol{e}_h\|_0^2 \le 2L\|I_He_h\|_0^2 + 2L\|(I - I_H)e_h\|_0^2.$$

Since we are assuming that $\beta \ge 8L$ we have that $\beta/2 - 2L \ge \beta/4$, so that taking into account that $1+\overline{\mu} \ge 1$ and $(\mu/2)\|\nabla\cdot \boldsymbol{e}_h\| \ge 0$ we get

$$\begin{aligned} \frac{d}{dt}\|\boldsymbol{e}_h\|_0^2 + \nu\|\nabla \boldsymbol{e}_h\|_0^2 - 4L\|(I - I_H)\boldsymbol{e}_h\|_0^2 + \frac{\beta}{2}\|I_H\boldsymbol{e}_h\|_0^2 \\ \le 2\overline{k}\|\tau_h^2\|_0^2 + \beta c_0^2\|\boldsymbol{u} - \boldsymbol{w}_h\|_0^2 + \left((1-\overline{\mu})\frac{2\hat{c}_P}{\nu} + \frac{\overline{\mu}}{L}\right)\|\boldsymbol{\tau}_h^1\|_{-1+\overline{\mu}}^2. \end{aligned} \tag{3.19}$$

For the second and third terms on the left-hand side above, applying (2.20) to the latter, we write

$$\nu\|\nabla \boldsymbol{e}_h\|_0^2 - 4L\|(I - I_H)e_h\|_0^2 \ge \nu\|\nabla \boldsymbol{e}_h\|_0^2 - 4Lc_I^2H^2\|\nabla e_h\|_0^2 \ge \frac{\nu}{2}\|\nabla e_h\|_0^2 \tag{3.20}$$

whenever

$$H \le \frac{\nu^{1/2}}{(8L)^{1/2}c_I}. \tag{3.21}$$

Therefore, for the last three terms on the left-hand side of (3.19) we have

$$\nu\|\nabla \boldsymbol{e}_h\|_0^2 + \frac{\beta}{2}\|I_H\boldsymbol{e}_h\|_0^2 - 4L\|(I - I_H)\boldsymbol{e}_h\|_0^2 \ge \frac{\nu}{2}\|\nabla \boldsymbol{e}_h\|_0^2 + \frac{\beta}{2}\|I_H\boldsymbol{e}_h\|_0^2. \tag{3.22}$$

Now, applying (2.20) again to bound below the right-hand side above we have that

$$\begin{aligned} \frac{\nu}{2}\|\nabla \boldsymbol{e}_h\|_0^2 + \frac{\beta}{2}\|I_H\boldsymbol{e}_h\|_0^2 &\ge \frac{\nu}{2}c_I^{-2}H^{-2}\|(I - I_H)e_h\|_0^2 + \frac{\beta}{2}\|I_H\boldsymbol{e}_h\|_0^2 \\ &\ge \gamma(\|I_H\boldsymbol{e}_h\|_0^2 + \|(I - I_H)\boldsymbol{e}_h\|_0^2), \end{aligned} \tag{3.23}$$

where

$$\gamma = \min\left\{\frac{\nu}{2}c_I^{-2}H^{-2}, \frac{\beta}{2}\right\}. \tag{3.24}$$

Finally, since $\gamma(\|I_H\boldsymbol{e}_h\|_0^2+\|(I-I_H)\boldsymbol{e}_h\|_0^2) \geq (\gamma/2)\|e_h\|_0^2$, from (3.19), (3.22), and (3.23) it follows that

$$\frac{d}{dt}\|\boldsymbol{e}_h\|_0^2 + \frac{\gamma}{2}\|\boldsymbol{e}_h\|_0^2 \leq 2\overline{k}\|\tau_h^2\|_0^2 + \beta c_0^2\|\boldsymbol{u}-\boldsymbol{w}_h\|_0^2 + \left((1-\overline{\mu})\frac{2\hat{c}_P}{\nu} + \frac{\overline{\mu}}{L}\right)\|\boldsymbol{\tau}_h^1\|_{-1+\overline{\mu}}^2,$$

from which we reach (3.6). $\square$

We now obtain the error bounds of the standard Galerkin method (case $\mu = 0$).

THEOREM 3.2. *Assume that the solution of* (1.1) *satisfies that* $\boldsymbol{u} \in L^\infty(H^s(\Omega)^d)$, $p \in L^\infty(H^{s-1}(\Omega)/\mathbb{R})$, $\boldsymbol{u}_t \in L^\infty(H^{\max(2,s-1)}(\Omega)^d)$, *and* $p_t \in L^\infty(H^{\max(1,s-2)}(\Omega)/\mathbb{R})$ *for* $s \geq 2$. *Let* $\boldsymbol{u}_h$ *be the finite element approximation defined in* (3.3) *with* $\mu = 0$. *Then, if* $\beta \geq 8L$ *and* H *satisfies condition* (3.21), *the following bound holds for* $t \geq 0$ *and* $2 \leq r \leq s$*:*

$$\begin{aligned}\|\boldsymbol{u}(t)-\boldsymbol{u}_h(t)\|_0 \leq & e^{-\gamma t/2}\|\boldsymbol{u}_h(0)-\boldsymbol{u}(0)\|_0^2 \\ & + C\Big(\max_{0\leq\tau\leq t}\big((\beta/\gamma)^{1/2} + (\gamma\nu)^{-1/2}K_0(\boldsymbol{u},p,|\Omega|)\big)N_r(\boldsymbol{u},p) \\ & + (\gamma\nu)^{-1/2}\max_{0\leq\tau\leq t}|\Omega|^{(1+\hat{r}-r)/d}N_{\hat{r}}(\boldsymbol{u}_t,p_t)\Big)h^r,\end{aligned}$$

where γ *is defined in* (3.24), $K_0(\boldsymbol{u},p,|\Omega|)$ *is defined in* (3.28) *below, and* $\hat{r} = r-1$ *if* $r \geq 3$ *and* Ω *is of class* $\mathcal{C}^3$ *and* $\hat{r} = r$ *otherwise.*

Proof. Following [6] we compare $\boldsymbol{u}_h$ with $\boldsymbol{s}_h$, where $\boldsymbol{s}_h$ satisfies (2.8), for which we apply Lemma 3.1 with $\boldsymbol{w}_h = \boldsymbol{s}_h$. To bound the norms $\|\boldsymbol{s}_h\|_\infty$ and $\|\nabla\boldsymbol{s}_h\|_{L^{2d/(d-1)}}$ in (3.14) we apply (3.30) and (3.31).

