On the convergence order of the finite element error in the kinetic energy for high Reynolds number incompressible flows

Bosco García-Archilla[a], Volker John[b,c], Julia Novo[d,*]

[a] *Departamento de Matemática Aplicada II, Universidad de Sevilla, Sevilla, Spain*
[b] *Weierstrass Institute for Applied Analysis and Stochastics (WIAS), Mohrenstr. 39, 10117 Berlin, Germany*
[c] *Freie Universität Berlin, Department of Mathematics and Computer Science, Arnimallee 6, 14195 Berlin, Germany*
[d] *Departamento de Matemáticas, Universidad Autónoma de Madrid, Spain*

Abstract

The kinetic energy of a flow is proportional to the square of the $L^2(\Omega)$ norm of the velocity. Given a sufficient regular velocity field and a velocity finite element space with polynomials of degree r, then the best approximation error in $L^2(\Omega)$ is of order $r + 1$. In this survey, the available finite element error analysis for the velocity error in $L^\infty(0, T; L^2(\Omega))$ is reviewed, where T is a final time. Since in practice the case of small viscosity coefficients or dominant convection is of particular interest, which may result in turbulent flows, robust error estimates are considered, i.e., estimates where the constant in the error bound does not depend on inverse powers of the viscosity coefficient. Methods for which robust estimates can be derived enable stable flow simulations for small viscosity coefficients on comparatively coarse grids, which is often the situation encountered in practice. To introduce stabilization techniques for the convection-dominated regime and tools used in the error analysis, evolutionary linear convection–diffusion equations are studied at the beginning. The main part of this survey considers robust finite element methods for the incompressible Navier–Stokes equations of order $r - 1$, r, and $r + 1/2$ for the velocity error in $L^\infty(0, T; L^2(\Omega))$. All these methods are discussed in detail. In particular, a sketch of the proof for the error bound is given that explains the estimate of important terms which determine finally the order of convergence. Among them, there are methods for inf–sup stable pairs of finite element spaces as well as for pressure-stabilized discretizations. Numerical studies support the analytic results for several of these methods. In addition, methods are surveyed that behave in a robust way but for which only a non-robust error analysis is available. The conclusion of this survey is that the problem of whether or not there is a robust method with optimal convergence order for the kinetic energy is still open.

* Corresponding author.
    E-mail addresses: bosco@esi.us.es (B. García-Archilla), john@wias-berlin.de (V. John), julia.novo@uam.es (J. Novo).

1. Introduction

Incompressible flows are modeled by considering the conservation of linear momentum and the conservation of mass, leading to the incompressible Navier–Stokes equations, given here already in dimensionless form,

$$
\begin{aligned}
\partial_t \boldsymbol{u} - \nu \Delta \boldsymbol{u} + (\boldsymbol{u} \cdot \nabla)\boldsymbol{u} + \nabla p &= \boldsymbol{f} && \text{in } (0, T] \times \Omega, \\
\nabla \cdot \boldsymbol{u} &= 0 && \text{in } (0, T] \times \Omega, \\
\boldsymbol{u} &= \boldsymbol{0} && \text{in } [0, T] \times \partial \Omega, \\
\boldsymbol{u}(0, \cdot) &= \boldsymbol{u}_0(\cdot) && \text{in } \Omega,
\end{aligned}
\tag{1}
$$

where Ω is a bounded domain in $\mathbb{R}^d$, $d \in \{2, 3\}$, with polyhedral and Lipschitz continuous boundary $\partial \Omega$, $T \in \mathbb{R}$ is a final time, $\boldsymbol{u}$ is the velocity field, p the pressure, $\nu > 0$ the kinematic viscosity coefficient, $\boldsymbol{u}_0$ a given initial velocity, and $\boldsymbol{f}$ represents the external body accelerations acting on the fluid. Eq. (1) is already given as an equation without physical dimensions, as it is used in numerical analysis and simulations. Consider now the underlying physical problem. Let ρ [kg/m³] be the density of the fluid, μ [kg/m s] be its dynamic viscosity and define a characteristic length scale L [m] and a characteristic velocity scale U [m/s]. Then, the dimensionless number

$$
Re = \frac{\rho U L}{\mu}
$$

is called Reynolds number. Utilizing the characteristic time scale L/U [s], then one finds the relation $\nu = Re^{-1}$. Thus, a high Reynolds number flow leads to a small coefficient ν in (1).

An important physical quantity of incompressible flows is the contained energy, which is the sum of the kinetic energy $\frac{1}{2}\|\boldsymbol{u}(t)\|_0^2$ at time $t \in (0, T]$, where $\|\cdot\|_0$ denotes the norm in $L^2(\Omega)$, and the dissipative energy. It is known that a distributional solution exists, satisfying the energy inequality

$$
\frac{1}{2}\|\boldsymbol{u}(t)\|_0^2 + \nu \|\nabla \boldsymbol{u}\|_{L^2((0,t), L^2(\Omega))}^2 \leq \frac{1}{2}\|\boldsymbol{u}(0)\|_0^2 + \int_0^t \int_\Omega \boldsymbol{f} \cdot \boldsymbol{u} \, d\boldsymbol{x} \, dt,
$$

for all $t \in (0, T]$, e.g., see [1, Chapter V, Theorem 3.1.1] or [2, Lemma 7.21]. In the present paper, we will assume sufficient regularity so that weak solutions are indeed strong solutions and, hence, unique. More precisely, Leray–Hopf solutions are considered, e.g., see [3]. The unique weak solution in two dimensions satisfies even an energy equality. The kinetic energy of a flow belongs to the quantities of interest in many simulations.

This paper surveys available analytic results for the order of convergence of the finite element error in the kinetic energy, more precisely, of the convergence of the velocity in $L^\infty(L^2)$, where

$$
L^p(L^s) := L^p(0, T; L^s(\Omega)) = \left\{ v \; : \; \left(\int_0^T \left(\int_\Omega |v|^s \, d\boldsymbol{x} \right)^{p/s} dt \right)^{1/p} < \infty, \; p, s \in [1, \infty) \right\},
$$

with the application to each component of a vector-valued function and the usual replacement of the integrals with the essential supremum if $p = \infty$ or $s = \infty$. Since in applications often the case of small viscosity coefficients is of importance, which leads eventually to turbulent flows, especially error estimates are considered where the constant of the error bound does not explicitly depend on inverse powers of ν, such that it does not blow up for $\nu \to 0$. The case where ν is small (relative to the product of a typical velocity scale and a typical length) is often referred to as convection-dominated, since, if $\nabla \boldsymbol{u}$ and $\Delta \boldsymbol{u}$ have norms of similar sizes, the first order term $(\boldsymbol{u} \cdot \nabla)\boldsymbol{u}$ dominates the second order term $-\nu \Delta \boldsymbol{u}$.

Our aim is to investigate finite element approximations of the Navier–Stokes equations that are suitable for the convection-dominated regime. This requires to analyze how the different terms in the Navier–Stokes equations affect their finite element approximation at high Reynolds numbers. To this end, we begin this survey with a simpler case, time-dependent linear convection–diffusion–reaction equations, in order to study a linear convective term. Then, we consider the incompressibility condition and finally we also add the nonlinear term of the Navier–Stokes equations. It will be analyzed how the different terms affect the rate of convergence of finite element methods for a regular solution and $\nu \to 0$. Typically, in the error bounds, it appears a product of an error constant and of norms of the solution in Sobolev spaces. Our error analysis aims for error bounds with constants that are independent of inverse powers of ν. Of course, one must be aware that in practical situations the norms of the solutions will grow as $\nu \to 0$, but error constants that do not explode as $\nu \to 0$ can be considered as the minimal requirement or

necessary condition for a robust performance of a method in practical convection-dominated problems. Hence, this survey provides both, some insight into which methods are expected to fail in practice, because they do not meet the minimum requirement, and in particular a survey of those methods which are expected to be robust. Finally, from the theoretical point of view, the question of deriving local error bounds for the methods is still open. Since for realistic problems, a high global regularity of the solution cannot be expected, such local error estimates would be more appropriate. That means, to prove bounds for the methods in a subset of the domain where the solution is smooth, and to get optimal bounds for the error depending on norms of the solution in this subdomain. To the best of our knowledge, there is almost no error analysis in this direction and the techniques needed probably have not been developed yet. For evolutionary problems, the only references we know are [4] and [5], where local error bounds for the SUPG method applied to linear evolutionary convection–reaction–diffusion equations are obtained.

Accordingly, we first consider the following time-dependent linear convection–diffusion–reaction problem

$$
\begin{aligned}
\partial_t u - \nu \Delta u + \boldsymbol{b} \cdot \nabla u + cu &= f && \text{in } (0, T] \times \Omega, \\
u &= 0 && \text{on } [0, T] \times \partial \Omega, \\
u(0, \boldsymbol{x}) &= u_0(\boldsymbol{x}) && \text{in } \Omega,
\end{aligned}
\tag{2}
$$

where $\boldsymbol{b}(t, \boldsymbol{x})$, $c(t, \boldsymbol{x})$, and $f(t, \boldsymbol{x})$ are given functions, $\nu > 0$ is a constant diffusion coefficient, and $u_0(\boldsymbol{x})$ is the initial data. Problem (2) describes the passive transport of a scalar quantity, like temperature or concentration, by molecular motion, modeled with the diffusive term $-\nu \Delta u$, and by a convective transport via a flow field $\boldsymbol{b}$. One speaks of a convection-dominated regime if $\nu \ll \|\boldsymbol{b}\|_{L^\infty(L^\infty)} L$, where L is a characteristic length scale. To provide an idea of a typical range for diffusion and convection, in the temperature balance of the crystallization process studied in [6], the diffusion coefficient is around $1.5 \cdot 10^{-7}$ m²/s and the maximal velocity is of order 0.1 m/s. If $L = 1$ m is used for deriving non-dimensionalized equations, the convection in these equations is stronger than the diffusion by around six orders of magnitude.

A common feature of the solutions of both the Navier–Stokes equations (1) and the linear convection–diffusion–reaction equation (2) in the convection-dominated regime is the appearance of (very) small structures. The smallest physically important structures or scales of turbulent incompressible flows are the so-called Kolmogorov scales whose size, away from walls, is proportional to $Re^{-3/4}L$. The solution of (2) typically possesses layers, i.e., regions of a small width of $\mathcal{O}(\nu)$ to $\mathcal{O}(\nu^{1/2})$ with very steep slopes of the gradient.

Standard spatial discretizations of (1) and (2), like central finite differences or the Galerkin finite element method, try to predict the behavior of all important scales of the solution. However, in the convection-dominated regime, the smallest scales are usually much smaller than the affordable mesh width, such that most of the small scales even cannot be represented on such grids. Of course, scales that cannot be represented cannot be simulated. Standard discretizations cannot cope with this situation and they lead usually either to a blow up of the simulations or to meaningless numerical solutions that are globally polluted with spurious oscillations, compare also the numerical results presented in Section 4.1.3.

The remedy consists in using so-called stabilized discretizations, which are usually denoted as turbulence models in the case of turbulent flows. Such discretizations introduce in some sophisticated way numerical (artificial) diffusion/viscosity in the discrete equations. A mile stone in this direction was the proposal of the Streamline-Upwind Petrov–Galerkin (SUPG) method in [7,8] for steady-state convection–diffusion–reaction equations. The first finite element error analysis for the steady-state as well the time-dependent convection–diffusion equations can be found in [4,9]. The analysis was extended to the steady-state Navier–Stokes equations in [10,11] and the time-dependent equations in [12,13]. The SUPG method is the core of a variational multiscale turbulence model proposed in [14]. For other stabilized discretizations, like the Continuous Interior Penalty (CIP) method or Local Projection Stabilization (LPS) methods, one can observe a similar development: they were first proposed for the steady-state convection–diffusion–reaction equation and then extended to other equations, including (2) and (1). Since the same ideas can be utilized for discretizing scalar convection–diffusion–reaction equations and incompressible flow equations, these classes of equations are sometimes jointly presented in the literature, like in the most often cited monograph in the field [15]. For completeness, it should be noted that, besides the purely mathematical approaches for turbulence modeling mentioned above, there are many turbulence models whose derivation is based on physical insight in turbulent flows, e.g., see [16], which do not have a counterpart for the scalar equation (2). Often, a numerical analysis for such turbulence models is not available.

Finite element convergence analysis of partial differential equations usually assumes that the solution of these equations is sufficiently regular, i.e., that it belongs to certain Sobolev spaces. Then, the goal consists in deriving

so-called optimal estimates, where the order of convergence equals the order of the best approximation error in the corresponding norm. The techniques for proving such estimates differ for different norms. For convection-dominated problems, there is an additional important aspect: one has to track the dependency of the error bound on the data of the problem, in particular on the coefficient ν. The goal is to obtain error bounds where the constant does not contain inverse powers of ν and therefore it does not blow up as $\nu \to 0$. Such estimates are called robust or semi-robust. The latter notion is inspired from the fact that the error bounds contain norms of the solution of the partial differential equation in Sobolev spaces, which are usually of higher order with respect to the spatial variable, these norms might depend implicitly on ν, and they might grow as $\nu \to 0$. Another aspect for convection-dominated problems is that the convergence of the method, in the sense that the mesh width tends to zero, is not of interest, because if the mesh is sufficiently fine such that all important scales are resolved, a stabilization is not longer necessary. Instead, the behavior on coarse grids is of importance, especially in practice. To distinguish these aspects, one speaks often of 'order of error reduction' or 'order of error decay' instead of 'order of convergence' in the latter situation.

The main topic of this survey is the provable order of error reduction of the velocity error, where with respect to the spatial variable the space $L^2(\Omega)^d$ is considered. Given a finite element velocity space with continuous piecewise polynomials of order r and assuming for the velocity solution of (1) that $\boldsymbol{u}(t) \in H^{r+1}(\Omega)^d$, then the best approximation error in $L^2(\Omega)^d$ is of order $r + 1$. Since for the kinetic energy error, applying the Cauchy–Schwarz inequality, one has

$$\|\boldsymbol{u}(t)\|_0^2 - \|\boldsymbol{u}_h(t)\|_0^2 \leq \|\boldsymbol{u}(t) - \boldsymbol{u}_h(t)\|_0 \|\boldsymbol{u}(t) + \boldsymbol{u}_h(t)\|_0,$$

this error is proportional to the $L^2(\Omega)$ norm of the velocity error and the optimal order of error reduction for the kinetic energy is $(r + 1)$. The finite element error analysis of time-dependent problems leads usually to estimates where a sum of errors in different norms is finally bounded. One of these terms, emerging from the temporal derivative, is the velocity error in $L^\infty(L^2)$. To facilitate the comparison with the literature, also in this survey the error analysis will be presented in terms of the velocity error in $L^\infty(L^2)$ instead of the kinetic energy. One of our motivations for compiling this survey is an open problem stated recently in [17]: *It is open whether optimal and semi-robust $L^\infty(0, T; L^2(\Omega)^d)$ error bounds for the velocity can be proved for some method.*

The present paper provides a survey on the state of the art of error estimates in $L^\infty(L^2)$[1] for the convection-dominated regime of problems (1) and (2). The available analysis was revised for this survey such that the formulas are dimensionally correct, i.e., if all quantities in a formula are equipped with physical units, then the operations in this formula are well defined. In this way, the numerical analysis becomes physically consistent. The complete error analysis of finite element discretizations of (1) and (2) needs usually several pages of estimates. In this survey, we restrict the presentation only to the key steps of the analysis, which lead to the dominating term in the error bounds. With this presentation, also important differences of the analysis between different methods are illustrated. Further aspects of the realization of the methods, e.g., fully implicit approaches vs. implicit–explicit (IMEX) methods, solution as a coupled velocity–pressure problem vs. projection methods, etc. will not be discussed here to focus the presentation on the main topic. Several results from the finite element error analysis are supported with numerical examples. Here, the expectations are illustrated for an academic problem with smooth solution, because only with such a setup, it is possible to support the rates of convergence predicted by the analysis. Numerical simulations for realistic problems, i.e., turbulent flows, possess usually quantities of interest that are not covered by the numerical analysis and such simulations are outside the scope of this survey.

Section 2 sets up weak formulations of (1) and (2), introduces some notations, and provides some tools that are used in the numerical analysis. Discretizations of the convection–diffusion–reaction equation (2) are studied in Section 3. The bulk of the paper, Section 4, is devoted to finite element discretizations of the incompressible Navier–Stokes equations (1). The current state of the art is summarized in Section 5.

2. Weak formulations and notations

Throughout the paper, standard notations will be used for Lebesgue and Sobolev spaces. The dimension is not indicated in symbols of time–space function spaces and the corresponding norms.

---

[1] When surveying the literature, error bounds will be discussed that were obtained for fully discrete problems. Strictly speaking, these bounds are valid only in the discrete time instants and not in the whole time interval. For not unnecessarily complicating the notation, we will use $L^\infty(L^2)$ also in the case of fully discrete problems.

For the convection–diffusion–reaction equation (2), it will be assumed that $\boldsymbol{b}$ and c are sufficiently smooth functions with respect to $\boldsymbol{x}$ and that

$$\mu_0 \le \mu(t, \boldsymbol{x}) = \left(c - \frac{1}{2} \nabla \cdot \boldsymbol{b} \right)(t, \boldsymbol{x}) \quad \forall\, (t, \boldsymbol{x}) \in [0, T] \times \Omega \tag{3}$$

is satisfied. Let $V = H_0^1(\Omega)$. A variational form of (2) reads as follows: Given $f \in H^{-1}(\Omega)$, find $u \,:\, (0, T] \to V$ such that

$$(\partial_t u, v) + \nu(\nabla u, \nabla v) + (\boldsymbol{b} \cdot \nabla u, v) + (cu, v) = \langle f, v \rangle \quad \forall\, v \in V, \tag{4}$$

and $u(0, \boldsymbol{x}) = u_0(\boldsymbol{x})$. Here, $(\cdot, \cdot)$ denotes the inner product in $L^2(\Omega)^d$ and $\langle \cdot, \cdot \rangle$ the dual pairing of $H_0^1(\Omega)^d$ and $H^{-1}(\Omega)^d$, $d \in \{1, 2, 3\}$. Inner products in $L^2(\omega)$ with $\omega \ne \Omega$ will be denoted with $(\cdot, \cdot)_\omega$ and a similar convention is used for the notation of norms.

Using for the incompressible Navier–Stokes equations (1) the function spaces

$$\boldsymbol{V} = H_0^1(\Omega)^d, \quad Q = L_0^2(\Omega) = \left\{ q \in L^2(\Omega) \,:\, (q, 1) = 0 \right\},$$

a weak formulation of problem (1) is as follows: Given $\boldsymbol{f} \in H^{-1}(\Omega)^d$, find $(\boldsymbol{u}, p) \,:\, (0, T] \to \boldsymbol{V} \times Q$ such that for all $(\boldsymbol{v}, q) \in \boldsymbol{V} \times Q$,

$$(\partial_t \boldsymbol{u}, \boldsymbol{v}) + \nu(\nabla \boldsymbol{u}, \nabla \boldsymbol{v}) + ((\boldsymbol{u} \cdot \nabla)\boldsymbol{u}, \boldsymbol{v}) - (\nabla \cdot \boldsymbol{v}, p) + (\nabla \cdot \boldsymbol{u}, q) = \langle \boldsymbol{f}, \boldsymbol{v} \rangle, \tag{5}$$

and $\boldsymbol{u}(0, \cdot) = \boldsymbol{u}_0(\cdot) \in \boldsymbol{V}$. An important subspace of $\boldsymbol{V}$ is the space of weakly divergence-free functions

$$\boldsymbol{V}^{\mathrm{div}} = \{\boldsymbol{v} \in \boldsymbol{V} \,:\, (\nabla \cdot \boldsymbol{v}, q) = 0 \quad \forall\, q \in Q\}. \tag{6}$$

Let $\{\mathcal{T}_h\}$ be a family of triangulations of Ω, where $h > 0$ indicates the fineness of $\mathcal{T}_h$. The mesh cells of $\mathcal{T}_h$ are denoted by K with h_K being the diameter of K and $h = \max_{K \in \mathcal{T}_h} h_K$. It will be assumed that the meshes are conforming and quasi-uniform. In practice, anisotropic grids are of considerable importance. However, the analysis of finite element discretizations on such grids has many open questions, in particular there are no contributions with respect to the topic of this survey.

In numerical simulations, the infinite-dimensional spaces are replaced by finite-dimensional ones, e.g., V by $V_{h,r}$, $\boldsymbol{V}$ by $\boldsymbol{V}_{h,r}$, and Q by $Q_{h,l}$ where $r, l \in \mathbb{N}$ denote the degrees of the local finite element polynomials. Most of the available finite element analysis is for the case of conforming finite element spaces, i.e., the finite element spaces are subspaces of the corresponding infinite-dimensional spaces, e.g., $\boldsymbol{V}_{h,r} \subset \boldsymbol{V}$.

For the Navier–Stokes equations, the space of discretely divergence-free functions is defined by

$$\boldsymbol{V}_h^{\mathrm{div}} = \left\{ \boldsymbol{v}_h \in \boldsymbol{V}_{h,r} \,:\, (\nabla \cdot \boldsymbol{v}_h, q_h) = 0 \quad \forall\, q_h \in Q_{h,l} \right\}. \tag{7}$$

Clearly, there are less conditions for a function to be contained in $\boldsymbol{V}_h^{\mathrm{div}}$ than to belong to $\boldsymbol{V}^{\mathrm{div}}$. In fact, for many pairs $\boldsymbol{V}_{h,r}/Q_{h,l}$ of finite element spaces, it turns out that $\boldsymbol{V}_h^{\mathrm{div}} \not\subset \boldsymbol{V}^{\mathrm{div}}$.

The following inverse inequality holds for each $v_h \in V_{h,r}$, e.g., see [18, Theorem 3.2.6],

$$\|v_h\|_{W^{m,q}(K)} \le c_{\mathrm{inv}} h_K^{l-m-d\left(\frac{1}{q'} - \frac{1}{q}\right)} \|v_h\|_{W^{l,q'}(K)}, \tag{8}$$

where $0 \le l \le m \le 1$, $1 \le q' \le q \le \infty$, and $\|\cdot\|_{W^{m,q}(K)}$ is the norm in $W^{m,q}(K)$.

We will denote by $\pi_h u \in V_{h,r}$ the elliptic projection defined by

$$(\nabla \pi_h u, \nabla \varphi_h) = (\nabla u, \nabla \varphi_h) \quad \forall \varphi_h \in V_{h,r}. \tag{9}$$

The following bound holds for any $u \in H_0^1(\Omega) \cap H^{r+1}(\Omega)$, e.g., see [18],

$$\|u - \pi_h u\|_0 + h\|u - \pi_h u\|_1 \le Ch^{r+1}\|u\|_{r+1}. \tag{10}$$

Generic constants that do not depend neither on the mesh width h nor on the diffusion/viscosity coefficient ν will be denoted by C.

To avoid technicalities that are not significant for the topic of this survey, the actual regularity assumptions on the weak solution will not be stated explicitly in the formulation of the theorems, but it will be assumed that the solution is sufficiently regular such that all terms appearing in the error bounds are well defined.

3. Convection–diffusion–reaction equations

The first part of this section discusses briefly the Galerkin finite element discretization of the convection–diffusion–reaction equation (4). If the best approximation error in $L^2(\Omega)$ is $r+1$, then an error bound in $L^\infty(L^2)$ with constants independent of inverse powers of the diffusion coefficient of order r can be proved. Stabilized methods, with the emphasis on the SUPG method, are the topic of the second part of the section. For these methods, robust estimates of order $r + 1/2$ can be derived.

3.1. The Galerkin method

The time-continuous finite element problem aims to find a function $u_h \in V_{h,r}$ that fulfills a problem of form (4) for all test functions from $V_{h,r}$. More precisely, $u_h : (0, T] \to V_{h,r}$ is called Galerkin approximation if

$$(\partial_t u_h, v_h) + \nu(\nabla u_h, \nabla v_h) + (\boldsymbol{b} \cdot \nabla u_h, v_h) + (c u_h, v_h) = \langle f, v_h \rangle \quad \forall \, v_h \in V_{h,r}, \tag{11}$$

with $u_h(0, \boldsymbol{x})$ being an appropriate approximation of $u_0(\boldsymbol{x})$.

It is well known that the standard Galerkin method leads to numerical solutions with non-physical spurious oscillations in the convection-dominated regime when the solution of (4) possesses layers. However, if the solution of (4) does not exhibit layers and if it is sufficiently smooth, the Galerkin method produces meaningful approximations that tend to the solution of (4) with an order of error reduction r in $L^\infty(L^2)$, independently of the value of ν, which will be proved in Theorem 3.1 and illustrated numerically in Example 3.2.

In most works on convection-dominated problems, the case $\mu_0 > 0$, the constant in (3), is assumed; see [15]. In Theorem 3.1, we consider both cases $\mu_0 > 0$ and $\mu_0 \le 0$.

Theorem 3.1 (*Error Estimate for the Galerkin Method*). *Let u be the sufficiently smooth solution of* (2) *and u_h be the finite element approximation defined in* (11). *Assume $\boldsymbol{b} \in L^\infty(L^\infty)$ and $c \in L^\infty(L^\infty)$, then the following bound holds for $0 < t \le T$*

$$\|u(t) - u_h(t)\|_0 \le e^{-\mu_0 t} \|u_h(0) - \pi_h u_0\|_0^2 + C h^r K_u \frac{\left|1 - e^{-\mu_0 t}\right|}{|\mu_0|} + C h^{r+1} \|u\|_{L^\infty(H^{r+1})}, \tag{12}$$

the bound being also valid for $\mu_0 = 0$ if $\left|1 - e^{-\mu_0 t}\right|/|\mu_0|$ is replaced by t. The constant is

$$K_u = \|\partial_t u\|_{L^\infty(H^r)} + \left(\|\boldsymbol{b}\|_{L^\infty(L^\infty)} + h\|c\|_{L^\infty(L^\infty)}\right) \|u\|_{L^\infty(H^{r+1})}. \tag{13}$$

Proof. The error is split into $e = u - u_h = (u - \pi_h u) - (u_h - \pi_h u) = \eta - e_h$, where $\pi_h u$ is the elliptic projection defined in (9). Subtracting (11) from (2) gives the error equation

$$(\partial_t e_h, v_h) + \nu(\nabla e_h, \nabla v_h) + (\boldsymbol{b} \cdot \nabla e_h, v_h) + (c e_h, v_h) = (\tau_1 + \tau_2, v_h), \tag{14}$$

with the truncation errors $\tau_1 = (\partial_t u - \pi_h \partial_t u) + c \eta$ and $\tau_2 = \boldsymbol{b} \cdot \nabla \eta$. Here, the commutation of elliptic projection and temporal differentiation $\partial_t(\pi_h u) = \pi_h \partial_t u$ was used. Choosing $v_h = e_h$ in the error equation (14) and integrating by parts yields

$$\frac{1}{2} \frac{d}{dt} \|e_h\|_0^2 + \mu_0 \|e_h\|_0^2 \le (\tau_1 + \tau_2, e_h). \tag{15}$$

Because of using the elliptic projection, the right-hand side of (15) does not depend on ν. The non-negative term with ν on the left-hand side was neglected. Altogether, estimate (15) does not contain ν such that the arising error bound will also not depend on ν.

Applying the Cauchy–Schwarz inequality gives

$$(\tau_1 + \tau_2, e_h) \le (\|\tau_1\|_0 + \|\tau_2\|_0) \|e_h\|_0,$$

and then, arguing as in [19, Lemma 3.1] leads to

$$\|e_h\|_0 \frac{d}{dt} \|e_h\|_0 + \mu_0 \|e_h\|_0^2 = \frac{1}{2} \frac{d}{dt} \|e_h\|_0^2 + \mu_0 \|e_h\|_0^2 \le (\|\tau_1\|_0 + \|\tau_2\|_0) \|e_h\|_0,$$

which yields

$$\frac{d}{dt}\|e_h\|_0 + \mu_0\|e_h\|_0 \le \|\tau_1\|_0 + \|\tau_2\|_0.$$

Multiplying by $e^{\mu_0 t}$ and integrating with respect to time, one obtains

$$\|e_h(t)\|_0 \le e^{-\mu_0 t}\|e_h(0)\|_0 + \int_0^t e^{-\mu_0(t-s)}(\|\tau_1(s)\|_0 + \|\tau_2(s)\|_0)\, ds$$

$$\le e^{-\mu_0 t}\|e_h(0)\|_0 + \frac{(1 - e^{-\mu_0 t})}{\mu_0} \max_{0 \le s \le T}(\|\tau_1(s)\|_0 + \|\tau_2(s)\|_0), \tag{16}$$

where, in the last term, the factor $(1 - e^{-\mu_0 t})/\mu_0$ must be replaced by t if $\mu_0 = 0$. Utilizing (10) and the assumptions on the regularity of the coefficient functions leads to

$$\max_{0 \le s \le T}(\|\tau_1(s)\|_0 + \|\tau_2(s)\|_0) \le Ch^r K_u, \tag{17}$$

where K_u is the constant in (13). Inserting (17) in (16) and applying the triangle inequality together with (10) gives (12). □

A similar result to that of Theorem 3.1 can be found in [20], where the error of the time derivative $\|\partial_t u - \partial_t u_h\|_0$ is proved to be $\mathcal{O}(h^r)$ with a constant independent of inverse powers of ν, assuming sufficient regularity for the solution.

Example 3.2 (*Galerkin Method*). Let $\Omega = (0, 1)^2$ and consider the convection–diffusion–reaction equation (2) with the prescribed solution

$$u(x, y) = \sin(\pi t)\sin(2\pi x)\sin(2\pi y),$$

the convection field $\boldsymbol{b} = (1 - y, -2/3 + x)^T$, the reaction field $c = 1$, the final time $T = 2$, and Dirichlet boundary conditions on $\partial\Omega$.

Fig. 1 displays the initial grid (level 1). This grid was refined uniformly using the standard regular refinement by dividing each triangle into four triangles. Piecewise linear P_1 and piecewise quadratic P_2 finite elements were used, see Table 1 for information on the numbers of degrees of freedom.

To mimic the continuous-in-time situation, a second order time stepping scheme, the Crank–Nicolson scheme, was applied with the very small time step $\Delta t = 10^{-5}$. We checked that almost the same results were obtained for $\Delta t = 10^{-4}$, such that the temporal error can be considered to be negligible.

If not mentioned otherwise, the linear systems of equations in all examples of this paper were solved with the sparse direct solver UMFPACK [21].

The computational results obtained with the Galerkin method for four small values of the diffusion coefficient are presented in Fig. 2. First of all, one can see that the errors in $L^\infty(L^2)$ are small, which indicates that for this smooth problem even the standard Galerkin method computes meaningful numerical solutions. For P_1 finite elements, the optimal second order of convergence can be observed in Fig. 2. In contrast, for P_2 finite elements, the error reduction on coarse grids is not optimal. In the cases $\nu = 10^{-6}$ and $\nu = 10^{-8}$, it is only of second order, exactly as predicted by the analysis. For $\nu = 10^{-4}$, it can be already seen that the order of convergence improves to third order when the grids become sufficiently fine. Finally, if the diffusion is sufficiently large, $\nu = 10^{-2}$, the optimal order is observed already on very coarse grids. Since similar observations were made in all other numerical studies for convection–diffusion–reaction equations and the emphasis of this survey is on the convection-dominated regime, only numerical results for $\nu \le 10^{-4}$ will be presented in the remainder of this section. □

3.2. Stabilized methods of order $r + 1/2$ in $L^\infty(L^2)$

As mentioned before, the Galerkin method produces numerical solutions in the convection-dominated regime that are globally polluted with spurious oscillations if the solution of (4) possesses layers. In this situation, one has to introduce a stabilizing mechanism in the discretization. In the ideal case, this mechanism removes all spurious oscillations and gives numerical solutions with sharp layers. However, numerical assessments in [22,23] show that there are only very few finite element methods with these properties, which are methods that use a so-called algebraic stabilization. Most stabilized finite element methods localize the spurious oscillations in a vicinity of the layers and reduce them considerably compared with the Galerkin method.
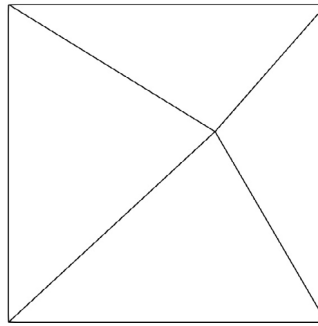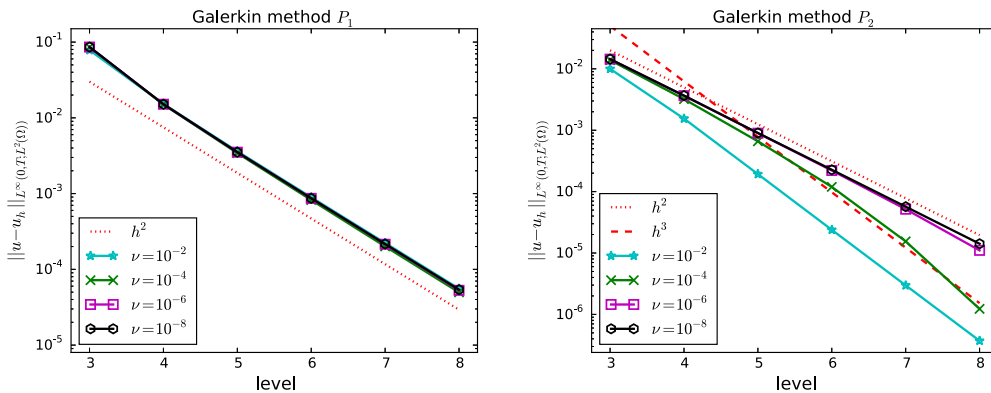
Fig. 1. Initial grid (level 1).