We observe that (3.5) holds with $\mu = 0$ and $\boldsymbol{\tau}_h^2 = 0$ and

$$(\boldsymbol{\tau}_h^1, \boldsymbol{\varphi}_h) = (\boldsymbol{u}_t - \dot{\boldsymbol{s}}_h, \boldsymbol{\varphi}_h) + b_h(\boldsymbol{u},\boldsymbol{u},\boldsymbol{\varphi}_h) - b_h(\boldsymbol{s}_h,\boldsymbol{s}_h,\boldsymbol{\varphi}_h) \quad \forall \boldsymbol{\varphi}_h \in V_{h,r}.$$

Then from (3.6) we get

$$\begin{aligned}\|\boldsymbol{e}_h(t)\|_0^2 \leq e^{-\gamma t/2}\|\boldsymbol{e}_h(0)\|_0^2 + \int_0^t e^{-\gamma(t-s)/2}\frac{2\hat{c}_P}{\nu}\|\boldsymbol{\tau}_h^1\|_{-1}^2\, ds \\ + \int_0^t e^{-\gamma(t-s)/2}\beta c_0^2\|\boldsymbol{u}(s)-\boldsymbol{w}_h(s)\|_0^2\, ds.\end{aligned}$$

Consequently,

$$\|\boldsymbol{e}_h(t)\|_0^2 \leq e^{-\gamma t/2}\|\boldsymbol{e}_h(0)\|_0^2 + \frac{4\hat{c}_P}{\nu\gamma}\max_{0\leq\tau\leq t}\|\tau_h(\tau)\|_{-1}^2 + 2c_0^2\frac{\beta}{\gamma}\max_{0\leq\tau\leq t}\|\boldsymbol{u}(\tau)-\boldsymbol{s}_h(\tau)\|_0^2.$$

To bound the last term on the right-hand side of above we apply (2.9) to get

$$\max_{0\leq\tau\leq t}\|\boldsymbol{u}(\tau)-\boldsymbol{s}_h(\tau)\|_0^2 \leq Ch^{2r}\max_{0\leq\tau\leq t}N_r(\boldsymbol{u}(\tau),p(\tau)).$$

For the truncation error, applying (2.12) we can bound

$$\|\boldsymbol{u}_t - \dot{\boldsymbol{s}}_h\|_{-1} \leq Ch^rN_{r-1}(\boldsymbol{u}_t,p_t),$$

or, in case we use the minielement or the boundary is not of class $\mathcal{C}^3$, applying (2.9) again we get

$$\|\boldsymbol{u}_t - \dot{\boldsymbol{s}}_h\|_{-1} \leq C|\Omega|^{1/d}h^rN_r(\boldsymbol{u}_t,p_t).$$

Also, applying Lemma 3.6 below we have

$$\sup_{\|\boldsymbol{\varphi}\|_1=1} |b_h(\boldsymbol{u},\boldsymbol{u},\boldsymbol{\varphi}) - b_h(\boldsymbol{s}_h,\boldsymbol{s}_h,\boldsymbol{\varphi})| \le K_0(\boldsymbol{u},p,|\Omega|)\|\boldsymbol{u}-\boldsymbol{s}_h\|_0,$$

so that we conclude the proof by applying again (2.9). ∎

We observe from Theorem 3.2 that the rate of convergence of the method is optimal $O(h^s)$ and, as in [37], we have obtained uniform in time error estimates. In the following theorem we bound the error of the Galerkin method with grad-div stabilization (case $\mu > 0$). Comparing with Theorem 3.2 we show that adding grad-div stabilization allows us to remove the dependence of the error constants on inverse powers of the viscosity ν.

THEOREM 3.3. *Assume that the solution of* (1.1) *satisfies that* $\boldsymbol{u} \in L^\infty(H^s(\Omega)^d) \cap W^{1,\infty}(\Omega)^d$, $p \in L^\infty(H^{s-1}(\Omega)/\mathbb{R})$, *and* $\boldsymbol{u}_t \in L^\infty(H^{s-1}(\Omega)^d)$ *for* $s \ge 2$. *Let* $\boldsymbol{u}_h$ *be the finite element approximation defined in* (3.3) *with grad-div stabilization* ($\mu \neq 0$). *Then, if* $\beta \ge 8L$ *and* H *satisfies condition* (3.21), *the following bound holds for* $t \ge 0$ *and* $2 \le r \le s$:

$$\begin{aligned}\|\boldsymbol{u}(t)-\boldsymbol{u}_h(t)\|_0 \le\,& e^{-\gamma t/2}\|\boldsymbol{u}_h(0)-\boldsymbol{u}(0)\|_0^2\\ &+\frac{C}{L^{1/2}}h^{r-1}\max_{0\le\tau\le t}\biggl(\Bigl(\beta^{1/2}h+\mu^{1/2}+\frac{K_1(\boldsymbol{u},|\Omega|)}{L^{1/2}}\Bigr)\|\boldsymbol{u}\|_r\\ &+\frac{1}{L^{1/2}}\|\boldsymbol{u}_t\|_{r-1}+\frac{1}{\mu^{1/2}}\|p\|_{H^{r-1}/\mathbb{R}})\biggr),\end{aligned}$$

where γ *is defined in* (3.24) *and* $K_1(\boldsymbol{u},|\Omega|)$ *is defined in* (3.29) *below.*

Proof. Following [15], [16] we compare $\boldsymbol{u}_h$ with $\boldsymbol{s}_h^m$, where $\boldsymbol{s}_h^m$ satisfies (2.13). We first observe that the norms in (3.15) are bounded since for $\|\boldsymbol{s}_h^m\|_\infty$ we apply (3.32) and applying (2.15), $\|\nabla \boldsymbol{s}_h^m\|_\infty \le C\|\nabla\boldsymbol{u}\|_\infty$.

Then, we apply Lemma 3.1 with $\boldsymbol{w}_h = \boldsymbol{s}_h^m$. We observe that (3.5) holds with

$$(\boldsymbol{\tau}_h^1,\boldsymbol{\varphi}_h) = (\dot{\boldsymbol{u}}-\dot{\boldsymbol{s}}_h^m,\boldsymbol{\varphi}_h) + b_h(\boldsymbol{u},\boldsymbol{u},\boldsymbol{\varphi}_h) - b_h(\boldsymbol{s}_h^m,\boldsymbol{s}_h^m,\boldsymbol{\varphi}_h) \quad \forall \boldsymbol{\varphi}_h \in V_{h,r}$$

and

$$(\boldsymbol{\tau}_h^2,\nabla\cdot\boldsymbol{\varphi}_h) = (\pi_h p - p,\nabla\cdot\boldsymbol{\varphi}_h) + \mu(\nabla\cdot(\boldsymbol{u}-\boldsymbol{s}_h^m),\nabla\cdot\boldsymbol{\varphi}_h),$$

and then from (3.6), we get

$$\begin{aligned}\|\boldsymbol{e}_h(t)\|_0^2 \le e^{-\gamma t/2}\|\boldsymbol{e}_h(0)\|_0^2 &+ \int_0^t e^{-\gamma(t-s)/2}\frac{\|\boldsymbol{\tau}_h^1\|_0^2}{L}\,ds\\ &+\int_0^t e^{-\gamma(t-s)/2}\left(\beta c_0^2\|\boldsymbol{u}(s)-\boldsymbol{s}_h^m(s)\|_0^2+\frac{2}{\mu}\|\tau_h^2\|_0^2\right)\,ds.\end{aligned}$$

Consequently,

$$\begin{aligned}\|\boldsymbol{e}_h(t)\|_0^2 \le e^{-\gamma t/2}\|\boldsymbol{e}_h(0)\|_0^2 + \frac{2}{\gamma L}\max_{0\le\tau\le t}\|\tau_h^1(\tau)\|_0^2 + 2c_0^2\frac{\beta}{\gamma}\max_{0\le\tau\le t}\|\boldsymbol{u}(\tau)-\boldsymbol{s}_h^m(\tau)\|_0^2\\ +\frac{4}{\mu\gamma}\max_{0\le\tau\le t}\|\boldsymbol{\tau}_h^2(\tau)\|_0^2.\end{aligned}$$

From (3.21) and (3.24) and taking into account that we are assuming $\beta \ge 8L$, we get $1/\gamma \le \max(2/\beta, 1/(4L)) = 1/(4L)$ and $\beta/\gamma \le \max(2,\beta/(4L)) = \beta/(4L)$. Then, it follows that

$$\begin{aligned}\|\boldsymbol{e}_h(t)\|_0^2 \le e^{-\gamma t/2}\|\boldsymbol{e}_h(0)\|_0^2 + \frac{1}{2L^2}\max_{0\le\tau\le t}\|\tau_h^1(\tau)\|_0^2 + \frac{\beta}{2L}c_0^2\max_{0\le\tau\le t}\|\boldsymbol{u}(\tau)-\boldsymbol{s}_h^m(\tau)\|_0^2 \\ + \frac{1}{\mu L}\max_{0\le\tau\le t}\|\boldsymbol{\tau}_h^2(\tau)\|_0^2. \qquad (3.25)\end{aligned}$$

To bound the second term on the right-hand side of (3.25) we apply (2.14) to get

$$\max_{0\le\tau\le t}\|\boldsymbol{u}(\tau)-\boldsymbol{s}_h(\tau)\|_0^2 \le Ch^{2r}\max_{0\le\tau\le t}\|\boldsymbol{u}(\tau)\|_r^2.$$

For the first term in the truncation error $\boldsymbol{\tau}_h^1$ we apply (2.14) again to get

$$\max_{0\le\tau\le t}\|\boldsymbol{u}_t(\tau)-\dot{\boldsymbol{s}}_h(\tau)\|_0^2 \le Ch^{2(r-1)}\max_{0\le\tau\le t}\|\boldsymbol{u}_t(\tau)\|_{r-1}^2.$$

For the second term in the truncation error $\boldsymbol{\tau}_h^1$, applying Lemma 3.6 below we have

$$\sup_{\|\boldsymbol{\varphi}\|_0=1}|b_h(\boldsymbol{u},\boldsymbol{u},\boldsymbol{\varphi})-b_h(\boldsymbol{s}_h^m,\boldsymbol{s}_h^m,\boldsymbol{\varphi})| \le K_1(\boldsymbol{u},|\Omega|)\|\boldsymbol{u}-\boldsymbol{s}_h^m\|_1 \le CK_1(\boldsymbol{u},|\Omega|)h^{r-1}\|\boldsymbol{u}\|_r,$$

where in the last inequality we have applied (2.14). Finally, from (2.16) and (2.14) we obtain

$$\|\boldsymbol{\tau}_h^2\|_0 \le Ch^{r-1}\|p\|_{H^{r-1}/\mathbb{R}} + C\mu h^{r-1}\|\boldsymbol{u}\|_r,$$

which concludes the proof. ∎

Remark 3.4. Some works in the literature, [34], [37], use $\beta(I_H(\boldsymbol{u}_h-\boldsymbol{u}),\boldsymbol{\varphi}_h)$ as a nudging term instead of the one in (3.1), which is also used in [40]. Using (3.1) in the present paper is essential for the error analysis of the method. In the case where I_H is the orthogonal projection in L^2, since $\beta(I_H(\boldsymbol{u}_h-\boldsymbol{u}),\boldsymbol{\varphi}_h)=\beta(I_H(\boldsymbol{u}_h-\boldsymbol{u}),I_H\boldsymbol{\varphi}_h)$ the analysis presented above obviously covers both nudging terms.

Remark 3.5. By adding $+\mu(\nabla\cdot\boldsymbol{s}_h,\nabla\cdot\boldsymbol{\varphi})$ to the left-hand side of the first equation in (2.8), and repeating the arguments in the proof of Theorem 3.2 (with obvious changes), one can obtain an $O(h^s)$ error bound also when $\mu>0$, but where, as in Theorem 3.2 and opposed to Theorem 3.3, error constants depend on inverse powers of ν and, hence, are useful in practice only when ν is not too small (see Figure 2 below).

Lemma 3.6. *The following bounds hold:*

$$\sup_{\|\boldsymbol{\varphi}\|_1=1}|b_h(\boldsymbol{u},\boldsymbol{u},\boldsymbol{\varphi})-b_h(\boldsymbol{s}_h,\boldsymbol{s}_h,\boldsymbol{\varphi})| \le K_0(\boldsymbol{u},p,|\Omega|)\|\boldsymbol{u}-\boldsymbol{s}_h\|_0, \qquad (3.26)$$

$$\sup_{\|\boldsymbol{\varphi}\|_0=0}|b_h(\boldsymbol{u},\boldsymbol{u},\boldsymbol{\varphi})-b_h(\boldsymbol{s}_h^m,\boldsymbol{s}_h^m,\boldsymbol{\varphi})| \le K_1(\boldsymbol{u},|\Omega|)\|\boldsymbol{u}-\boldsymbol{s}_h^m\|_1, \qquad (3.27)$$

where

(3.28)

$$K_0(\boldsymbol{u},p,|\Omega|) = C\Big(K_1(\boldsymbol{u},|\Omega|) + N_1(\boldsymbol{u},p)^{1/2}\big(N_{d-1}(\boldsymbol{u},p)+|\Omega|^{(3-d)/d}N_2(\boldsymbol{u},p)\big)^{1/2}\Big),$$

$$K_1(\boldsymbol{u},|\Omega|) = C\Big((\|\boldsymbol{u}\|_{d-2}\|\boldsymbol{u}\|_2)^{1/2} + |\Omega|^{(3-d)/(2d)}(\|\boldsymbol{u}\|_1\|\boldsymbol{u}\|_2)^{1/2}\Big), \qquad (3.29)$$

and $N_j(\boldsymbol{u},p)$ is the quantity in (2.10).

Proof. Applying [17, Lemma 5] we have

$$|b_h(\boldsymbol{u},\boldsymbol{u},\boldsymbol{\varphi})-b_h(\boldsymbol{s}_h,\boldsymbol{s}_h,\boldsymbol{\varphi})| \leq C\big(\|\nabla\boldsymbol{u}\|_{L^{2d/(d-1)}}+\|\nabla\boldsymbol{s}_h\|_{L^{2d/(d-1)}}\big)\|\boldsymbol{u}-\boldsymbol{s}_h\|_0\|\boldsymbol{\varphi}\|_{L^{2d}} + \big(\|\boldsymbol{u}\|_\infty+\|\boldsymbol{s}_h\|_\infty\big)\|\boldsymbol{u}-\boldsymbol{s}_h\|_0\|\nabla\boldsymbol{\varphi}\|_0.$$

To bound $\|\nabla\boldsymbol{u}\|_{L^{2d/(d-1)}}$ and $\|\boldsymbol{u}\|_\infty$ we apply (2.2) and (2.3), respectively, and applying Sobolev's inequality (2.1) we have $\|\boldsymbol{\varphi}\|_{L^{2d}} \leq c_1|\Omega|^{(3-d)/(2d)}\|\boldsymbol{\varphi}\|_1$. The proof of (3.26) is finished by applying Lemma 3.7 below.