Fig. 2. Example 3.2 Galerkin method for the convection–diffusion–reaction equation.

Table 1
Numerical examples for convection–diffusion–reaction
equation, number of degrees of freedom (including
Dirichlet nodes).

Level	P_1	P_2
3	41	145
4	145	545
5	545	2113
6	2113	8321
7	8321	33025
8	33025	131585

3.2.1. The SUPG method

The most popular method of this kind is the streamline-upwind Petrov–Galerkin (SUPG) method. This residual-based stabilization adds artificial diffusion along the streamlines of the solution. It reads as follows: Given $f \in L^2(\Omega)$, find $u_h : (0, T] \to V_{h,r}$ such that

$$(\partial_t u_h, v_h) + a_{\text{SUPG}}(u_h, v_h) + \sum_{K \in \mathcal{T}_h} \delta_K (\partial_t u_h, \boldsymbol{b} \cdot \nabla v_h)_K = (f, v_h) + \sum_{K \in \mathcal{T}_h} \delta_K (f, \boldsymbol{b} \cdot \nabla v_h)_K \quad \forall \, v_h \in V_{h,r}, \quad (18)$$

with $u_h(0, \boldsymbol{x})$ being an appropriate approximation of $u_0(\boldsymbol{x})$ and

$$a_{\text{SUPG}}(u_h, v_h) = \nu (\nabla u_h, \nabla v_h) + (\boldsymbol{b} \cdot \nabla u_h, v_h) + (c u_h, v_h) + \sum_{K \in \mathcal{T}_h} \delta_K (-\varepsilon \Delta u_h + \boldsymbol{b} \cdot \nabla u_h + c u_h, \boldsymbol{b} \cdot \nabla v_h)_K.$$

Here, $\{\delta_K\}$ are local parameters that have to be chosen appropriately.

Let u have the physical unit unit(u), then the term $(\partial_t u_h, v_h)$ possesses the unit $\text{unit(u)}^2 \text{m}^d / \text{s}$, where the unit m^d comes from integration and the term $1/\text{s}$ from the differentiation with respect to time. For all other terms in (18), one finds the same physical unit $\text{unit(u)}^2 \text{m}^d / \text{s}$ by taking into account the units for the diffusion coefficient m^2 / s, the convection field m/s, and the reaction field $1/\text{s}$. The stabilization term of the SUPG method has the physical unit $\text{unit}(\delta_K) \text{unit(u)}^2 \text{m}^d / \text{s}^2$. In order that adding this term to the other terms is correct from the dimensional point of view, it has to have the same unit as the other terms. Consequently, it follows that $\text{unit}(\delta_K) = \text{s}$. Thus, the stabilization parameter is a time scale.

The numerical analysis for a stabilized method is performed usually in an associated norm. For $u_h \in V_{h,r}$, this norm for the SUPG method is as follows

$$\|u_h\|_{\text{SUPG}} := \left(\nu \|\nabla u_h\|_0^2 + \sum_{K \in \mathcal{T}_h} \delta_K \|\boldsymbol{b} \cdot \nabla u_h\|_{0,K}^2 + \|\mu^{1/2} u_h\|_0^2 \right)^{1/2}.$$

The SUPG method for evolutionary convection–diffusion equations is analyzed in [24]. A main topic in [24], for a full discretization in time and space, is the discussion whether the stabilization parameters $\{\delta_K\}$ should depend on the length of the time step or on the mesh width. The latter situation can be studied by considering the continuous-in-time method, because this method does not contain a time step. Assume that

$$\boldsymbol{b}(t, \boldsymbol{x}) = \boldsymbol{b}(\boldsymbol{x}), \quad \nabla \cdot \boldsymbol{b}(\boldsymbol{x}) = 0, \quad c(t, \boldsymbol{x}) = c(\boldsymbol{x}), \tag{19}$$

$\mu_0 > 0$, the mesh is uniform with mesh width h, and the stabilization parameters are the same for all mesh cells, i.e., $\delta_K = \delta$. Also, considered only the convection-dominated regime, i.e., it is assumed that $\nu \leq \|\boldsymbol{b}\|_{L^\infty} h$. The stabilization parameter is defined to be

$$\delta = \min \left\{ \frac{h}{4 c_{\text{inv}} \|\boldsymbol{b}\|_{L^\infty}} \min \left\{ \frac{1}{2}, \frac{\mu_0}{4 \|c\|_{L^\infty}}, \frac{\mu_0^{1/2}}{\|c\|_{L^\infty}^{1/2}}, \frac{\|\boldsymbol{b}\|_{L^\infty} h}{4 \nu c_{\text{inv}}} \right\}, \frac{1}{\mu_0}, \frac{1}{\|c\|_{L^\infty}} \right\}. \tag{20}$$

Note that the physical unit of δ is seconds. Formula (20) for the stabilization parameter can be obtained with a slight revision of the analysis of [24] by carefully taking into account the dimensional issue. Then, the following error estimate is derived in [24, Theorem 5.2].

Theorem 3.3 (*Error Estimate for the SUPG Method*). *Let $t \leq T < \infty$, let $\nu < \|\boldsymbol{b}\|_{L^\infty} h$, let the solution of (2) be sufficiently regular, and let (19) and (20) be satisfied. Then, there exists a positive constant M_u depending on $\|u_0\|_{r+1}$, $\|u\|_{L^\infty(H^{r+1})}$, $\|\partial_t u(0)\|_{L^\infty(H^{r+1})}$, $\|\partial_{tt} u\|_{L^2(H^{r+1})}$, and the coefficients of the problem, but not on inverse powers of ν, such that the following error estimate holds:*

$$\|(u - u_h)(t)\|_0 + \|u - u_h\|_{L^2(\text{SUPG})} \leq M_u h^{r+1/2}. \tag{21}$$

Proof (*Sketch*). The proof of (21) in [24] covers several pages. Here, only a sketch is given that concentrates on explaining the reason why estimate (21) leads to the order $r + 1/2$ in $L^\infty(L^2)$.

As usual, the error is decomposed into an interpolation or projection error and a discrete remainder. The analysis in [24] defines for this decomposition a function $\Pi_h u(t)$ which is the solution of a formally steady-state problem. The idea is to consider the solution $u(t)$ of the evolutionary problem (2) as the solution of a steady-state problem with right-hand side $f(t) - \partial_t u(t)$. Then, given $t \in (0, T]$, $\Pi_h u(t) \in V_{h,r}$ satisfies

$$a_{\text{SUPG}} (\Pi_h u(t), v_h) = a_{\text{SUPG}} (u(t), v_h) \quad \forall \, v_h \in V_{h,r}, \tag{22}$$

i.e., $\Pi_h u(t)$ is the SUPG approximation to a steady-state problem whose solution is $u(t)$. An error estimate for the solution of the steady-state problem (22) is well known, e.g., see [15, Thm. III, 3.27],

$$\|u(t) - \Pi_h u(t)\|_{\mathrm{SUPG}} \leq C_1 h^{r+1/2} \|u(t)\|_{r+1}, \tag{23}$$

where $C_1 = C\left(\|\boldsymbol{b}\|_{L^\infty} + \delta_0^{-1} + \|c\|_{L^\infty} h\right)^{1/2}$ and δ_0 satisfies $0 < \delta_0 < (\delta/h)$.

One obtains with the triangle inequality

$$\|u(t) - u_h(t)\|_{\mathrm{SUPG}} \leq \|u(t) - \Pi_h u(t)\|_{\mathrm{SUPG}} + \|\Pi_h u(t) - u_h(t)\|_{\mathrm{SUPG}}. \tag{24}$$

The first term on the right-hand side is of order $r + 1/2$ by (23). The estimate of the second term is technically involved. In the first step, a weaker norm is bounded: $\|\Pi_h u(t) - u_h(t)\|_0$. To bound this norm requires an estimate of the error $\|\partial_t u(t) - \partial_t(\Pi_h u)(t)\|_0$. Using assumption (19), the convection–diffusion–reaction equation (2) can be differentiated with respect to time to give a convection–diffusion–reaction equation for $\partial_t u$ of the same form as for u and it is straightforward to show that $\partial_t(\Pi_h u)(t) = \Pi_h(\partial_t u)(t)$. Then, applying (23) to $\partial_t u(t)$ yields

$$\|\partial_t u(t) - \partial_t(\Pi_h u)(t)\|_{\mathrm{SUPG}} = \|\partial_t u(t) - \Pi_h(\partial_t u)(t)\|_{\mathrm{SUPG}} \leq C_1 h^{r+1/2} \|\partial_t u(t)\|_{r+1}. \tag{25}$$

The expression $\|\partial_t u(t) - \partial_t(\Pi_h u)(t)\|_0$ can be estimated with (25), since the $L^2(\Omega)$ norm is part of the SUPG norm. In the second step, a bound for $\|\partial_t(\Pi_h u)(t) - \partial_t u_h(t)\|_0$ is obtained using again (19) and applying the error analysis of the first step. Both steps lead finally to a bound for the second term on the right-hand side of (24) in the stronger SUPG norm with order $r + 1/2$.   □

Since the result of Theorem 3.3 holds for $\nu < \|\boldsymbol{b}\|_{L^\infty} h$, such that the situation $h \to 0$ is not covered, the notion error reduction will be used for discussing the error estimate, instead of convergence. The error estimate (21) of Theorem 3.3 states that the order of error reduction of the SUPG method in $L^\infty(L^2)$ is $r + 1/2$, i.e., it is half an order better than for the Galerkin method, but still suboptimal by half an order, while the error reduction of the L^2(SUPG) norm of the streamline derivative is optimal of order r. In [24], also fully discrete cases are considered and error bounds are derived for stabilization parameters that do not depend on the length of the time step. As shown in the sketch of the proof of Theorem 3.3, an essential idea consists in reducing the error estimate for the time-dependent convection–diffusion–reaction equation to an error estimate for a stabilized discretization of a steady-state convection–diffusion–reaction equation. The finite element error analysis of stabilizations for the steady-state equation is quite far developed, see [15]. In particular, checking the convergence analysis of the SUPG method, e.g., [15, Thm. III, 3.27], one finds that the estimates of several terms, namely the diffusive term, the convective term, and terms coming from the stabilization, lead only to order $r + 1/2$. The sharpness of this order for the steady-state problem will be discussed at the end of this section.

Both, a weakly consistent SUPG method and a SUPG method using a reconstructed Laplacian are analyzed in [25], where the order $r + 1/2$ for the error in $L^\infty(L^2)$ is proved. This order comes from using a result from the numerical analysis of the stationary problem.

Example 3.4 (*SUPG Method*). The behavior of the SUPG method is illustrated at the same example that was utilized for the standard Galerkin method, see Example 3.2. Exactly the same setup was used with respect to the temporal discretization and the spatial meshes. The stabilization parameter, constant in each mesh cell, was set to be

$$\delta_K = \frac{1}{16} \begin{cases} h_K & \text{if } \nu \leq h_K, \\ h_K^2/\nu & \text{else.} \end{cases}$$

The choice for the convection-dominated regime corresponds to the proposal of the parameter (20) with respect to the mesh width.

The computational results for three small values of the diffusion coefficient are presented in Fig. 3. It can be observed that for P_1 finite elements, the optimal order of convergence of the error in $L^\infty(L^2)$ can be observed. For P_2 finite elements, the order of error reduction on coarse grids, where coarse is to be understood with respect to the diffusion coefficient, is suboptimal. The analysis predicts an order of $r + 1/2 = 2.5$, which can be clearly observed for $\nu = 10^{-6}$ and 10^{-8}. Thus, the numerical results match the analytic predictions very well.   □
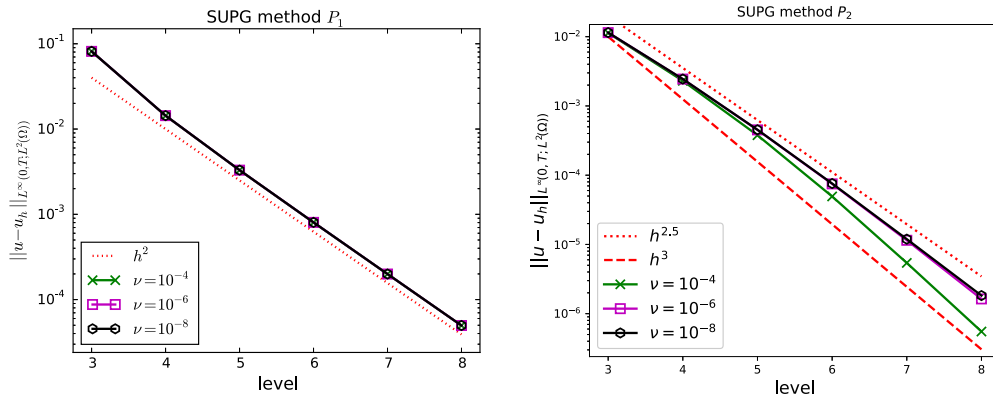
Fig. 3. Example 3.4 SUPG method for the convection–diffusion–reaction equation.

3.2.2. Other stabilized methods

Apart from the SUPG method, other types of stabilizations can be applied and are analyzed in the literature. Galerkin Least Squares (GLS) stabilizations are studied in [26] in combination with one-step θ-schemes as temporal discretization. The continuous-in-time case is not considered. The GLS stabilization term has the form

$$\sum_{K \in \mathcal{T}_h} \delta_K \left(f - (\partial_t u - \nu \Delta u + \boldsymbol{b} \cdot \nabla u + cu), -\nu \Delta v + \boldsymbol{b} \cdot \nabla v + cv \right)_K .$$

Assumption (19) is made in the analysis of the GLS method and the spatial stabilization parameter is assumed to be $\delta_K = \delta$ for all mesh cells K with $\delta = \mathcal{O}(\Delta t)$, Δt being the length of the time step. This type of assumption is sometimes called inverse CFL condition. Observe that under this assumption, the spatial stability vanishes in the continuous-in-time limit $\Delta t \to 0$. Also for the GLS stabilization, the physical unit of the stabilization parameter is a time unit (second) since it is a residual-based stabilization. For applying the GLS stabilization in numerical simulations in [26], a parameter is designed independently of the constraint coming from the length of the time step, by minimizing the 'elliptic part' of the error bound, i.e., those terms that do not arise from the discretization of the temporal derivative. This parameter depends on the mesh width. If the mesh width is sufficiently fine compared with the length of the time step, such that the analytic results are applicable, one obtains in the convection-dominated regime order $r + 1/2$ for the $L^\infty(L^2)$ error.