To prove (3.27), we replace $\boldsymbol{s}_h$ by $\boldsymbol{s}_h^m$ in the arguments above and use the skew-symmetric property of b to interchange the roles of φ and $\boldsymbol{u}-\boldsymbol{s}_h$. We finish by applying Lemma 3.8 below. ∎

LEMMA 3.7. *There exist a positive constant C_0 such that the following bounds hold:*

$$\|\boldsymbol{s}_h\|_\infty \leq C_0\left((\|\boldsymbol{u}\|_{d-2}\|\boldsymbol{u}\|_2)^{1/2}+\big(N_1(\boldsymbol{u},p)N_{d-1}(\boldsymbol{u},p)\big)^{1/2}\right), \tag{3.30}$$

$$\|\nabla\boldsymbol{s}_h\|_{L^{2d/(d-1)}} \leq C_0\big(N_1(\boldsymbol{u},p)N_2(\boldsymbol{u},p)\big)^{1/2}. \tag{3.31}$$

Proof. For the L^∞ bound, applying inverse inequality (2.17), we write

$$\|\boldsymbol{s}_h\|_\infty \leq C\|I_h(\boldsymbol{u})\|_\infty+\|\boldsymbol{s}_h-I_h(\boldsymbol{u})\|_\infty \leq C\|\boldsymbol{u}\|_\infty+c_{\rm inv}h^{-d/2}\big(\|\boldsymbol{s}_h-\boldsymbol{u}\|_0+\|\boldsymbol{u}-I_h(\boldsymbol{u})\|_0\big)$$

and apply (2.3) to bound $\|\boldsymbol{u}\|_\infty$. In the case $d=2$ we have

$$\|\boldsymbol{u}-I_h(\boldsymbol{u})\|_0 \leq Ch\|\boldsymbol{u}\|_1 \leq Ch(\|\boldsymbol{u}\|_0\|\boldsymbol{u}\|_2)^{1/2} = Ch^{d/2}(\|\boldsymbol{u}\|_{d-2}\|\boldsymbol{u}\|_2)^{1/2},$$

where we have applied (2.20) for $H=h$ and also $\|\boldsymbol{u}-I_h(\boldsymbol{u})\|_0 \leq Ch^2\|\boldsymbol{u}\|_2$. By (2.9),

$$\|\boldsymbol{s}_h-\boldsymbol{u}\|_0 \leq CN_1(\boldsymbol{u},p)h.$$

In the case $d=3$,

$$\|\boldsymbol{u}-I_h(\boldsymbol{u})\|_0 \leq Ch^{3/2}(\|\boldsymbol{u}\|_1\|\boldsymbol{u}\|_2)^{1/2} = Ch^{d/2}(\|\boldsymbol{u}\|_{d-2}\|\boldsymbol{u}\|_2)^{1/2}$$

and

$$\|\boldsymbol{s}_h-\boldsymbol{u}\|_0 \leq Ch^{d/2}(N_1(\boldsymbol{u},p)N_2(\boldsymbol{u},p))^{1/2}.$$

For $\|\nabla\boldsymbol{s}_h\|_{L^{2d/(d-1)}}$, since $\|\nabla\boldsymbol{s}_h\|_{L^q} \leq C(\|\nabla\boldsymbol{u}\|_{L^q}+\nu^{-1}\|p\|_{L^q})$ for $q=2,\infty$ [12], by the Riesz–Thorin interpolation theorem and applying (2.2),

$$\begin{aligned}\|\nabla\boldsymbol{s}_h\|_{L^{2d/(d-1)}} &\leq C(\|\nabla\boldsymbol{u}\|_{L^{2d/(d-1)}}+\nu^{-1}\|p\|_{L^{2d/(d-1)}}\\ &\leq C((\|\boldsymbol{u}\|_1\|\boldsymbol{u}\|_2)^{1/2}+\nu^{-1}(\|p\|_0\|p\|_1)^{1/2}) \leq C(N_1(\boldsymbol{u},p)N_2(\boldsymbol{u},p))^{1/2}.\end{aligned}$$ ∎

LEMMA 3.8. *There exists a positive constant C_1 such that the following bounds hold:*

$$\|\boldsymbol{s}_h^m\|_\infty \leq C_1(\|\boldsymbol{u}\|_{d-2}\|\boldsymbol{u}\|_2)^{1/2}, \tag{3.32}$$

$$\|\nabla\boldsymbol{s}_h^m\|_{L^{2d/(d-1)}} \leq C_1\big(\|\boldsymbol{u}\|_1\|\boldsymbol{u}\|_2\big)^{1/2}. \tag{3.33}$$

Proof. We argue exactly as in the proof of Lemma 3.7, replacing (2.9) by (2.14). ∎

Remark 3.9. In the case $\mu = 0$, according to Lemma 3.7, we have that $\beta \geq 8L$ when $\boldsymbol{w}_h = \boldsymbol{s}_h$ if for $t \geq 0$,

$$\beta \geq 8 \left(2 \frac{\left((\hat{c}_1 C_0)^2 N_1(\boldsymbol{u},p) N_2(\boldsymbol{u},p)\right)^{d/3}}{\nu^{(2d-3)/3}} + C_0^2 \frac{\|\boldsymbol{u}\|_{d-2}\|\boldsymbol{u}\|_2 + N_1(\boldsymbol{u},p)N_{d-1}(\boldsymbol{u},p)}{\nu} \right) \tag{3.34}$$

with C_0 the constant in Lemma 3.7. In case $\mu \neq 0$ from (2.15) and (3.32) we have that $\beta \geq 8L$ when $\boldsymbol{w}_h = \boldsymbol{s}_h^m$ if for $t \geq 0$

$$\beta \geq 16 \left(C_1 \|\nabla \boldsymbol{u}\|_\infty + C_1^2 \frac{\|\boldsymbol{u}\|_{d-2}\|\boldsymbol{u}\|_2}{4\mu} \right). \tag{3.35}$$

3.1. The Lagrange interpolant. In this section we consider when $I_H \boldsymbol{u} = I_H^{La} \boldsymbol{u}$. With the help of the following lemmas we will show that the analogue of Theorems 3.2 and 3.3 (Theorem 3.12 below) also holds in this case.

LEMMA 3.10. *For $v_h \in X_{h,r}$ the following holds:*

$$\|\boldsymbol{v}_h - I_H^{La}\boldsymbol{v}_h\|_0 \leq c_{La} H \|\nabla \boldsymbol{v}_h\|_0, \tag{3.36}$$

where

$$c_{La} = C\,(H/h)^{\frac{d(p-2)}{2p}}, \tag{3.37}$$

where C is a generic constant and $p = 3$ if $d = 2$ and $p = 4$ if $d = 3$.

Proof. For $\boldsymbol{v}_h \in X_{h,r}$ we write

$$\|\boldsymbol{v}_h - I_H^{La}\boldsymbol{v}_h\|_0^2 = \sum_{K \in T_H} \|\boldsymbol{v}_h - I_H^{La}\boldsymbol{v}_h\|_{L^2(K)}^2 \leq C \sum_{K \in T_H} H^{\frac{d(p-2)}{p}} \|\boldsymbol{v}_h - I_H^{La}\boldsymbol{v}_h\|_{L^p(K)}^2, \tag{3.38}$$

the last inequality being a consequence of Hölder's inequality and of the fact that $|K| \leq CH^d$. Applying (2.18) and (2.17) we get

$$\|\boldsymbol{v}_h - I_H^{La}\boldsymbol{v}_h\|_{L^p(K)} \leq c_{\rm int} H \|\nabla \boldsymbol{v}_h\|_{L^p(K)} \leq c_{\rm int} c_{\rm inv} H \|\nabla \boldsymbol{v}_h\|_{L^2(K)} h^{-\frac{d(p-2)}{2p}},$$

so that inserting the above inequality into (3.38) we reach (3.36). ∎

LEMMA 3.11. *Let $\boldsymbol{s}_h$ be the Stokes projection defined in* (2.8)*. Then the following bound holds:*

$$\|(I - I_H^{La})(\boldsymbol{s}_h - \boldsymbol{u})\|_0 \leq CH^2 h^{r-2} \|\boldsymbol{u}\|_r, \tag{3.39}$$

where C is a generic constant.

Proof. We write

$$(I - I_H^{La})(\boldsymbol{s}_h - \boldsymbol{u}) = (I - I_H^{La})(\boldsymbol{s}_h - I_h^{La}\boldsymbol{u}) + (I - I_H^{La})(I_h^{La}\boldsymbol{u} - \boldsymbol{u}).$$

Applying (3.36) and (3.37) to $\boldsymbol{v}_h = \boldsymbol{s}_h - I_h^{La}\boldsymbol{u}$ and then (2.9) and (2.18) we get

$$\begin{aligned} \|(I - I_H^{La})(\boldsymbol{s}_h - I_h^{La}\boldsymbol{u})\|_0 &\leq C\,(H/h)^{\frac{d(p-2)}{2p}}\, H \|\nabla(\boldsymbol{s}_h - I_h^{La}\boldsymbol{u})\|_0 \\ &\leq C\,(H/h)^{\frac{d(p-2)}{2p}}\, H h^{r-1} \|\boldsymbol{u}\|_r \leq CH^2 h^{r-2}\|\boldsymbol{u}\|_r, \end{aligned} \tag{3.40}$$

where in the last inequality we have bounded $(H/h)^{d(p-2)/(2p)}$ by H/h. For the other term we argue as in (3.38) and apply (2.18) to get

$$\begin{aligned}\|(I-I_H^{La})(I_h^{La}\boldsymbol{u}-\boldsymbol{u})\|_0^2 &\le Cc_{\rm int}H^2H^{\frac{d(p-2)}{p}}\sum_{K\in\tau_H}\|\nabla(I_h^{La}\boldsymbol{u}-\boldsymbol{u})\|_{L^p(K)}^2\\ &\le Cc_{\rm int}H^2H^{\frac{d(p-2)}{p}}h^{2(r-2)}\sum_{K\in\tau_H}|\boldsymbol{u}|_{r-1,p,K}^2. \qquad (3.41)\end{aligned}$$

Applying (2.1) with $s=1$ and taking into account $CH^d\le|K|\le CH^d$ we get $\|\boldsymbol{u}\|_{L^p(K)}\le CH^{1-\frac{d(p-2)}{2p}}\|\boldsymbol{u}\|_{1,K}$, from which

$$|\boldsymbol{u}|_{r-1,p,K}^2\le CH^{2-\frac{d(p-2)}{p}}\|\boldsymbol{u}\|_{r,2,K}^2.$$

Inserting the above inequality into (3.41) we reach

$$\|(I-I_H^{La})(I_h^{La}\boldsymbol{u}-\boldsymbol{u})\|_0\le CH^2h^{r-2}\|\boldsymbol{u}\|_r. \qquad (3.42)$$

Finally, (3.39) follows from (3.40) and (3.42). ∎

THEOREM 3.12. *In the same conditions as Theorem* 3.2 *(resp., Theorem* 3.3*), if* I_H *is replaced by* I_H^{La}, H *satisfies condition* (3.21) *with* c_I *replaced by* c_{La} *defined in* (3.37)*, and* H/h *remains bounded, then the statement of Theorem* 3.2 *(resp., Theorem* 3.3*) holds with* γ *defined in* (3.24) *with* c_I *replaced by* c_{La}.

Proof. The proof of the theorem can be obtained arguing exactly as in the proof of Theorem 3.2 (resp., 3.3) with only two differences that we now state. We first observe that assuming H/h remains bounded we can apply (3.36) instead of (2.20) in (3.20) and (3.23). We also observe that since (2.19) does not hold for $I_H=I_H^{la}$ we cannot apply (3.17). Instead, adding and subtracting $\boldsymbol{u}-\boldsymbol{s}_h$ and using (3.39) we get

$$\begin{aligned}&\beta|(I_H\boldsymbol{u}-I_H\boldsymbol{s}_h,I_H\boldsymbol{e}_h)|\le\beta|(I_H-I)(\boldsymbol{u}-\boldsymbol{s}_h),I_H\boldsymbol{e}_h)|+\beta|(\boldsymbol{u}-\boldsymbol{s}_h,I_H\boldsymbol{e}_h)|\\&\quad\le\beta(CH^2h^{r-2}\|\boldsymbol{u}\|_r+\|\boldsymbol{u}-\boldsymbol{s}_h\|_0)\|I_H\boldsymbol{e}_h\|_0\le\beta(Ch^r\|\boldsymbol{u}\|_r+\|\boldsymbol{u}-\boldsymbol{s}_h\|_0)\|I_H\boldsymbol{e}_h\|_0,\end{aligned}$$

where in the last inequality we have applied that since H/h is bounded then $H\le Ch$. Then we replace (3.17) in the proof of Lemma 3.1 and consequently in the proof of Theorem 3.2 (resp., 3.3) by the inequality

$$\beta|(I_H\boldsymbol{u}-I_H\boldsymbol{s}_h,I_H\boldsymbol{e}_h)|\le\frac{\beta}{2}(Ch^r\|\boldsymbol{u}\|_r+\|\boldsymbol{u}-\boldsymbol{s}_h\|_0)^2+\frac{\beta}{2}\|I_H\boldsymbol{e}_h\|_0^2,$$

and we can conclude applying the same arguments. ∎

4. Numerical experiments. We check the results of the previous section with some numerical experiments. As is customary for these purposes, we use an example with a known solution. In particular, we consider the Navier–Stokes equations in $\Omega=[0,1]^2$, with the forcing term $\boldsymbol{f}$ chosen so that the solutions $\boldsymbol{u}$ and p are given by

$$\boldsymbol{u}(x,y,t)=\frac{6+4\cos(4t)}{10}\left[\begin{array}{c}8\sin^2(\pi x)(2y(1-y)(1-2y)\\-8\pi\sin(2*\pi x)(y(1-y))^2\end{array}\right], \qquad (4.1)$$

$$p(x,y,t)=\frac{6+4\cos(4t)}{10}\sin(\pi x)\cos(\pi y), \qquad (4.2)$$

which means that $\boldsymbol{f}=\partial_t\boldsymbol{u}-\nu\Delta\boldsymbol{u}+(\boldsymbol{u}\cdot\nabla)\boldsymbol{u}+\nabla p$ with the above expressions of $\boldsymbol{u}$ and p and a value of ν that will be specified for every experiment. For the spatial discretization we used P_2/P_1 elements on a regular triangulation with SW-NE diag-