An orthogonal subscale (OSS) method is analyzed in [27]. This method adds the stabilization term, in the fully discrete case as considered in [27],

$$\sum_{K \in \mathcal{T}_h} \left(\frac{1}{\delta_K} + \frac{1}{\Delta t} \right)^{-1} \left(\Pi^\perp (\boldsymbol{b} \cdot \nabla u^{n+1}), \Pi^\perp (\boldsymbol{b} \cdot \nabla v) \right)_K - \sum_{K \in \mathcal{T}_h} \left(\frac{1}{\delta_K} + \frac{1}{\Delta t} \right)^{-1} \left(\tilde{u}^n, \boldsymbol{b} \cdot \nabla v \right)_K ,$$

where Π is the $L^2(\Omega)$ projection in the finite element space and $\tilde{u}^n$ are so-called orthogonal subscales from the previous time instant. These subscales can be computed locally. The analysis in [27] is performed for the backward Euler scheme. It assumes that the stabilization parameters are constant $\delta_K = \delta$ and that they are bounded by a multiple of the length of the time step: $\delta \leq C \Delta t$ for some fixed positive C. Then, the error analysis leads to a convergence of order $r + 1/2$ in $L^\infty(L^2)$. The reduction comes from estimating the diffusion, the convective, and the stabilization terms with respect to a stabilized norm. In practice, a parameter choice of the form

$$\delta_K = \left(C_1 \frac{\nu}{h_K^2} + C_2 \frac{\|\boldsymbol{b}\|_{L^\infty(K)}}{h_K} + \|c\|_{L^\infty(K)} \right)^{-1}$$

is proposed, which is a time scale. In the convection-dominated regime, the stabilization parameter behaves like $h_K/(C_2 \|\boldsymbol{b}\|_{L^\infty(K)})$.

The local projection stabilization (LPS) method, as analyzed in [28,29], adds a stabilization term of the form

$$\sum_{K \in \mathcal{T}_h} \delta_K \left((\mathrm{Id} - \Pi_K) \nabla u_h, (\mathrm{Id} - \Pi_K) \nabla v_h \right)_K , \tag{26}$$

where Π_K is a local projection operator and Id is the identity operator. Clearly, (26) is a diffusive term, with the additional diffusion acting only on the so-called fluctuations defined by (Id $- \Pi_K$). The goals of both papers [28,29] are to analyze the LPS method for different classes of temporal discretizations. In addition, an analysis for the continuous-in-time situation is provided in [28]. The analysis for the LPS method is performed for a norm that contains a contribution from the stabilization term (26), where the stabilization parameter should be chosen to be $\delta_K \sim h_K$ (note that from the dimensional point of view something like $\delta_K \sim h_K \|\boldsymbol{b}\|_{L^\infty(K)}$ is appropriate since δ_K represents a diffusion). Like in [24], it is assumed that $\boldsymbol{b}$ and c are time-independent functions. As usual, the error is decomposed into an interpolation error, or here projection error, and a discrete remainder. From the results obtained in [28,29], one finds that the order of error reduction in $L^\infty(L^2)$ is $r + 1/2$ in the convection-dominated regime $\nu \leq \|\boldsymbol{b}\|_{L^\infty} h$. In [28], already the considered projection error possesses this order, whereas in [29] this order comes from bounding the discrete remainder. Inspecting the proofs, one finds that in both cases estimates of the diffusive, the convective, and the stabilization term result in this order. A LPS method with fluctuations of the streamline derivative $\boldsymbol{b} \cdot \nabla u_h$ and additional nonlinear crosswind diffusion is analyzed in [30]. The analysis covers the case of time-dependent coefficients. Also for this method, the order of error reduction $r + 1/2$ is obtained for the error in $L^\infty(L^2)$ in the convection-dominated regime $\nu \leq \|\boldsymbol{b}\|_{L^\infty} h$.

A class of symmetric stabilization methods is studied in [31]. Several general assumptions on the stabilizing term are made. A stabilization that satisfies these assumptions is the continuous interior penalty (CIP) method whose stabilization term is given by

$$\frac{\delta}{2} \sum_{K \in \mathcal{T}_h} \sum_{F \in \partial K} \int_F h_F^2 |\boldsymbol{b} \cdot \boldsymbol{n}_F| \, [\![\nabla u_h \cdot \boldsymbol{n}_F]\!]_F \, [\![\nabla v_h \cdot \boldsymbol{n}_F]\!]_F \; ds,$$

where $\delta > 0$ is a constant, F is a facet of K with diameter h_F, $\boldsymbol{n}_F$ is a fixed unit normal of F, and $[\![\cdot]\!]_F$ denotes the jump across F. Other examples that fit into the framework of [31] are certain LPS methods and the OSS method. The analysis is performed for the imposition of Dirichlet boundary conditions via Nitsche's method. Two fully discrete cases are analyzed. Both of them lead to the order of error reduction $r + 1/2$ in $L^\infty(L^2)$ if $\nu \leq \|\boldsymbol{b}\|_{L^\infty} h$. This order is determined by estimates of the diffusive, convective, and stabilization terms.

In [32], a so-called subgrid stabilization is analyzed. To this end, a decomposition $V_h = V_H \oplus V_h^H$ of a finite element space is utilized, where V_H is the standard space, e.g., P_1. The space V_h is constructed with the help of a finer grid and V_h^H is the subgrid space. Then, the stabilization is defined by a bilinear form $V_h^H \times V_h^H \to \mathbb{R}$, which is assumed to be in an appropriate sense coercive and bounded. A straightforward realization consists in defining this bilinear form to be a diffusion acting on the subgrid scales. Another realization uses the streamline derivative of the subgrid scales. The numerical analysis presented in [32] leads to the order of error reduction $r + 1/2$ in $L^\infty(L^2)$ if $\nu \leq \|\boldsymbol{b}\|_{L^\infty} h$. This order comes from interpolation estimates in certain norms that contain the diffusive and the convective term.

Altogether, the principal motivation for the subgrid stabilization from [32], the OSS method analyzed in [27], and LPS methods is the same: to introduce additional (streamline) diffusion on some appropriately defined small or subgrid scales. In this way, all these methods fit in the framework of the variational multiscale method proposed in [33].

Also for the discontinuous Galerkin (DG) approach, a robust finite element analysis is available, see [34] or [35, Chapter 4.6]. In the considered methods, an interior penalty (IP) discretization of the diffusion term is applied. For the coercivity of the corresponding bilinear form, a parameter in a term containing jumps across facets has to be positive for the non-symmetric IP and it has to be also sufficiently large for the symmetric and incomplete IP. The convective term is discretized with an upwind method. It is well known that upwind discretizations in the context of DG methods lead to a control of the streamline derivative of the error, e.g., see [36]. In the situation that $\nu \leq \|\boldsymbol{b}\|_{L^\infty} h$, a robust estimate for the error in $L^\infty(L^2)$ of order $r + 1/2$ can be proved, see [35, Theorem 4.28].

The error analysis of convection–diffusion equations aims at deriving error bounds where the constants are independent of inverse powers of the diffusion coefficient ν. Since the case of small diffusion coefficients is of interest, a natural question is: What is known for the limit case $\nu = 0$, i.e., for linear transport equations? To survey the literature concerning the error analysis of this type of equations is beyond the scope of this paper. That is why, only some selected results will be mentioned here. The SUPG method is studied for linear transport equations in [37]. For fully discrete cases, error bounds are derived such that the error in $L^\infty(L^2)$ is of order $r + 1/2$, where the stabilization parameter depends on the length of the time step. A numerical example for linear finite elements
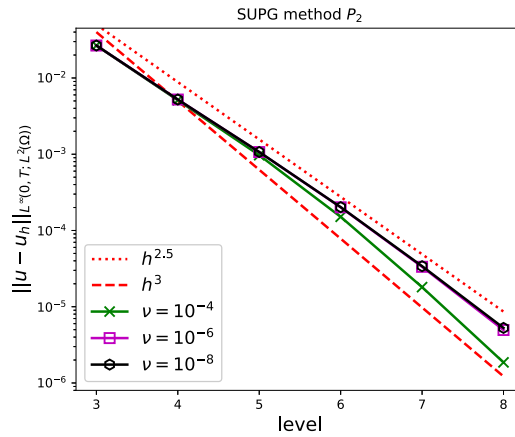
Fig. 4. SUPG method for the convection–diffusion–reaction equation from Example 3.4 on a family of triangulations of Friedrichs–Keller type. The initial grid (level 0) was obtained by dividing the unit square with a diagonal from bottom left to top right. It was refined uniformly using the division of each triangle into four triangles.

is presented which shows second order convergence for the error in $L^\infty(L^2)$. In [32], the subgrid scale method is analyzed also for the transport equation. The order of convergence in $L^\infty(L^2)$ is $r + 1/2$. The work [31] on symmetric stabilizations considers also the case $\nu = 0$, for which the order $r + 1/2$ in $L^\infty(L^2)$ is proved. A certain kind of LPS method is analyzed in [38], for which an error bound for $L^\infty(L^2)$ of order $r + 1/2$ is derived and second order convergence is observed in numerical simulations.

This section will finish with a brief summary.

- For many popular stabilized methods, the order of error reduction $r + 1/2$ for the error in $L^\infty(L^2)$ can be proved in the convection-dominated regime.
- The estimate for the error in $L^\infty(L^2)$ is just a byproduct of the error analysis for a sum of (semi-) norms that includes, besides the error in the $L^\infty(L^2)$ norm, the error term of $\nabla(u - u_h)$ in $L^2(L^2)$ multiplied with $\nu^{1/2}$ and some error term that is related to the studied stabilization.
- The reduction of the convergence order is caused by estimating terms that appear already in the stabilized method for the steady-state problem. In fact, there is a principal open problem with respect to the steady-state case, which is formulated in [39], see also [17]: no stabilized discretization of steady-state convection–diffusion–reaction equations is known for which one can prove an optimal robust error estimate in $L^2(\Omega)$. For the SUPG method and P_1 finite elements in two dimensions, it is known that the order of convergence 1.5 for the error in $L^2(\Omega)$ is sharp, see [40], which is proved by considering special types of grids. However, on many other grids, second order convergence can be observed such that for the steady-state problem optimal convergence in $L^2(\Omega)$ might be possible to prove for certain finite element spaces and with some appropriate assumptions on the mesh. For the time-dependent problem, optimal order error reduction in $L^\infty(L^2)$ for the SUPG method and the P_1 finite element could be already observed in Example 3.4. However, for P_2 finite elements, even using a structured family of triangulations of Friedrichs–Keller type leads only to the order 2.5 predicted by Theorem 3.3 for $\nu = 10^{-6}$ and $\nu = 10^{-8}$, compare Fig. 4.

4. The incompressible Navier–Stokes equations

The first part of this section considers the Galerkin finite element discretization of the incompressible Navier–Stokes equations. A robust estimate for the $L^\infty(L^2)$ error of the velocity is derived whose order is $r - 1$. In Section 4.2, methods are surveyed for which robust estimates of order r can be proved. Then, Section 4.3 presents stabilized methods where robust error bounds of order $r + 1/2$ could be derived. To the best or our knowledge, the coverage of methods, for which provable robust estimates of order $r - 1$, r, or $r + 1/2$, respectively, can be found in the literature so far, is complete. As it is already noticed in the introduction, it is an open problem whether or not there are methods which allow robust estimates of optimal order $r + 1$. Finally, Section 4.4 presents briefly methods that behave in a robust way but for which only estimates are derived so far which are not robust.

4.1. The Galerkin method

Section 4.1.1 reviews results from the classical finite element convergence theory. In this theory, optimal order of convergence of the $L^\infty(L^2)$ velocity error can be proved, however, the error bounds are not robust. Then, in Section 4.1.2, the role of the continuity equation in the finite element error analysis and its impact on the convergence order is highlighted. Considering a linearized flow equation and assuming in $L^2(\Omega)$ a best approximation error for the velocity of order $r + 1$, for certain inf–sup stable pairs of finite element spaces, including the popular Taylor–Hood pairs, only order $r - 1$ can be proved for the velocity error in $L^\infty(L^2)$. The presented numerical example demonstrates that robust estimates provide in fact correct information on the order of error reduction on coarse grids. Section 4.1.3 discusses the effect of the discretization of the nonlinear convective term on the stability of the discretization. It is shown that the so-called EMAC form of the discrete nonlinear term leads to a robust estimate of order $r - 1$.

4.1.1. Classical finite element convergence theory

The convergence theory of the Galerkin finite element method for the evolutionary Navier–Stokes equations (5) was developed in the seminal papers of Heywood and Rannacher [41–44]. In all of these papers, the viscosity coefficient is assumed to be $\nu = 1$. Hence, there is no need to track the dependency of the error bounds on this coefficient. Applying the analysis developed in [41–44] for small viscosity coefficients, it turns out that the corresponding bounds are not robust. We like to emphasize that the small viscosity case was not a concern of the authors of [41–44] in the development of their theory.

Let $\{V_h \subset V\}$ and $\{Q_h \subset Q\}$ be two families of finite element spaces that correspond to a family of partitions $\{\mathcal{T}_h\}$ of Ω. The papers [41–44] consider pairs of finite element spaces that satisfy the discrete inf–sup condition

$$\inf_{q_h \in Q_h} \sup_{v_h \in V_h} \frac{(\nabla \cdot v_h, q_h)}{\|\nabla v_h\|_0 \|q_h\|_0} \geq \beta_0 > 0, \tag{27}$$

where β_0 is independent of the mesh width h and the polynomial degree r.

In the first paper, [41], a bound of second order is proved for the $L^\infty(L^2)$ error of the velocity. This order is optimal for inf–sup stable pairs with velocity finite element spaces of first order, e.g., for the MINI element [45]. The constants in the error bound grow exponentially in time, in general, even in the case in which the regularity of the solution is uniform in time. For this reason, they are called local error estimates, regarding to their blow-up for $t \to \infty$. The error analysis in [41] is restricted to second order convergence, independently of the polynomial degree of the finite element spaces, because getting higher order error bounds would require the solution of the Navier–Stokes equations to satisfy non-local compatibility conditions at the initial time, which are not assumed.

The second paper [42] considers the case in which a stable solution is approximated in order to derive uniform bounds for the error constants as t goes to infinity. Several concepts of stability are introduced: exponential stability, quasi-exponential stability, and contractive stability to a tolerance. Error bounds of second order in space are proved with constants that remain bounded as $t \to \infty$ for exponentially stable solutions. Analogous bounds are also derived in the case of quasi-exponentially stable solutions with the difference that $u(t, x)$ in the error is replaced by $u(t + \beta_h(t), x + w_h(t, x))$, where w_h is a spatial rotation and β_h is a time shift with both time derivatives being of order h^2. Finally, for the case of a solution with a contractive stability to a tolerance, the discrete approximations are proven to remain within a tolerance neighborhood of the solution of the continuous equation (5) and, with appropriate additional assumptions, to inherit the property of contractive stability to a tolerance.

In the third paper, [43], optimal bounds of order up to five are proved for the $L^\infty(L^2)$ error of the velocity, without assuming local compatibility conditions. Thus, the presented analysis applies, e.g., to the Taylor–Hood pairs of finite element spaces P_r/P_{r-1}, $r \in \{2, 3, 4\}$, [46]. As in [41], the constants in the error bounds in [43] grow exponentially in time. In addition, they possess a factor $t^{1-r/2}$, r being the rate of convergence, so that except for the case $r = 2$, considered in [41], the bounds explode as t tends to zero. Under the assumption that the solution is exponentially stable, the exponentially growing behavior in time could be removed from the constants of the error bound. Whether the limitation to fifth order convergence in space can be avoided is left as an open question, although in the special case of weakly divergence-free elements, it is stated that the results could be extended to convergence order six up to a logarithmic term.

Finally, the fourth paper [44] provides an error analysis of a fully discrete method with the Crank–Nicolson scheme as time integrator. A second order in time error bound is proved for the $L^\infty(L^2)$ error of the velocity

between the fully discrete approximation and the continuous-in-time approximation. The constant in the error estimate behaves as t^{-1} as t goes to zero. The bounds are valid for $t \to \infty$ when an exponentially stable solution is approximated.

In summary, the classical finite element convergence theory of the Galerkin discretization of the evolutionary Navier–Stokes equations provides, with some restrictions as mentioned above, optimal estimates for the velocity error in $L^\infty(L^2)$. However, the error bounds are not robust. In the following, it will be shown that taking robustness into account leads to an order reduction of the error bounds and that, in fact, the reduced order can be observed for small viscosity coefficients on coarse grids, i.e., in the convection-dominated regime. Nevertheless, it would be interesting to study carefully the classical analysis keeping track of the dependency on ν^{-1}, and, in particular, to check if it is possible to obtain error bounds without exponential factor in the case of exponentially stable solutions.