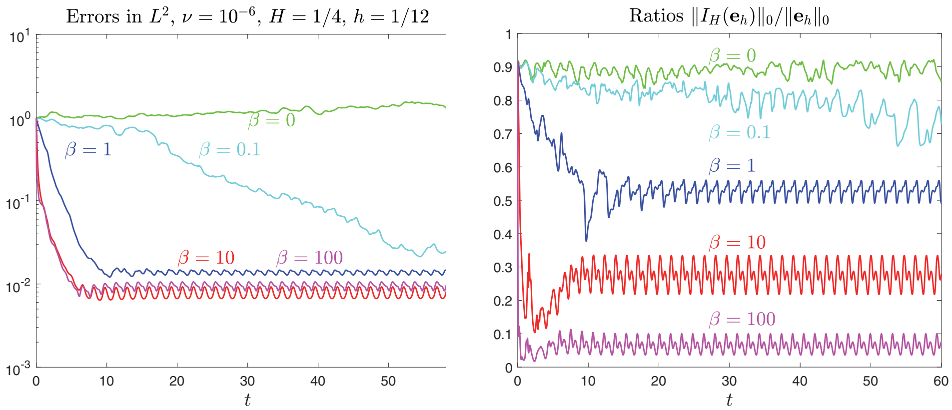

FIG. 1. *Velocity errors vs. time.*

onals, with the same number of subdivisions on each coordinate direction. For coarse mesh interpolation we take piecewise constants. The time integration was done with an IMEX method based on the second order backward differentiation formula, where, to avoid solving nonlinear steady problems at each step, linear extrapolation of the form $b_h(2\boldsymbol{u}_h(t-\Delta t)-\boldsymbol{u}_h(t-2\Delta t), \boldsymbol{u}_h(t), \boldsymbol{\varphi}_h)$ was used in the convection term, where Δt is the time step, except in the first step, where $b_h(\boldsymbol{u}_h(t-\Delta t), \boldsymbol{u}_h(t), \boldsymbol{\varphi}_h)$ was used. The time step was chosen so that the error arising from the spatial discretization was dominant. To check that this was the case, we made sure that results were not essentially altered if recomputed with a smaller Δt. For $h = 1/12, 1/18, 1/24, 1/36$, and $1/48$, the values of Δt used were, respectively, $\Delta t = 1/160, 1/160, 1/320, 1/320$, and $1/640$. Unless stated otherwise, in what follows the initial condition was set to $\boldsymbol{u}_h = \mathbf{0}$ and $p = 0$, so that there is an $O(1)$ error at time $t = 0$.

We first check that there is no upper bound on the nudging parameter β. The left plot in Figure 1 shows the velocity errors in L^2 vs. time for different values of β for $\nu = 10^{-6}$, including $\beta = 100$. A clear difference can be seen between $\beta = 0$, where the initial errors do not decay with time, and $\beta > 0$ where they do, and for the four largest values of β shown, they do so exponentially in time, until an asymptotic regime is reached. We also notice that the results are little altered for $\beta \geq 10$. In view of (3.24), one may be temped to question the advantage of taking $\beta \geq \nu(c_I H)^{-2}$, since the rate of decay of the initial errors, γ, is unaltered for larger values of β. If we assume that c_I is of order one, then the value of $\nu(c_I H)^{-2}$ in the present example is unlikely to be larger than 10^{-5}, so that Figure 1 (and more examples in [34] and [40]) seems to suggest that there is some advantage in taking $\beta \geq \nu(c_I H)^{-2}$ if we want a faster decay of the initial errors. Since this is in apparent contradiction with the analysis in the previous section, we now propose an alternative explanation.

Let us consider for some integer $k \geq 2$ the value $r = \|I_H(\boldsymbol{u}(0))\|_0/(k\|\boldsymbol{u}(0)\|_0)$. If we take the intial condtion $\boldsymbol{u}_h = 0$, then by continuity there exist $t_0 > 0$ such that

$$\|I_H(\boldsymbol{e}_h(t)\|_0 \geq r\|\boldsymbol{e}_h(t)\|_0, \qquad t \in [0, t_0]. \tag{4.3}$$

Consequently, from (3.18) it follows that

$$\frac{d}{dt}\|\boldsymbol{e}_h\|_0^2 + (\beta r^2 - 2L)\|\boldsymbol{e}_h\|_0^2 \leq 2\overline{k}\|\tau_h^2\|_0^2 + \beta c_0^2\|\boldsymbol{u}-\boldsymbol{w}_h\|_0^2 + \left((1-\overline{\mu})\frac{2\hat{c}_P}{\nu} + \overline{\mu}L\right)\|\boldsymbol{\tau}_h^1\|_{-1+\overline{\mu}}^2$$

for $t \in [0, t_0]$, which, for $\beta > 2L/r^2$, could explain that initial errors decay faster when

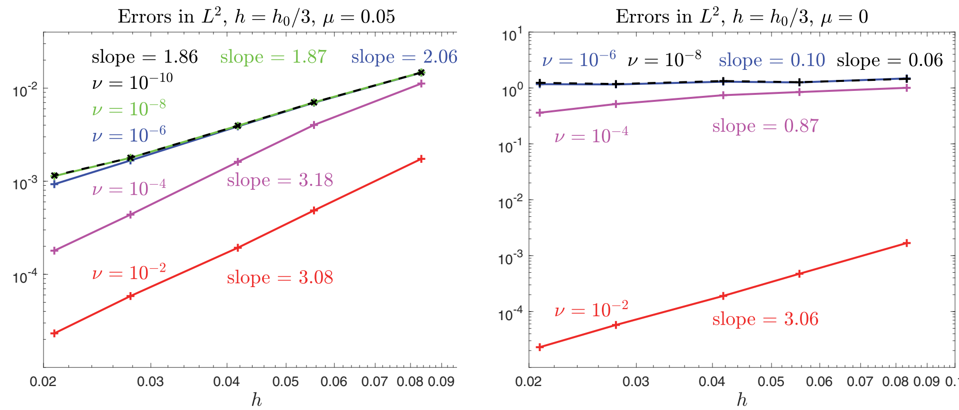

FIG. 2. *Velocity errors. Left,* $\mu = 0.05$*. Right,* $\mu = 0$.

larger values of β are taken. In Figure 1 we also show the ratios $\|I_H(\boldsymbol{e}_h)\|_0/\|\boldsymbol{e}_h\|_0$. It can be seen that although they became smaller as β is increased, they are sufficiently away from zero to suggest that the analysis in the present section may explain the faster rates of decay of the initial errors when larger values of β are taken.

In the remainder of this section we take $\beta = 1$. Since, as shown in Figure 1, after an initial decay, the errors show an oscillatory behavior, in the examples below by L^2 errors we mean the maximum of errors $\|\boldsymbol{u}_h(t) - \boldsymbol{u}(t)\|_0$ for values of t after the asymptotic regime has shown itself.

We now check the rates of convergence proved in the present paper. In Figure 2 we show errors vs. h for different values of the diffusion parameter ν and compare the cases of positive μ ($\mu = 0.05$) and $\mu = 0$. The value of H is $H = 3h$ and β is set to $\beta = 1$. Results corresponding to the smallest value of ν are represented with discontinuous lines in both plots so that they can be seen superimposed to those corresponding to larger values of ν. Slopes of least squares fit to the results corresponding to each value of ν are shown, so that the order of convergence can be checked. In both cases, $\mu = 0.05$ and $\mu = 0$, $O(h^3)$ errors are obtained for large values of ν, which is what Theorem 3.2 and Remark 3.5 predict. However, for smaller values of ν, while the errors with positive μ become $O(h^2)$ and independent of ν, as Theorem 3.3 predicts, for $\mu = 0$ the method does not have convergent behavior for the values of h shown. (Presumably, the method will show convergence for $h \leq \nu$.)