4.1.2. On the impact of the continuity equation

This section studies the role of the continuity equation in the Navier–Stokes equations. To this end, the evolutionary Oseen equations are considered, which contain the continuity equation but there is no influence of a nonlinear convection term.

The time-dependent Oseen problem reads as follows

$$
\begin{aligned}
\partial_t \boldsymbol{u} - \nu \Delta \boldsymbol{u} + (\boldsymbol{b} \cdot \nabla)\boldsymbol{u} + \nabla p &= \boldsymbol{f} && \text{in } (0, T] \times \Omega, \\
\nabla \cdot \boldsymbol{u} &= 0 && \text{in } (0, T] \times \Omega, \\
\boldsymbol{u} &= \boldsymbol{0} && \text{on } (0, T] \times \partial\Omega, \\
\boldsymbol{u}(0, \cdot) &= \boldsymbol{u}_0(\cdot) && \text{in } \Omega,
\end{aligned}
\tag{28}
$$

where the functions have the same meaning as for the Navier–Stokes equations (1) and $\boldsymbol{b} : (0, T] \times \Omega \to \mathbb{R}^d$ is a solenoidal vector field, i.e., $\nabla \cdot \boldsymbol{b}(t, \boldsymbol{x}) = 0$ in Ω and for all times.

A weak formulation of problem (28) is: Find $(\boldsymbol{u}, p) : (0, T] \to V \times Q$ such that for all $(\boldsymbol{v}, q) \in V \times Q$,

$$
(\partial_t \boldsymbol{u}, \boldsymbol{v}) + \nu(\nabla \boldsymbol{u}, \nabla \boldsymbol{v}) + ((\boldsymbol{b} \cdot \nabla)\boldsymbol{u}, \boldsymbol{v}) - (\nabla \cdot \boldsymbol{v}, p) + (\nabla \cdot \boldsymbol{u}, q) = \langle \boldsymbol{f}, \boldsymbol{v} \rangle,
\tag{29}
$$

and $\boldsymbol{u}(0, \cdot) = \boldsymbol{u}_0(\cdot) \in V$. Note that from the assumption $\nabla \cdot \boldsymbol{b} = 0$, it follows that the convective term is skew-symmetric, i.e.,

$$
((\boldsymbol{b} \cdot \nabla)\boldsymbol{v}, \boldsymbol{w}) = -((\boldsymbol{b} \cdot \nabla)\boldsymbol{w}, \boldsymbol{v}) \quad \forall \, \boldsymbol{v}, \boldsymbol{w} \in V.
\tag{30}
$$

The space of weakly divergence-free functions is defined in (6) and the space of discretely divergence-free functions in (7). Note that for many pairs of finite element spaces, discretely divergence-free functions are in general not weakly divergence-free, i.e., $V_h^{\mathrm{div}} \not\subset V^{\mathrm{div}}$. This property causes a number of difficulties in the analysis and simulation of incompressible flow problems, compare [47]. In this section, only the situation $V_h^{\mathrm{div}} \not\subset V^{\mathrm{div}}$ is considered, for the other case see Section 4.2.1.

To fix ideas in the exposition, we will study the Taylor–Hood pair of mixed finite elements in which $V_h = V_{h,r}$ is the space of piecewise continuous polynomials of degree $r \geq 2$, satisfying homogeneous Dirichlet boundary conditions, and $Q_h = Q_{h,r-1}$ is the space of piecewise continuous polynomials of degree $r - 1$ with vanishing integral mean value. For the Taylor–Hood pairs, it holds that $V_h^{\mathrm{div}} \not\subset V^{\mathrm{div}}$. The Taylor–Hood pair with $r = 2$ is probably the most popular inf–sup stable pair of finite element spaces.

In the analysis, the $L^2(\Omega)$ projection $\Pi_h^{r-1} p$ of the pressure p in (28) onto $Q_{h,r-1}$ is used. It is known that the following estimate holds for $m \in \{0, 1\}$

$$
\| p - \Pi_h^{r-1} p \|_m \leq C h^{r-m} \| p \|_r \quad \forall \, p \in H^r(\Omega).
\tag{31}
$$

Following [48], a modified Stokes projection $\boldsymbol{s}_h : V \to V_h^{\mathrm{div}}$ is considered for the velocity, satisfying

$$
(\nabla \boldsymbol{s}_h, \nabla \boldsymbol{v}_h) = (\nabla \boldsymbol{u}, \nabla \boldsymbol{v}_h) \quad \forall \, \boldsymbol{v}_h \in V_h^{\mathrm{div}}.
\tag{32}
$$

The following bound holds for $m \in \{0, 1\}$, compare [48],

$$
\| \boldsymbol{u} - \boldsymbol{s}_h \|_m \leq C h^{r+1-m} \| \boldsymbol{u} \|_{r+1} \quad \forall \, \boldsymbol{u} \in V \cap H^{r+1}(\Omega)^d.
\tag{33}
$$

It holds, see [49, (3.32)], [50, (21)], [51], that

$$
\| \boldsymbol{s}_h \|_{L^\infty} \leq C \left(\| \boldsymbol{u} \|_{d-2} \| \boldsymbol{u} \|_2 \right)^{1/2}, \quad \| \nabla \boldsymbol{s}_h \|_{L^\infty} \leq C \| \nabla \boldsymbol{u} \|_{L^\infty}.
\tag{34}
$$

Using s_h as test function in (32) gives $\|\nabla s_h\|_0 \le \|\nabla u\|_0$. Then, one can conclude with (34) and the Riesz–Thorin interpolation theorem that

$$\|\nabla s_h\|_{L^p} \le C\|\nabla u\|_{L^p} \quad \text{for } p \in (2, \infty). \tag{35}$$

Assuming the necessary smoothness in time, a Stokes projection like (32) can be defined for $\partial_t u$. Then, an error bound of form (33) can be derived also for $\partial_t(u - s_h)$.

In this section, the Galerkin method for the Oseen equations (29) without any added stabilization is analyzed. Let $u_h : (0, T] \to V_{h,r}$, $p_h : (0, T] \to Q_{h,r-1}$ be the Galerkin approximation satisfying

$$(\partial_t u_h, v_h) + \nu(\nabla u_h, v_h) + ((b \cdot \nabla)u_h, v_h) - (\nabla \cdot v_h, p_h) = \langle f, v_h \rangle \quad \forall\, v_h \in V_{h,r},$$
$$(\nabla \cdot u_h, q_h) = 0 \quad \forall\, q_h \in Q_{h,r-1}. \tag{36}$$

Since the pair $V_{h,r}/Q_{h,r-1}$ satisfies the discrete inf–sup condition (27), this problem is well posed. The following theorem states an error bound for the Galerkin approximation to the velocity in $L^\infty(L^2)$ in which the constant in the error bound does not depend on inverse powers of the viscosity.

Theorem 4.1 (*Error Estimate for the Galerkin Method*). *Let (u, p) be the sufficiently smooth solution of (29) and assume $b \in L^\infty(L^\infty)$. Let u_h be the velocity finite element approximation defined in (36). Then, the following bound holds for $0 < t \le T$*

$$\|u(t) - u_h(t)\|_0 \le \|u_h(0) - s_h(0)\|_0 + Cth^{r-1}M_{u,p} + Ch^{r+1}\|u\|_{L^\infty(H^{r+1})},$$

where

$$M_{u,p} = h\left(\|\partial_t u\|_{L^\infty(H^r)} + \|b\|_{L^\infty(L^\infty)}\|u\|_{L^\infty(H^{r+1})}\right) + \|p\|_{L^\infty(H^r)}. \tag{37}$$

Proof. The proof is similar to the proof of Theorem 3.1. The error is split into an interpolation error between u and its modified Stokes projection s_h and the discrete remainder $\phi_h = u_h - s_h \in V_h^{\text{div}}$. A straightforward calculation, taking into account the definition of the Stokes projection, leads to the error equation

$$(\partial_t \phi_h, v_h) + \nu(\nabla \phi_h, \nabla v_h) + ((b \cdot \nabla)\phi_h, v_h) = (\tau_1, v_h) + (\nabla \cdot v_h, \tau_2) \quad \forall\, v_h \in V_h^{\text{div}}, \tag{38}$$

where

$$\tau_1 = (\partial_t u - \partial_t s_h) + ((b \cdot \nabla)(u - s_h)), \quad \tau_2 = p - \Pi_h^{r-1}p.$$

Taking $v_h = \phi_h$ in (38), neglecting a non-negative term on the left-hand side, and using the skew-symmetry (30), which gives in particular $((b \cdot \nabla)\phi_h, \phi_h) = 0$, yields

$$\|\phi_h\|_0 \frac{d}{dt}\|\phi_h\|_0 \le (\tau_1, \phi_h) + (\nabla \cdot \phi_h, \tau_2). \tag{39}$$

This inequality does not contain the viscosity ν any more, hence the resulting estimate will not depend explicitly on ν. The key step for estimating the last term consists in utilizing integration by parts

$$(\nabla \cdot \phi_h, \tau_2) = (\nabla \cdot \phi_h, p - \Pi_h^{r-1}p) = -\left(\phi_h, \nabla(p - \Pi_h^{r-1}p)\right). \tag{40}$$

Now, the estimate of the right-hand side of (39) continues by applying the Cauchy–Schwarz inequality, (33), and (31) leading to

$$(\tau_1, \phi_h) + (\nabla \cdot \phi_h, \tau_2) \le CM_{u,p}h^{r-1}\|\phi_h\|_0,$$

where $M_{u,p}$ is the constant in (37). To conclude proof, one argues as in the proof of Theorem 3.1, applies the triangle inequality and again (33). $\quad\square$

In view of Theorem 4.1, one observes that for the Galerkin approximation to the velocity in the Oseen equations, the suboptimal order of convergence $r - 1$ is proved with constants independent of inverse powers of ν while, as Theorem 3.1 shows, for the convection–reaction–diffusion model an order of convergence r can be proved. The reduction of one order comes from the error bound of the second term in (39), which is the truncation error of the pressure, see (40). Alternatively, an inverse estimate could be used, which leads finally to the same error bound.

Table 2
Numerical examples for the Stokes and Navier–Stokes equations, number of degrees of freedom (including Dirichlet nodes) for inf–sup stable pairs of finite element spaces on triangular grids.

	P_2/P_1		P_2/P_1^{disk}, barycentric		P_3/P_2	
Level	Velocity	Pressure	Velocity	Pressure	Velocity	Pressure
2			210	144	170	41
3	290	41	802	576	626	145
4	1090	145	3138	2304	2402	545
5	4226	545	12418	9216	9410	2113
6	16642	2113	49410	36864	37250	8321
7	66050	8321				
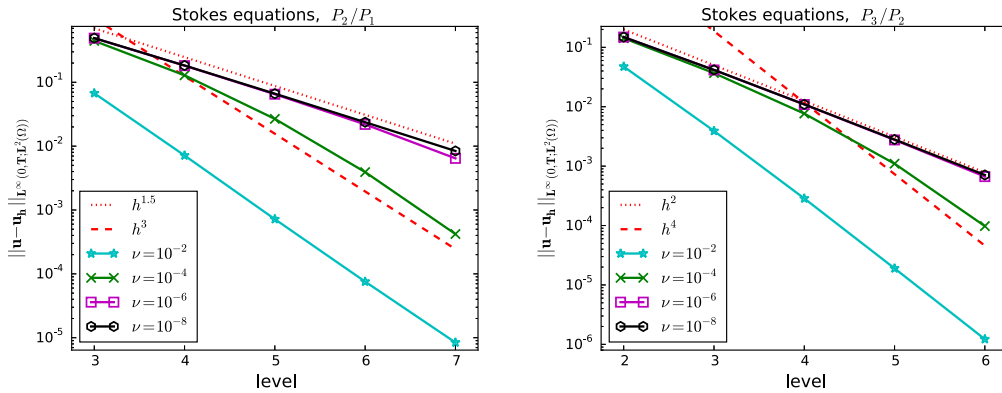
Fig. 5. Example 4.2 Galerkin method for the evolutionary Stokes equations, P_2/P_1 (left) and P_3/P_2 (right).

A first way to improve the situation is the use of inf–sup stable pairs of finite element spaces where the term $(\nabla \cdot \boldsymbol{\phi}_h, \tau_2)$ does not appear. This is the case if discretely divergence-free functions are weakly divergence-free, see Section 4.2.1. Another way consists in adding appropriate stabilization terms to the Galerkin method, see Section 4.2.2.

Example 4.2 (*Galerkin Discretization of the Evolutionary Stokes Equations*). The order of convergence predicted by Theorem 4.1 is determined by the pressure term in $M_{\boldsymbol{u},p}$, the last term in (37). As already mentioned, the term $(\nabla \cdot \boldsymbol{\phi}_h, \tau_2)$, a coupling term of the finite element velocity and pressure that does not vanish if the finite element velocity solution is not weakly divergence-free, is responsible for the appearance of the pressure in $M_{\boldsymbol{u},p}$. Dominating convection does not play any role in this respect. To support this observation, a numerical study for the evolutionary Stokes equations, i.e., for (28) with $\boldsymbol{b} = \boldsymbol{0}$, will be presented here.

The considered example is defined in $\Omega = (0, 1)^2$, the final time is $T = 2$, and the prescribed solution is given by

$$\boldsymbol{u}(t; x, y) = 2\pi \sin(\pi t) \begin{pmatrix} \sin^2(\pi x) \sin(\pi y) \cos(\pi y) \\ -\sin^2(\pi y) \sin(\pi x) \cos(\pi x) \end{pmatrix}, \quad p(t, x, y) = 20 \sin(\pi t) \left(x^2 y - \frac{1}{6} \right). \tag{41}$$

The same spatial meshes and the same temporal discretization as in Example 3.2 were used. For the spatial discretization, the Taylor–Hood pairs P_2/P_1 and P_3/P_2 of finite element spaces were utilized. The numbers of degrees of freedom are given in Table 2. A quadrature rule was applied that is exact for polynomials of degree eight on triangles. This order of the quadrature rule is not necessary for assembling the matrices but it keeps the quadrature error small for the calculation of the right-hand side and the $L^2(\Omega)$ error of the velocity. Simulations were performed for $\nu \in \{10^{-2}, 10^{-4}, 10^{-6}, 10^{-8}\}$.

Fig. 5 presents the numerical results. There are two important observations. First, if the mesh is sufficiently fine with respect to the viscosity coefficient, one can see numerically the optimal order of convergence, as predicted by the classical finite element convergence theory. But second, in the other situation, there is a clear decrease of

the order of error reduction for $\nu \leq 10^{-4}$. Note that coarse meshes with respect to the viscosity coefficient are the typical situation for high Reynolds number flows. In this case, the order of error reduction is described obviously better by robust error bounds. From the results for the P_3/P_2 pair of spaces, one can conclude that the order $r - 1$ is a sharp prediction. $\quad\square$

4.1.3. On the impact of the discretization of the nonlinear term

In this section, the Galerkin approximation to the Navier–Stokes equations (5) is studied. Again, the situation $V_h^{\mathrm{div}} \not\subset V^{\mathrm{div}}$ will be considered and for fixing ideas, as in previous section, the analysis will be presented for the inf–sup stable Taylor–Hood pair of mixed finite element spaces.

Given $\boldsymbol{f} \in L^2(\Omega)^d$, let $\boldsymbol{u}_h \; : \; (0, T] \to V_{h,r}$, $p_h \; : \; (0, T] \to Q_{h,r-1}$ be the solution of the following Galerkin approximation of (5)

$$(\partial_t \boldsymbol{u}_h, \boldsymbol{v}_h) + \nu(\nabla \boldsymbol{u}_h, \nabla \boldsymbol{v}_h) + ((\boldsymbol{u}_h \cdot \nabla)\boldsymbol{u}_h, \boldsymbol{v}_h) - (\nabla \cdot \boldsymbol{v}_h, p_h) = (\boldsymbol{f}, \boldsymbol{v}_h) \quad \forall \, \boldsymbol{v}_h \in V_{h,r},$$
$$(\nabla \cdot \boldsymbol{u}_h, q_h) = 0 \quad \forall \, q_h \in Q_{h,r-1}. \tag{42}$$

The nonlinear term in (42) is called convective form of the nonlinear term in [2, Chapter 6]. It is not skew-symmetric for Taylor–Hood pairs of finite element spaces. For $(\boldsymbol{u}_h, p_h)$ satisfying (42), one cannot even prove boundedness of the approximations, since in the stability analysis, the term $((\boldsymbol{u}_h \cdot \nabla)\boldsymbol{u}_h, \boldsymbol{u}_h)$ cannot be bounded.