Finally, we check that the requirement H/h bounded is required for convergence if Lagrange interpolants are used. In Figure 3 we show velocity errors when $h \to 0$ in two different scenarios: $h = H/3$ (left) and H fixed to $H = 0.25$. We see that while $H = 3h$, the method converges as predicted by Theorem 3.12 (the value of $\nu = 10^{-6}$ and that of $\mu = 0.05$). If H is kept fixed, however, the method using the Lagrange Interpolant does not exhibit convergent behavior. We remark, however, that with larger values of β or ν, convergence is not altered as much as in Figure 3 when H/h grows. Nevertheless, this example shows the risks of not keeping (H/h) bounded with Lagrange interpolants.

5. Conclusions.

We have analyzed a semidiscretization in space by inf-sup stable mixed finite elements of a continuous downscaling data assimilation method for the two- and three-dimensional Navier–Stokes equations. The data assimilation method, introduced in [7], combines observational data (measurements) on large spatial scales

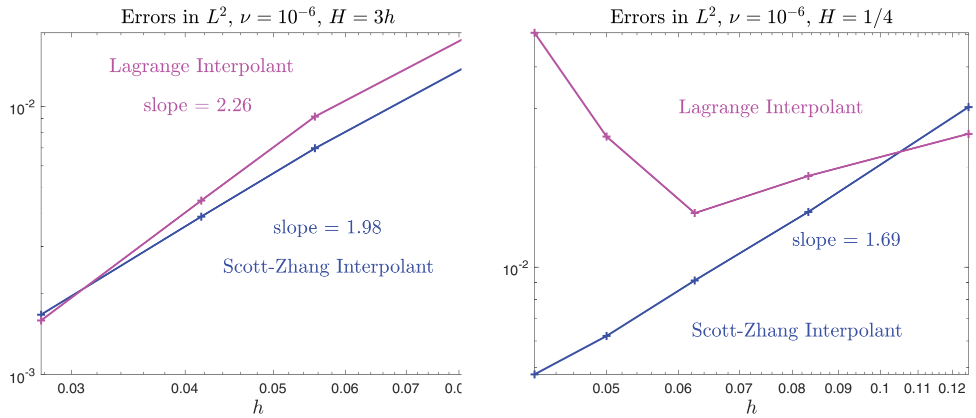

FIG. 3. *Velocity errors. Left, $H = 3h$. Right, $H = 0.25$.*

or coarse mesh, $I_H \boldsymbol{u}$, with simulations in order to improve predictions of the physical phenomenon being studied. We have considered the Galerkin method with and without grad-div stabilization. Uniform error bounds in time have been obtained for the approximation to velocity field, under standard assumptions in finite element analysis. The order of convergence proved for the method without stabilization is optimal, in the sense that it is the best that can be obtained with the finite element space being used (i.e., errors of the same order as interpolation). For the Galerkin method with grad-div stabilization error bounds in which the constants are independent on inverse powers of the viscosity are proved. Convergence rates and the dependence or independence of ν are corroborated in numerical experiments. As opposed to previous works in the literature, our analysis also covers the case in which $I_H \boldsymbol{u}$ is the standard Lagrange interpolant, where we show that H/h must be kept bounded in order to get convergence. Also, the upper bound on the nudging parameter assumed in previous references is removed. The techniques of analysis used in the present paper allow us to improve the available error bounds for a closely related finite element method in [34].

REFERENCES

[1] R. A. Adams and J. J. F. Fournier, *Sobolev Spaces*, 2nd ed., Pure Appl. Math. 140 (Amst.), Elsevier/Academic Press, Amsterdam, 2003.

[2] S. Agmon, *Lectures on Elliptic Boundary Value Problems*, AMS Chelsea Publishing, Providence, RI, 2010, https://doi.org/10.1090/chel/369.

[3] M. U. Altaf, E. S. Titi, T. Gebrael, O. M. Knio, L. Zhao, M. F. McCabe, and I. Hoteit, *Downscaling the 2d Bénard convection equations using continuous data assimilation*, Comput. Geosci., 21 (2017), pp. 393–410, https://doi.org/10.1007/s10596-017-9619-2.

[4] M. Asch, M. Bocquet, and M. Nodet, *Data Assimilation*, Methods, Algorithms, and Applications, Fundam. Algorithms 11, SIAM, Philadelphia, 2016, https://doi.org/10.1137/1.9781611974546.pt1.

[5] B. Ayuso and B. García-Archilla, *Regularity constants of the Stokes problem. Application to finite-element methods on curved domains*, Math. Models Methods Appl. Sci., 15 (2005), pp. 437–470, https://doi.org/10.1142/S021820250500042X.

[6] B. Ayuso, B. García-Archilla, and J. Novo, *The postprocessed mixed finite-element method for the Navier–Stokes equations*, SIAM J. Numer. Anal., 43 (2005), pp. 1091–1111, https://doi.org/10.1137/040602821.

[7] A. Azouani, E. Olson, and E. S. Titi, *Continuous data assimilation using general interpolant observables*, J. Nonlinear Sci., 24 (2014), pp. 277–304, https://doi.org/10.1007/s00332-013-9189-y.

[8] C. Bernardi and V. Girault, *A local regularization operator for triangular and quadrilateral finite elements*, SIAM J. Numer. Anal., 35 (1998), pp. 1893–1916, https://doi.org/10.1137/S0036142995293766.

[9] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, 2nd ed., Texts Appl. Math. 15, Springer-Verlag, New York, 2002, https://doi.org/10.1007/978-1-4757-3658-8.

[10] F. Brezzi and R. S. Falk, *Stability of higher-order Hood–Taylor methods*, SIAM J. Numer. Anal., 28 (1991), pp. 581–590, https://doi.org/10.1137/0728032.

[11] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, Springer Ser. Comput. Math. 15, Springer-Verlag, New York, 1991, https://doi.org/10.1007/978-1-4612-3172-1.

[12] H. Chen, *Pointwise error estimates for finite element solutions of the Stokes problem*, SIAM J. Numer. Anal., 44 (2006), pp. 1–28, https://doi.org/10.1137/S0036142903438100.

[13] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, Classics Appl. Math. 40, SIAM, Philadelphia, 2002, https://doi.org/10.1137/1.9780898719208.

[14] P. Constantin and C. Foias, *Navier-Stokes Equations*, Chicago Lectures in Math. University of Chicago Press, Chicago, IL, 1988.

[15] J. de Frutos, B. García-Archilla, V. John, and J. Novo, *Grad-div stabilization for the evolutionary Oseen problem with inf-sup stable finite elements*, J. Sci. Comput., 66 (2016), pp. 991–1024, https://doi.org/10.1007/s10915-015-0052-1.

[16] J. de Frutos, B. García-Archilla, V. John, and J. Novo, *Analysis of the grad-div stabilization for the time-dependent Navier-Stokes equations with inf-sup stable finite elements*, Adv. Comput. Math., 44 (2018), pp. 195–225, https://doi.org/10.1007/s10444-017-9540-1.

[17] J. de Frutos, B. García-Archilla, and J. Novo, *Error analysis of projection methods for non inf-sup stable mixed finite elements: The Navier-Stokes equations*, J. Sci. Comput., 74 (2018), pp. 426–455, https://doi.org/10.1007/s10915-017-0446-3.