Next, in addition to (42), a Galerkin approximation with a skew-symmetric bilinear form of the nonlinear term is considered: Find $\boldsymbol{u}_h \; : \; (0, T] \to V_{h,r}$, $p_h \; : \; (0, T] \to Q_{h,r-1}$ such that

$$(\partial_t \boldsymbol{u}_h, \boldsymbol{v}_h) + \nu(\nabla \boldsymbol{u}_h, \nabla \boldsymbol{v}_h) + b(\boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{v}_h) - (\nabla \cdot \boldsymbol{v}_h, p_h) = (\boldsymbol{f}, \boldsymbol{v}_h) \quad \forall \, \boldsymbol{v}_h \in V_{h,r},$$
$$(\nabla \cdot \boldsymbol{u}_h, q_h) = 0 \quad \forall \, q_h \in Q_{h,r-1}, \tag{43}$$

where, the so-called divergence form of the nonlinear term is given by

$$b(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = ((\boldsymbol{u} \cdot \nabla)\boldsymbol{v}, \boldsymbol{w}) + \frac{1}{2}((\nabla \cdot \boldsymbol{u})\boldsymbol{v}, \boldsymbol{w}) \quad \forall \, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V. \tag{44}$$

Notice the well-known property

$$b(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = -b(\boldsymbol{u}, \boldsymbol{w}, \boldsymbol{v}) \quad \forall \, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V, \tag{45}$$

such that, in particular, $b(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{v}) = 0$ for all $\boldsymbol{u}, \boldsymbol{v} \in V$. There are several ways for estimating the convective term (44) in the error analysis, depending on the smoothness assumptions on its arguments. Assuming somewhat more than the minimal regularity, it holds that

$$b(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) \leq \|\boldsymbol{u}\|_0 \|\nabla \boldsymbol{v}\|_{L^\infty} \|\boldsymbol{w}\|_0 + \frac{1}{2}\|\nabla \cdot \boldsymbol{u}\|_0 \|\boldsymbol{v}\|_{L^\infty} \|\boldsymbol{w}\|_0 \quad \forall \, \boldsymbol{u}, \boldsymbol{w} \in V, \boldsymbol{v} \in W^{1,\infty}(\Omega)^d. \tag{46}$$

A standard stability analysis proves that the velocity approximation using the skew-symmetric form (44) is bounded in $L^\infty(L^2)$ with a constant that does not depend on inverse powers of ν.

Theorem 4.3 (*Stability for Skew-Symmetric Form of the Nonlinear Term*). *Let $\boldsymbol{u}_h$ be the velocity approximation defined in* (43). *Then, the following stability estimate is valid for all $t \in [0, T]$*

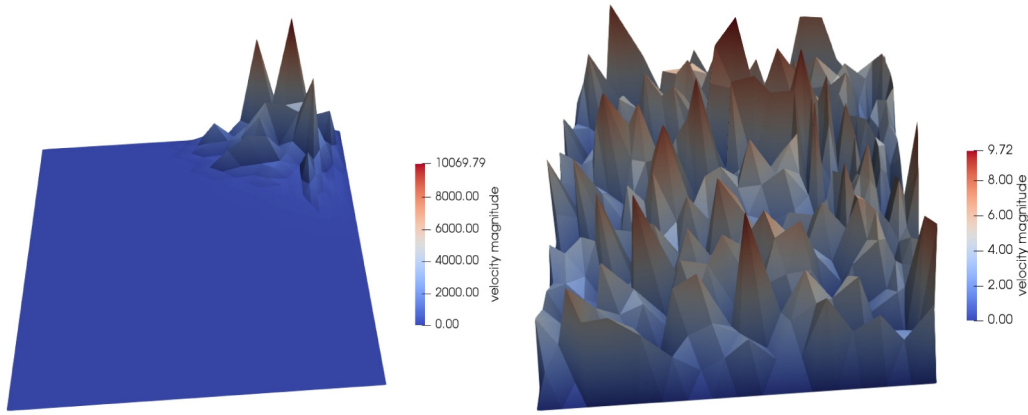$$\|\boldsymbol{u}_h(t)\|_0 \leq \|\boldsymbol{u}_h(0)\|_0 + \int_0^t \|\boldsymbol{f}(s)\|_0 \, ds. \tag{47}$$

Proof. From the theory of the Galerkin method developed in [3], it is known that (43) has a unique velocity solution $\boldsymbol{u}_h$. Taking $\boldsymbol{v}_h = \boldsymbol{u}_h$ in (43), using the skew-symmetric property (45) and the Cauchy–Schwarz inequality gives

$$\|\boldsymbol{u}_h\|_0 \frac{d}{dt}\|\boldsymbol{u}_h\|_0 + \nu\|\nabla \boldsymbol{u}_h\|_0^2 = \frac{1}{2}\frac{d}{dt}\|\boldsymbol{u}_h\|_0^2 + \nu\|\nabla \boldsymbol{u}_h\|_0^2 \leq \|\boldsymbol{f}\|_0\|\boldsymbol{u}_h\|_0.$$

Neglecting the second term on the left-hand side to get rid of ν, dividing by $\|\boldsymbol{u}_h\|_0$ and integrating in $(0, T)$ lead to (47). $\quad\square$

Despite the stability, it is not possible to derive error bounds with constants that are independent of inverse powers of the viscosity. The important terms of the error equation are

$$\frac{1}{2}\frac{d}{dt}\|\boldsymbol{\phi}_h\|_0^2 + \nu\|\nabla\boldsymbol{\phi}_h\|_0^2 = -b(\boldsymbol{\phi}_h, \boldsymbol{s}_h, \boldsymbol{\phi}_h) - b(\boldsymbol{u}_h, \boldsymbol{\phi}_h, \boldsymbol{\phi}_h) + \cdots = -b(\boldsymbol{\phi}_h, \boldsymbol{s}_h, \boldsymbol{\phi}_h) + \cdots, \tag{48}$$

Fig. 6. Example 4.4 Galerkin method for the Navier–Stokes equations, $\nu = 10^{-6}$, level 5, solution at $t = 1.3409$, left: convective form of the convective term (42), right: divergence form of the convective term (44); note the different scalings.

where $\boldsymbol{\phi}_h = \boldsymbol{u}_h - \boldsymbol{s}_h$, with $\boldsymbol{s}_h$ being the Stokes projection defined in (32). The second term on the right-hand side of (48) vanishes due to (45). Then, using the estimate (46) for the convective term, Young's inequality, and bounding the norm of the divergence by the norm of the gradient yields

$$\frac{1}{2}\frac{d}{dt}\|\boldsymbol{\phi}_h\|_0^2 + \nu\|\nabla\boldsymbol{\phi}_h\|_0^2 \leq \|\nabla\boldsymbol{s}_h\|_{L^\infty}\|\boldsymbol{\phi}_h\|_0^2 + \frac{1}{2}\|\boldsymbol{s}_h\|_{L^\infty}\|\nabla\cdot\boldsymbol{\phi}_h\|_0\|\boldsymbol{\phi}_h\|_0 + \cdots$$

$$\leq \frac{\nu}{2}\|\nabla\boldsymbol{\phi}_h\|_0^2 + \left(\|\nabla\boldsymbol{s}_h\|_{L^\infty} + \frac{1}{8\nu}\|\boldsymbol{s}_h\|_{L^\infty}^2\right)\|\boldsymbol{\phi}_h\|_0^2 + \cdots . \tag{49}$$

Hence, the application of Gronwall's lemma leads to the appearance of ν^{-1} in the exponential factor. Note that with minimal regularity assumptions on $\boldsymbol{u}$, one even gets ν^{-3} in the exponential factor, e.g., see [2, Theorem 7.35]. An alternative estimate that can be considered uses the inverse inequality (8)

$$\|\nabla\cdot\boldsymbol{\phi}_h\|_0 \leq \|\nabla\boldsymbol{\phi}_h\|_0 \leq \frac{c_{\mathrm{inv}}}{h}\|\boldsymbol{\phi}_h\|_0,$$

which, however, leads finally to an inverse power of the mesh width in the exponential factor of the error bound.

Example 4.4 (*Galerkin Method: Impact of the Discretization of the Nonlinear Term: Convective and Divergence Form*). For this example, in principle the same configuration is considered as in Example 4.2. Only, the right-hand side of the problem is different since the prescribed solution (41) is now inserted in the evolutionary Navier–Stokes equations (1) instead of the evolutionary Stokes equations. In each time instant, a nonlinear problem has to be solved. For this purpose, a standard fixed point iteration (Picard iteration) was employed. This iteration was stopped if the Euclidean norm of the residual vector was smaller than 10^{-9}.

Simulations were performed again on levels 3 to 7 and for $\nu \in \{10^{-4}, 10^{-6}, 10^{-8}\}$. Using the convective form of the nonlinear term (42), usually the numerical simulations blew up. Only if the spatial resolution was sufficiently fine with respect to the viscosity, there was no blow-up: here for $\nu = 10^{-4}$ on levels 5–7 and for $\nu = 10^{-6}$ on level 7.

In contrast, we could not observe any blow-up if the divergence form of the convective term (43) was applied. Fig. 6 presents a comparison of numerical solutions obtained with both discretizations of the convective term, shortly before the blow-up of the simulation with the convective form of the convective term. Even if we could not observe a blow-up if the divergence form was used, the numerical solution is globally polluted with spurious oscillations. The errors of the velocity in $L^\infty(L^2)$ for the simulations with the divergence form are presented in Fig. 7. One can observe that the errors are quite large, of order 1, which fits the theory since both $\|\boldsymbol{u}(t)\|_0$ and $\|\boldsymbol{u}_h(t)\|_0$ are bounded for a finite time interval, Theorem 4.3, so that also the maximal $L^2(\Omega)$ error is bounded. Apart of this fact, the Galerkin method with the divergence form of the nonlinear term cannot be used without stabilization in the convection-dominated regime on coarse grids, in view of Figs. 6 and 7, where again coarse is to be understood with respect to the size of the viscosity coefficient. $\square$
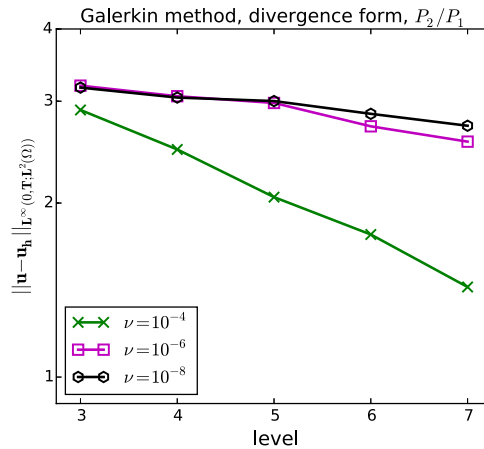
Fig. 7. Example 4.4 Galerkin method for the Navier–Stokes equations, P_2/P_1, divergence form of the convective term.

Recently, a new discrete form of the nonlinear term has been proposed in [52], the so-called EMAC form

$$c(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = (2\mathbb{D}(\boldsymbol{u})\boldsymbol{v}, \boldsymbol{w}) + ((\nabla \cdot \boldsymbol{u})\boldsymbol{v}, \boldsymbol{w}),$$

where $\mathbb{D}(\boldsymbol{u}) = (\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^T)/2$ is the velocity deformation tensor. The EMAC form is the only discrete form that conserves kinetic energy (for $\nu = 0, \boldsymbol{f} = \boldsymbol{0}$), momentum (for $\boldsymbol{f}$ with zero linear momentum), and angular momentum (for $\boldsymbol{f}$ with zero angular momentum) in the case that $\boldsymbol{u}$ is not weakly divergence-free. Using the EMAC form requires a modification of the pressure. In [53], the Galerkin discretization of the Navier–Stokes equations (5) with the EMAC form of the nonlinear term is analyzed. It is shown that under the assumption of a sufficiently smooth solution, the nonlinear term can be bounded in a robust way. Altogether, a non-robust estimate of order r is presented in [53], where the dependency on ν^{-1} appears in the pressure term of the error bound. However, it is possible to derive a robust estimate, but only of order $r - 1$. The numerical study in Example 4.6 will demonstrate that this estimate is sharp.

Theorem 4.5 (*Error Estimate for the Galerkin Method with EMAC Form of the Nonlinear Term*)**.** *Let the solution of the Navier–Stokes equations (5) be sufficiently smooth and define*

$$L(T) = C \int_0^T \|\nabla \boldsymbol{u}(s)\|_{L^\infty} \, ds,$$

$$M_{\boldsymbol{u}, p} = C \left(\int_0^T \left(h \|\partial_t \boldsymbol{u}(s)\|_r + h \left(\|\boldsymbol{u}(s)\|_1 \|\boldsymbol{u}(s)\|_2 \right)^{1/2} \|\boldsymbol{u}(s)\|_{r+1} + \|p(s)\|_r \right) \, ds \right).$$

Then, the following error estimate holds for the Galerkin method with EMAC form of the nonlinear term

$$\|\boldsymbol{u}(t) - \boldsymbol{u}_h(t)\|_0 \le e^{L(T)} \left(\|\boldsymbol{u}_h(0) - \boldsymbol{s}_h(0)\|_0 + M_{\boldsymbol{u}, p} h^{r-1} \right) + C h^{r+1} \|\boldsymbol{u}\|_{L^\infty(H^{r+1})},$$

with $\boldsymbol{s}_h$ being the Stokes projection. The constant $M_{\boldsymbol{u}, p}$ is dimensionally correct in three dimensions. For $M_{\boldsymbol{u}, p}$ to be dimensionally correct in two-dimensional problems, the term $(\|\boldsymbol{u}(s)\|_1 \|\boldsymbol{u}(s)\|_2)^{1/2}$ in $M_{\boldsymbol{u}, p}$ must be replaced by

$$(h \|\boldsymbol{u}(s)\|_1 \|\boldsymbol{u}(s)\|_2)^{1/2} + (\|\boldsymbol{u}(s)\|_0 \|\boldsymbol{u}(s)\|_2)^{1/2}. \tag{50}$$

Proof (*Sketch*)**.** The error is decomposed in the form $\boldsymbol{u} - \boldsymbol{u}_h = (\boldsymbol{u} - \boldsymbol{s}_h) - (\boldsymbol{u}_h - \boldsymbol{s}_h) = \boldsymbol{\eta} - \boldsymbol{\phi}_h$, where $\boldsymbol{s}_h$ is the Stokes projection defined in (32). Then, deriving the error equation in the usual way and using $\boldsymbol{\phi}_h$ as test function lead to

$$\|\boldsymbol{\phi}_h\|_0 \frac{d}{dt} \|\boldsymbol{\phi}_h\|_0 \le \left| c(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{\phi}_h) - c(\boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{\phi}_h) \right| + |(\nabla \cdot \boldsymbol{\phi}_h, P - q_h)| + \cdots \tag{51}$$

for arbitrary $q_h \in Q_{h, r-1}$.

The estimate of the difference of the convective terms in (51) follows [53]. A direct calculation yields

$$c(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{\phi}_h) - c(\boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{\phi}_h)$$

$$= c(\boldsymbol{\eta}, \boldsymbol{u}, \boldsymbol{\phi}_h) + c(\boldsymbol{s}_h, \boldsymbol{\eta}, \boldsymbol{\phi}_h) - c(\boldsymbol{s}_h, \boldsymbol{\phi}_h, \boldsymbol{\phi}_h) - c(\boldsymbol{\phi}_h, \boldsymbol{s}_h, \boldsymbol{\phi}_h) - c(\boldsymbol{\phi}_h, \boldsymbol{\phi}_h, \boldsymbol{\phi}_h). \tag{52}$$

The last term on the right-hand side vanishes due to a property of the EMAC form. Now, a part of the fourth term on the right-hand side is reformulated. Using the definition of the deformation tensor, integration by parts, the skew symmetry of the divergence form of the nonlinear term, and again the definition of the deformation tensor yields

$$2(\mathbb{D}\left(\boldsymbol{\phi}_h\right)\boldsymbol{s}_h, \boldsymbol{\phi}_h)$$
$$= ((\nabla\boldsymbol{\phi}_h)\boldsymbol{s}_h, \boldsymbol{\phi}_h) + (\boldsymbol{s}_h, (\nabla\boldsymbol{\phi}_h)\boldsymbol{\phi}_h)$$
$$= ((\boldsymbol{s}_h \cdot \nabla)\boldsymbol{\phi}_h, \boldsymbol{\phi}_h) + ((\boldsymbol{\phi}_h \cdot \nabla)\boldsymbol{\phi}_h, \boldsymbol{s}_h)$$
$$= -((\nabla \cdot \boldsymbol{s}_h)\boldsymbol{\phi}_h, \boldsymbol{\phi}_h) - ((\boldsymbol{s}_h \cdot \nabla)\boldsymbol{\phi}_h, \boldsymbol{\phi}_h) - ((\nabla \cdot \boldsymbol{\phi}_h)\boldsymbol{\phi}_h, \boldsymbol{s}_h) - ((\boldsymbol{\phi}_h \cdot \nabla)\boldsymbol{s}_h, \boldsymbol{\phi}_h)$$
$$= -\frac{1}{2}((\nabla \cdot \boldsymbol{s}_h)\boldsymbol{\phi}_h, \boldsymbol{\phi}_h) - ((\nabla \cdot \boldsymbol{\phi}_h)\boldsymbol{\phi}_h, \boldsymbol{s}_h) - ((\boldsymbol{\phi}_h \cdot \nabla)\boldsymbol{s}_h, \boldsymbol{\phi}_h)$$
$$= -\frac{1}{2}((\nabla \cdot \boldsymbol{s}_h)\boldsymbol{\phi}_h, \boldsymbol{\phi}_h) - ((\nabla \cdot \boldsymbol{\phi}_h)\boldsymbol{\phi}_h, \boldsymbol{s}_h) - (\mathbb{D}\left(\boldsymbol{s}_h\right)\boldsymbol{\phi}_h, \boldsymbol{\phi}_h).$$

Inserting this expression in (52) and collecting terms give

$$c(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{\phi}_h) - c(\boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{\phi}_h) = c(\boldsymbol{s}_h, \boldsymbol{\eta}, \boldsymbol{\phi}_h) + c(\boldsymbol{\eta}, \boldsymbol{u}, \boldsymbol{\phi}_h) - \frac{1}{2}((\nabla \cdot \boldsymbol{s}_h)\boldsymbol{\phi}_h, \boldsymbol{\phi}_h) - (\mathbb{D}\left(\boldsymbol{s}_h\right)\boldsymbol{\phi}_h, \boldsymbol{\phi}_h). \tag{53}$$

Applying Hölder's inequality, the stability (35) of the Stokes projection, a Sobolev embedding, and an interpolation in Sobolev spaces yields for the first term

$$c(\boldsymbol{s}_h, \boldsymbol{\eta}, \boldsymbol{\phi}_h) = (2\mathbb{D}\left(\boldsymbol{s}_h\right)\boldsymbol{\eta}, \boldsymbol{\phi}_h) + ((\nabla \cdot \boldsymbol{s}_h)\boldsymbol{\eta}, \boldsymbol{\phi}_h)$$
$$\leq 2\|\mathbb{D}\left(\boldsymbol{s}_h\right)\|_{L^3}\|\boldsymbol{\eta}\|_{L^6}\|\boldsymbol{\phi}_h\|_0 + \|\nabla \cdot \boldsymbol{s}_h\|_{L^3}\|\boldsymbol{\eta}\|_{L^6}\|\boldsymbol{\phi}_h\|_0$$
$$\leq C\|\nabla\boldsymbol{s}_h\|_{L^3}\|\boldsymbol{\eta}\|_{L^6}\|\boldsymbol{\phi}_h\|_0 \leq C\|\nabla\boldsymbol{s}_h\|_{L^3}\|\boldsymbol{\eta}\|_1\|\boldsymbol{\phi}_h\|_0 \leq C\|\nabla\boldsymbol{u}\|_{1/2}\|\boldsymbol{\eta}\|_1\|\boldsymbol{\phi}_h\|_0$$
$$\leq C\left(\|\boldsymbol{u}\|_1\|\boldsymbol{u}\|_2\right)^{1/2}\|\boldsymbol{\eta}\|_1\|\boldsymbol{\phi}_h\|_0.$$

To estimate the second term, Hölder's and Agmon's inequality are used to get

$$c(\boldsymbol{\eta}, \boldsymbol{u}, \boldsymbol{\phi}_h) \leq C\|\nabla\boldsymbol{\eta}\|_0\|\boldsymbol{u}\|_{L^\infty}\|\boldsymbol{\phi}_h\|_0 \leq C\left(\|\boldsymbol{u}\|_1\|\boldsymbol{u}\|_2\right)^{1/2}\|\boldsymbol{\eta}\|_1\|\boldsymbol{\phi}_h\|_0.$$

The third and the fourth term in (53) are bounded just with Hölder's inequality. One observes that there is no inverse power of the viscosity in the error bound of the discrete convective term.

The arguments above are dimensionally correct in three dimensions. In two dimensions, both $\| \cdot \|_{L^3}$ and $\| \cdot \|_{L^6}$ should be replaced by $\| \cdot \|_{L^4}$, and arguing similarly, one concludes that the term $(\|\boldsymbol{u}(s)\|_1 \| \boldsymbol{u}(s)\|_2)^{1/2}$ in $M_{\boldsymbol{u},p}$ should be replaced by (50).

The pressure term on the right-hand side of (51) can be bounded robustly in the same way as in the proof of Theorem 4.1, which gives a term of order $r - 1$. □

Example 4.6 (*Galerkin Method: EMAC Form of the Nonlinear Term*). The same problem as in Example 4.4 is considered, with the same setup. Only the EMAC form was used as discrete nonlinear form. The results are presented in Fig. 8.

Comparing with Fig. 7, it can be seen that there is a much better error reduction if the EMAC form is used instead of the divergence form. For small viscosity coefficients and on coarse grids, one can observe for the P_2/P_1 pair of spaces an order of error reduction of 1.5, which is the same observation as for the evolutionary Stokes problem in Example 4.2. In case of the P_3/P_2 pair of spaces, the order of error reduction is 2, which is exactly the order that is proposed by the robust error bound from Theorem 4.5. □

Theorem 4.5 and Example 4.6 allow to address another open problem stated in [17]: *From the theoretical point of view, the new EMAC form offers several attractive features. Further numerical studies are necessary to clarify whether this form should be preferred also in simulations.* With respect to robust velocity estimates in $L^\infty(L^2)$ for the evolutionary Navier–Stokes equations, the answer is twofold. On the one hand, for the Galerkin method and non weakly divergence-free and inf–sup stable pairs of finite element spaces, the EMAC form outperforms all other
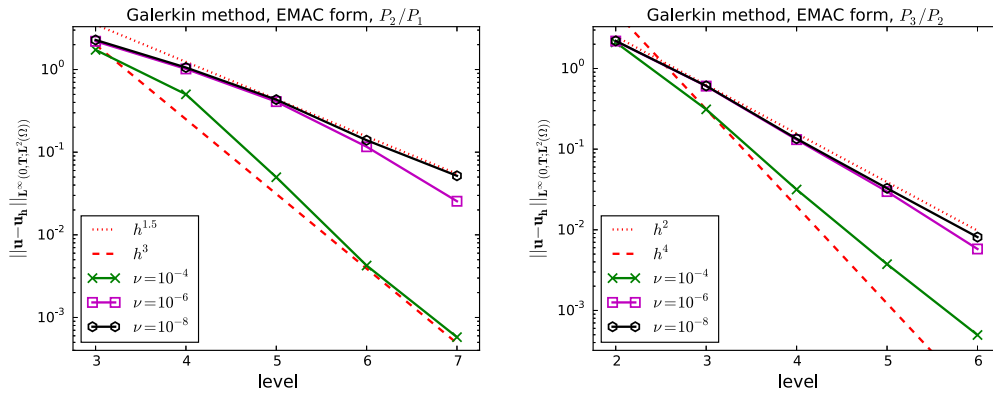
Fig. 8. Example 4.6 Galerkin method for the Navier–Stokes equations, EMAC form of the discrete convective term.

discrete nonlinear forms. It is the only form for which a robust estimate is possible, which is of order $r - 1$. But on the other hand, as it will be shown in Section 4.2, using weakly divergence-free pairs of spaces or applying a rather simple stabilization (grad–div stabilization) for non weakly divergence-free pairs with arbitrary form of the discrete convective term leads to robust estimates of order r, i.e., one order higher.

4.2. Methods of order r in $L^\infty(L^2)$

This section presents discretizations whose order of robust error reduction for the velocity error in $L^\infty(L^2)$ is r. First, the Galerkin discretization with inf–sup stable pairs of finite element spaces and weakly divergence-free velocity is studied, in Section 4.2.1. Second, Section 4.2.2 considers non weakly divergence-free and inf–sup stable pairs of finite element spaces, where the Galerkin method is augmented either with a grad–div stabilization or with a LPS stabilization for the velocity gradient. The final part, Section 4.2.3, discusses three methods for pairs of finite element spaces that are not inf–sup stable and thus require a pressure stabilization. Two of these methods achieve stability with respect to dominating convection with the same mechanisms as the methods from Section 4.2.2. For the last method, it is shown that an appropriate pressure stabilization might be sufficient for stabilizing also dominant convection.

4.2.1. Inf–sup stable and weakly divergence-free pairs of finite element spaces

This section considers the Galerkin method (42) for conforming inf–sup stable pairs of finite element spaces with $V_h^{\mathrm{div}} \subset V^{\mathrm{div}}$. There are a number of pairs that satisfy this property. The most popular one is the Scott–Vogelius pair [54], $P_r/P_{r-1}^{\mathrm{disk}}$, $r \geq d$. However, the discrete inf–sup condition for the case $r = d$ can be proved only for special grids, so-called barycentric-refined grids, see [55,56]. Barycentric-refined grids are constructed from a given simplicial grid by connecting the vertices of each mesh cell with the barycenter of that mesh cell. Notice that although $V_h^{\mathrm{div}} \subset V^{\mathrm{div}}$, one should not use the notion that the approximation $\boldsymbol{u}_h$ is pointwise divergence-free, since $\nabla \cdot \boldsymbol{u}_h$ is not continuous across boundaries of mesh cells.

The use of weakly divergence-free pairs of finite element spaces is attractive from the point of view of the physical consistency of numerical solutions. First, there is an exact conservation of mass in the sense that $\|\nabla \cdot \boldsymbol{u}_h(t)\|_0 = 0$ for all times. Second, provided that possible consistency errors do not depend on the pressure, the numerical solution satisfies a so-called fundamental invariance property: gradient forces on the right-hand side possess an impact only on the discrete pressure. Closely connected to this property is the so-called pressure robustness of the method, i.e., error bounds for the velocity do not depend on the pressure, compare (58). Thus, possibly large norms of the pressure are not of importance for velocity errors. All these properties are in contrast to pairs of finite element spaces that are not weakly divergence-free, see [47] for a comprehensive discussion of this topic.

If $V_h^{\mathrm{div}} \subset V^{\mathrm{div}}$, the finite element velocity solution $\boldsymbol{u}_h$ is weakly divergence-free, hence the second term of the divergence form (44) of the nonlinear term vanishes for $\boldsymbol{u}_h$ and the convective form is skew-symmetric. Consequently, one can use the convective form of the convective term in the numerical analysis as well as

in simulations. There are less terms to be estimated with the convective form compared with the divergence form.

The error analysis of the Galerkin method can be found in [57]. Error bounds of order r are proved for the $L^\infty(L^2)$ norm of the velocity. Here, a simplified proof is presented. In this proof, the velocity approximation is compared with the Stokes projection s_h defined in (32) and the decomposition $u - u_h = (u - s_h) - (u_h - s_h) = (u - s_h) - \phi_h$ is utilized. From the error equation, one obtains by using the definition of the Stokes projection and that ϕ_h is weakly divergence-free

$$\|\phi_h\|_0 \frac{d}{dt}\|\phi_h\|_0 \leq \left|\left((u_h \cdot \nabla)u_h, \phi_h\right) - \left((s_h \cdot \nabla)s_h, \phi_h\right)\right| + \left|(\tau_1, \phi_h)\right|, \tag{54}$$

where $\tau_1 = (\partial_t u - \partial_t s_h) + ((u \cdot \nabla)u - (s_h \cdot \nabla)s_h)$. Note that in contrast to the case where the finite element velocity is not weakly divergence-free, there is no pressure contribution in the error equation, i.e., a term like the last term on the right-hand side of (39) does not appear. It was shown in Section 4.1.2 that exactly this term is responsible for obtaining only order $r - 1$ for the velocity error in $L^\infty(L^2)$. Using the skew-symmetric property of the nonlinear term for weakly divergence-free functions and Hölder's inequality leads to

$$\left|\left((u_h \cdot \nabla)u_h, \phi_h\right) - \left((s_h \cdot \nabla)s_h, \phi_h\right)\right| \leq \left|\left((\phi_h \cdot \nabla)s_h, \phi_h\right)\right| + \left|\left((u_h \cdot \nabla)\phi_h, \phi_h\right)\right| \tag{55}$$
$$= \left|\left((\phi_h \cdot \nabla)s_h, \phi_h\right)\right| \leq \|\nabla s_h\|_{L^\infty}\|\phi_h\|_0^2.$$

The first term of the truncation error is bounded by using the Cauchy–Schwarz inequality and (33) for the temporal derivative

$$(\partial_t u - \partial_t s_h, \phi_h) \leq \|\partial_t u - \partial_t s_h\|_0 \|\phi_h\|_0 \leq Ch^r \|\partial_t u\|_r \|\phi_h\|_0.$$

For estimating the second term, Hölder's inequality, Sobolev embedding and interpolations, (34), and (33) are utilized

$$\left((u \cdot \nabla)u - (s_h \cdot \nabla)s_h, \phi_h\right) = \left(((u - s_h) \cdot \nabla)u - (s_h \cdot \nabla)(s_h - u), \phi_h\right)$$
$$\leq \left(\|u - s_h\|_{L^6}\|\nabla u\|_{L^3} + \|s_h\|_{L^\infty}\|\nabla(s_h - u)\|_0\right)\|\phi_h\|_0$$
$$\leq C\left(\|\nabla u\|_{1/2} + \|s_h\|_{L^\infty}\right)\|s_h - u\|_1\|\phi_h\|_0$$
$$\leq C\left((\|u\|_1\|u\|_2)^{1/2} + \|s_h\|_{L^\infty}\right)\|s_h - u\|_1\|\phi_h\|_0$$
$$\leq Ch^r \left(\|u\|_1\|u\|_2\right)^{1/2}\|u\|_{r+1}\|\phi_h\|_0. \tag{56}$$

Inserting the bounds for the truncation error and (55) in the error equation (54) and applying (34) yield

$$\frac{d}{dt}\|\phi_h\|_0 \leq \|\nabla s_h\|_{L^\infty}\|\phi_h\|_0 + Ch^r \left(\|\partial_t u\|_r + (\|u\|_1\|u\|_2)^{1/2}\|u\|_{r+1}\right)$$
$$\leq C\|\nabla u\|_{L^\infty}\|\phi_h\|_0 + Ch^r \left(\|\partial_t u\|_r + (\|u\|_1\|u\|_2)^{1/2}\|u\|_{r+1}\right).$$

Defining

$$L(T) = C\int_0^T \|\nabla u(s)\|_{L^\infty}\, ds, \quad M_u = C\left(\int_0^T \left(\|\partial_t u(s)\|_r + (\|u\|_1\|u\|_2)^{1/2}\|u(s)\|_{r+1}\right) ds\right), \tag{57}$$

and applying Gronwall's lemma give

$$\|\phi_h(t)\|_0 \leq e^{L(T)}\|\phi_h(0)\|_0 + e^{L(T)}M_u h^r.$$

The proof is finished by applying the triangle inequality. Note that all steps of the proof are dimensionally correct for $d = 3$, with the unit $\mathrm{m}^{d+2}/\mathrm{s}^3$ on both sides of all estimates.[2]

Theorem 4.7 (*Error Estimate for the Galerkin Method and Weakly Divergence-Free Pairs of Finite Element Spaces*). *For $T > 0$, assume that the solution of* (5) *is sufficiently regular and choose $\boldsymbol{u}_h(0) = \boldsymbol{s}_h(0)$. Then, it holds for $t \in (0, T]$*

$$\|\boldsymbol{u}(t) - \boldsymbol{u}_h(t)\|_0 \le e^{L(T)} M_{\boldsymbol{u}} h^r + C h^{r+1} \|\boldsymbol{u}\|_{L^\infty(H^{r+1})}, \tag{58}$$

where $L(T)$ and $M_{\boldsymbol{u}}$ are defined in (57).