[18] A. Farhat, H. Johnston, M. Jolly, and E. S. Titi, *Assimilation of nearly turbulent Rayleigh–Bénard flow through vorticity or local circulation measurements: A computational study*, J. Sci. Comput., 77 (2018), pp. 1519–1533, https://doi.org/10.1007/s10915-018-0686-x.

[19] A. Farhat, M. S. Jolly, and E. S. Titi, *Continuous data assimilation for the 2D Bénard convection through velocity measurements alone*, Phys. D, 303 (2015), pp. 59–66, https://doi.org/10.1016/j.physd.2015.03.011.

[20] A. Farhat, E. S. Lunasin, and E. S. Titi, *On the Charney conjecture of data assimilation employing temperature measurements alone: The paradigm of 3d planetary geostrophic model,*, Math. Clim. Weather Forecast., 2 (2016), pp. 59–66, https://doi.org/10.1515/mcwf-2016-0004.

[21] L. P. Franca and T. J. R. Hughes, *Two classes of mixed finite element methods*, Comput. Methods Appl. Mech. Engrg., 69 (1988), pp. 89–129, https://doi.org/10.1016/0045-7825(88)90168-5.

[22] G. P. Galdi, *An Introduction to the Mathematical Theory of the Navier-Stokes Equations*, 2nd ed., Springer Monogr. Math., Springer, New York, 2011, https://doi.org/10.1007/978-0-387-09620-9.

[23] B. García-Archilla, J. Novo, and E. S. Titi, *Postprocessing the Galerkin method: A novel approach to approximate inertial manifolds*, SIAM J. Numer. Anal., 35 (1998), pp. 941–972, https://doi.org/10.1137/S0036142995296096.

[24] B. García-Archilla, J. Novo, and E. S. Titi, *An approximate inertial manifolds approach to postprocessing the Galerkin method for the Navier-Stokes equations*, Math. Comp., 68 (1999), pp. 893–911, https://doi.org/10.1090/S0025-5718-99-01057-1.

[25] M. Gesho, E. Olson, and E. S. Titi, *A computational study of a data assimilation algorithm for the two-dimensional Navier-Stokes equations*, Commun. Comput. Phys., 19 (2016), pp. 1094–1110.

[26] V. Girault and J.-L. Lions, *Two-grid finite-element schemes for the transient Navier-Stokes problem*, ESAIM Math. Model. Numer. Anal., 35 (2001), pp. 945–980, https://doi.org/10.1051/m2an:2001145.

[27] V. Girault and P.-A. Raviart, *Finite Element Methods for Navier-Stokes Equations: Theory and Algorithms*, Springer Ser. Comput. Math. 5, Springer-Verlag, New York, 1986, https://doi.org/10.1007/978-3-642-61623-5.

[28] K. Hayden, E. Olson, and E. S. Titi, *Discrete data assimilation in the Lorenz and 2D Navier-Stokes equations*, Phys. D, 240 (2011), pp. 1416–1425, https://doi.org/10.1016/j.physd.2011.04.021.

[29] J. G. Heywood and R. Rannacher, *Finite element approximation of the nonstationary Navier-Stokes problem.* I. *Regularity of solutions and second-order error estimates for*

spatial discretization, SIAM J. Numer. Anal., 19 (1982), pp. 275–311, https://doi.org/10.1137/0719018.

[30] J. G. Heywood and R. Rannacher, *Finite element approximation of the nonstationary Navier–Stokes problem.* III. *Smoothing property and higher order error estimates for spatial discretization*, SIAM J. Numer. Anal., 25 (1988), pp. 489–512, https://doi.org/10.1137/0725032.

[31] H. A. Ibdah, C. F. Mondaini, and E. S. Titi, *Fully discrete numerical schemes of a data assimilation algorithm: uniform-in-time error estimates*, IMA J. Numer. Anal., (2019), https://doi.org/10.1093/imanum/drz043.

[32] V. John and A. Kindl, *Numerical studies of finite element variational multiscale methods for turbulent flow simulations*, Comput. Methods Appl. Mech. Engrg., 199 (2010), pp. 841–852, https://doi.org/10.1016/j.cma.2009.01.010.

[33] E. Kalnay, *Atmospheric Modeling, Data Assimilation and Predictability*, Cambridge University Press, Cambridge, 2002, https://doi.org/10.1017/CBO9780511802270.

[34] A. Larios, L. G. Rebholz, and C. Zerfas, *Global in Time Statbility and Accuracy of IMEX-FEM Data Assimilation Schemes for the Navier-Stokes Equations*, preprint, https://arxiv.org/abs/1805.04090, 2018.

[35] K. Law, A. Stuart, and K. Zygalakis, *Data Assimilation: A Mathematical Introduction*, Texts Appl. Math. 62, Springer-Verlag, New York, 2015, https://doi.org/10.1007/978-3-319-20325-6.

[36] E. Lunasin and E. S. Titi, *Finite determining parameters feedback control for distributed nonlinear dissipative systems—A computational study*, Evol. Equ. Control Theory, 6 (2017), pp. 535–557, https://doi.org/10.3934/eect.2017027.

[37] C. F. Mondaini and E. S. Titi, *Uniform-in-time error estimates for the postprocessing Galerkin method applied to a data assimilation algorithm*, SIAM J. Numer. Anal., 56 (2018), pp. 78–110, https://doi.org/10.1137/16M110962X.

[38] M. A. Olshanskii, *A low order Galerkin finite element method for the Navier-Stokes equations of steady incompressible flow: A stabilization issue and iterative methods*, Comput. Methods Appl. Mech. Engrg., 191 (2002), pp. 5515–5536, https://doi.org/10.1016/S0045-7825(02)00513-3.

[39] M. A. Olshanskii and A. Reusken, *Grad-div stabilization for Stokes equations*, Math. Comp., 73 (2004), pp. 1699–1718, https://doi.org/10.1090/S0025-5718-03-01629-6.

[40] L. G. Rebholz and C. Zerfas, *Simple and Efficient Continuous Data Assimilation of Evolution Equations via Algebraic Nudging*, https://arxiv.org/abs/1810.03512, 2018.

[41] S. Reich and C. Cotter, *Probabilistic Forecasting and Bayesian Data Assimilation*, Cambridge University Press, Cambridge, 2015, https://doi.org/10.1017/CBO9781107706804.

[42] L. Röhe and G. Lube, *Analysis of a variational multiscale method for large-eddy simulation and its application to homogeneous isotropic turbulence*, Comput. Methods Appl. Mech. Engrg., 199 (2010), pp. 2331–2342, https://doi.org/10.1016/j.cma.2010.02.015.

[43] A. H. Schatz, *Pointwise error estimates and asymptotic error expansion inequalities for the finite element method on irregular grids.* I. *Global estimates*, Math. Comp., 67 (1998), pp. 877–899, https://doi.org/10.1090/S0025-5718-98-00959-4.

[44] A. H. Schatz and L. B. Wahlbin, *On the quasi-optimality in L_∞ of the $\dot{H}^1$-projection into finite element spaces*, Math. Comp., 38 (1982), pp. 1–22, https://doi.org/10.2307/2007461.

[45] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493, https://doi.org/10.2307/2008497.

[46] C. Taylor and P. Hood, *A numerical solution of the Navier-Stokes equations using the finite element technique*, Internat. J. Comput. & Fluids, 1 (1973), pp. 73–100, https://doi.org/10.1016/0045-7930(73)90027-3.