It should be emphasized that the error bound (58) for the velocity error does not depend on the pressure, since, as already noted above, the pressure terms in the error equation vanish. This kind of error estimate is called pressure-robust.

In [58], a fully discrete finite element method for the incompressible Navier–Stokes equations is analyzed where Scott–Vogelius pairs of spaces $P_r/P_{r-1}^{\mathrm{disk}}$, $r \ge d$, are utilized for the spatial discretization and the so-called rotational form of the nonlinear convective term is used. The rotational form goes along with a modified pressure, the so-called Bernoulli pressure. Thanks to the use of weakly divergence-free finite elements for the velocity, the method is pressure robust. The velocity error in $L^\infty(L^2)$ is proved to be of order r, but the corresponding error bound is not robust for small values of the viscosity coefficient.

Example 4.8 (*Galerkin Method with a Weakly Divergence-Free Pair of Finite Element Spaces*). The same problem as in Example 4.4 is considered but with the Scott–Vogelius pair of spaces P_2/P_1^{disk}. A barycentric refinement needs to be applied to the grids arising from refining the initial grid from Fig. 1 leading to the number of degrees of freedom given in Table 2. The errors obtained for the Galerkin method (42) are shown in Fig. 9. One can see that the errors are small, where the order of error reduction is two on coarse grids and for small viscosity coefficients. This observation corresponds to the analytic prediction. Third order convergence can be observed for the large viscosity coefficient $\nu = 10^{-2}$ already on coarse grids and for smaller viscosity coefficients on finer grids. □

The optimal behavior for large viscosity coefficients on coarse grids was observed also for other methods and it will not be reported in further numerical studies.

Remark 4.9 (*Summary of the Results for the Plain Galerkin Method*). At this point, we can summarize the changes in the rate of error reduction of the $L^\infty(L^2)$ error of the velocity for the plain Galerkin method applied to the different equations, which are discretized with H^1-conforming velocity finite elements. For linear convection–reaction–diffusion equations, using piecewise polynomials of degree r, while the optimal rate $r + 1$ is achieved for ν large relative to the coefficients of the first order terms, only rate r can be obtained when $\nu \to 0$. For the Stokes and Oseen equations only the rate $r - 1$ is achieved, while for the Navier–Stokes equations, one cannot even prove robust error reduction unless the EMAC form of the nonlinear term is utilized (in that case the rate $r - 1$ is recovered) or weakly divergence-free pairs of finite elements are used (in that case the rate is increased to r).

By far most of the finite element analysis for evolutionary incompressible flow problems has been performed for H^1-conforming velocity finite element spaces. However, many pairs with such velocity spaces do not lead to weakly

---

[2] In order not to overload the presentation, dimensional correctness will be taken care of in the following only for $d = 3$. Only for this method, the changes for $d = 2$ are addressed in some detail once more. Instead of (56), the following estimate is performed, applying the same tools,

$$\big((\boldsymbol{u} \cdot \nabla)\boldsymbol{u} - (\boldsymbol{s}_h \cdot \nabla)\boldsymbol{s}_h, \boldsymbol{\phi}_h\big)$$
$$= \big(((\boldsymbol{u} - \boldsymbol{s}_h) \cdot \nabla)\boldsymbol{u} - (\boldsymbol{s}_h \cdot \nabla)(\boldsymbol{s}_h - \boldsymbol{u}), \boldsymbol{\phi}_h\big)$$
$$\le \big(\|\boldsymbol{u} - \boldsymbol{s}_h\|_{L^4}\|\nabla\boldsymbol{u}\|_{L^4} + \|\boldsymbol{s}_h\|_{L^\infty}\|\nabla(\boldsymbol{s}_h - \boldsymbol{u})\|_0\big)\|\boldsymbol{\phi}_h\|_0$$
$$\le C\big(\|\boldsymbol{u} - \boldsymbol{s}_h\|_{1/2}\|\nabla\boldsymbol{u}\|_{1/2} + \|\boldsymbol{s}_h\|_{L^\infty}\|\nabla(\boldsymbol{s}_h - \boldsymbol{u})\|_0\big)\|\boldsymbol{\phi}_h\|_0$$
$$\le C\big((\|\boldsymbol{u} - \boldsymbol{s}_h\|_0\|\boldsymbol{u} - \boldsymbol{s}_h\|_1)^{1/2}(\|\boldsymbol{u}\|_1\|\boldsymbol{u}\|_2)^{1/2} + (\|\boldsymbol{u}\|_0\|\boldsymbol{u}\|_2)^{1/2}\|\boldsymbol{u} - \boldsymbol{s}_h\|_1\big)\|\boldsymbol{\phi}_h\|_0$$
$$\le C h^r\left(h^{1/2}(\|\boldsymbol{u}\|_1\|\boldsymbol{u}\|_2)^{1/2} + (\|\boldsymbol{u}\|_0\|\boldsymbol{u}\|_2)^{1/2}\right)\|\boldsymbol{u}\|_{r+1}\|\boldsymbol{\phi}_h\|_0.$$

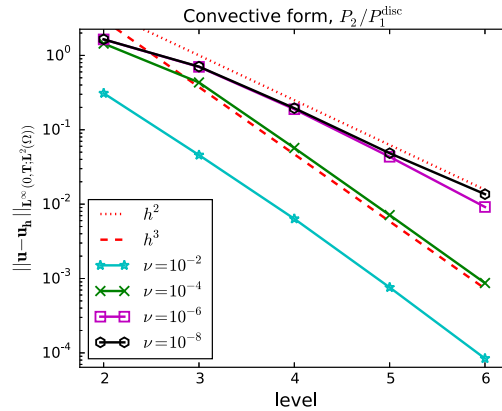Then, the expression for $M_{\boldsymbol{u}}$ changes accordingly.

Fig. 9. Example 4.8 Galerkin method for the Navier–Stokes equation, P_2/P_1^{disk} on a barycentric-refined grid, convective form of the convective term.

divergence-free discrete velocity solutions. Exceptions require usually a high polynomial degree and sometimes special grids, like the Scott–Vogelius pair of spaces for $r = d$. An attractive alternative approach consists in using $H(\text{div})$-conforming velocity spaces. Such spaces are actually used for simulations of flows with small viscosity, e.g., in [59], and one can find numerical analysis for stationary problems, e.g., in [60,61]. To the best of our knowledge, there are only few contributions that study the finite element analysis of weakly divergence-free $H(\text{div})$-conforming finite elements for evolutionary incompressible flow problems. In [62], this kind of discrete velocity spaces is considered together with pressure spaces such that a discrete inf–sup condition is satisfied. Concrete examples for such pairs are Raviart–Thomas (RT) spaces of order r for the velocity and discontinuous pressure spaces of order r on simplicial grids. Also, Brezzi–Douglas–Marini (BDM) spaces can be used for approximating the velocity. It is referred to [47, Sec. 4.4] for more details concerning these spaces. The analysis in [62] is performed for the evolutionary Oseen equations (29). Because $H(\text{div})$-conforming functions are generally not H^1-conforming, special care has to be taken for discretizing the viscous term such that the consistency error becomes sufficiently small. In [62], the symmetric interior penalty (SIP) form, which is well known from discontinuous Galerkin methods, is applied. Since weakly divergence-free velocity finite element functions are considered, the convective form of the convective term can be applied. It is augmented with an upwind term defined on the facets of the mesh cells. An error bound of order r with constants independent of inverse powers of the viscosity is proved for the $L^\infty(L^2)$ norm of the velocity. Like in the H^1-conforming case, (58), the bound is independent of the pressure. The restriction to order r is caused from the bound of the convective term. The results from [62] are extended to the Navier–Stokes equations in [63]. In [64], the authors improve the error bounds of [62], obtaining order $r+1/2$ for the Navier–Stokes equations. This result fits into Section 4.3 and a sketch of the proof will be presented there, see Theorem 4.23.

4.2.2. Non weakly divergence-free inf–sup stable pairs of finite element spaces

This section presents two methods that achieve robust error reduction of order r for the velocity error in $L^\infty(L^2)$ if the standard skew-symmetric form (44) of the nonlinear term is utilized. For one of these methods, even convergence for the convective form of the nonlinear term is shown.

The Galerkin method with grad–div stabilization and inf–sup stable pairs of mixed finite elements is analyzed in [50]. More precisely, the following discretization is considered: Find $\boldsymbol{u}_h : (0, T] \to \boldsymbol{V}_{h,r}$, $p_h : (0, T] \to Q_{h,r-1}$ satisfying

$$(\partial_t \boldsymbol{u}_h, \boldsymbol{v}_h) + \nu(\nabla \boldsymbol{u}_h, \nabla \boldsymbol{v}_h) + b(\boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{v}_h) + \mu(\nabla \cdot \boldsymbol{u}_h, \nabla \cdot \boldsymbol{v}_h) - (\nabla \cdot \boldsymbol{v}_h, p_h) = (\boldsymbol{f}, \boldsymbol{v}_h) \quad \forall\, \boldsymbol{v}_h \in \boldsymbol{V}_{h,r},$$

$$(\nabla \cdot \boldsymbol{u}_h, q_h) = 0 \quad \forall\, q_h \in Q_{h,r-1}, \tag{59}$$

where the nonlinear term is defined in (44) and $\mu > 0$ is the grad–div stabilization parameter. Note that the grad–div term vanishes for weakly divergence-free finite element velocities, which gives the Galerkin method considered in the previous subsection. The additional term in (59) can be derived by adding to the strong form of the momentum balance of the Navier–Stokes equations (1) the vanishing term $-\mu \nabla(\nabla \cdot \boldsymbol{u})$, integrating the corresponding weak

term by parts, and discretizing. This strong form is the origin for the notion grad–div stabilization, but sometimes one finds also the notion div–div stabilization in the literature. From the dimensional point of view, the stabilization parameter μ has the physical unit m^2/s, like a kinematic viscosity, e.g., compare the optimal parameters derived in [65] for the Stokes equations with viscosity.

Grad–div stabilization was proposed originally in [66] for penalizing the violation of mass conservation of pairs of finite element spaces where the velocity is not weakly divergence-free. This goal is usually achieved only to some extent, e.g., see [47]. However, it is shown in [50] that also dominant convection is stabilized with grad–div stabilization. The proof of the following error bound can be found in [50, Theorem 1].

Theorem 4.10 (*Error Estimate for the Galerkin Method with grad–div Stabilization*). *Consider $T > 0$, assume that the solution of (5) is sufficiently smooth, and let $r \geq 2$. Then, there exists a positive constant $M_{\boldsymbol{u},p}$ depending on $\|\boldsymbol{u}_0\|_r^2$ and*

$$\left(\int_0^T \left(\frac{\|p(s)\|_{H^r/\mathbb{R}}^2}{\mu} + T \|\partial_t \boldsymbol{u}(s)\|_r^2 + (\mu + T \|\boldsymbol{u}(s)\|_1 \|\boldsymbol{u}(s)\|_2) \|\boldsymbol{u}(s)\|_{r+1}^2 \right) ds \right)^{1/2}, \tag{60}$$

such that for $t \in [0, T]$ the following bound holds

$$\|\boldsymbol{u}(t) - \boldsymbol{u}_h(t)\|_0 \leq \exp(L(T)/2)h^r M_{\boldsymbol{u},p} + Ch^{r+1}\|\boldsymbol{u}\|_{L^\infty(H^{r+1})}, \tag{61}$$

where $L(T)$ is

$$L(T) = 1 + C \int_0^T \left(2 \|\nabla \boldsymbol{u}(s)\|_{L^\infty} + \frac{\|\boldsymbol{u}(s)\|_{L^\infty}^2}{2\mu} \right) ds.$$

For (60) to be dimensionally correct in the case $d = 2$, the term $\|\boldsymbol{u}(s)\|_1 \|\boldsymbol{u}(s)\|_2$ must be replaced by $h \|\boldsymbol{u}(s)\|_1 \|\boldsymbol{u}(s)\|_2 + \|\boldsymbol{u}(s)\|_0 \|\boldsymbol{u}(s)\|_2$.

Proof (*Sketch*)**.** Here, not the complete proof will be given, but only the usefulness of the grad–div stabilization in the derivation of the error bound will be explained. The grad–div term is used in the error analysis for absorbing some contributions from the nonlinear convective term. Considering only the important terms in the error equation for $\boldsymbol{\phi}_h = \boldsymbol{u}_h - \boldsymbol{s}_h$, see (48), yields with (46)

$$\frac{1}{2}\frac{d}{dt}\|\boldsymbol{\phi}_h\|_0^2 + \nu\|\nabla\boldsymbol{\phi}_h\|_0^2 + \mu\|\nabla \cdot \boldsymbol{\phi}_h\|_0^2 \leq -b(\boldsymbol{\phi}_h, \boldsymbol{s}_h, \boldsymbol{\phi}_h) + \cdots$$

$$\leq \|\nabla\boldsymbol{s}_h\|_{L^\infty}\|\boldsymbol{\phi}_h\|_0^2 + \frac{1}{2}\|\boldsymbol{s}_h\|_{L^\infty}\|\nabla \cdot \boldsymbol{\phi}_h\|_0\|\boldsymbol{\phi}_h\|_0 + \cdots$$

$$\leq \frac{\mu}{4}\|\nabla \cdot \boldsymbol{\phi}_h\|_0^2 + \left(\|\nabla\boldsymbol{s}_h\|_{L^\infty} + \frac{1}{4\mu}\|\boldsymbol{s}_h\|_{L^\infty}^2 \right) \|\boldsymbol{\phi}_h\|_0^2 + \cdots. \tag{62}$$

Another term that is bounded with the grad–div stabilization is the term coming from the pressure, i.e., the second term on the right-hand side in (39)

$$(\nabla \cdot \boldsymbol{\phi}_h, p - \Pi_h^{r-1}p) \leq \frac{\mu}{4}\|\nabla \cdot \boldsymbol{\phi}_h\|_0^2 + \frac{1}{\mu}\|p - \Pi_h^{r-1}p\|_0^2.$$

In contrast to (49), inverse powers of the viscosity are not present on the right-hand side of (62), but inverse powers of the grad–div stabilization parameter. □

The result of Theorem 4.10 holds true for inf–sup stable pairs of mixed finite elements where the polynomial degree r of the velocity space is larger by one than the degree of the pressure space. It is well known for the steady-state Stokes and Oseen problem from [65,67] that an appropriate choice of the grad–div stabilization parameter includes (semi-) norms of the solution in higher order Sobolev spaces, such that the parameter possesses the correct physical unit, but these (semi-) norms are not accessible. Thus, the usual way in practice for proposing a computable grad–div parameter consists in optimizing the derived error bound with respect to the order of convergence in space, such that the result is independent of inverse powers of the viscosity. Since the constant in (60) depends on μ and μ^{-1}, this approach gives $\mu \sim 1$ as asymptotic optimal choice for the grad–div stabilization parameter. In that case, an order of convergence r is achieved for the $L^\infty(L^2)$ error of the velocity, see (61). For inf–sup stable spaces with

the same polynomial degree, like the MINI element, Theorem 4.10 still holds, see [50, Remark 2]. In that case, assuming for the pressure $p \in H^2(\Omega)$, one can choose $\mu \sim h$ to achieve a linear order of convergence for the $L^\infty(L^2)$ error of the velocity.

Next, the same approach, namely the Galerkin finite element method with grad–div stabilization, will be considered, but using the convective form of the nonlinear term instead of the divergence form. Arguing as in (54) and (55) yields

$$
\frac{1}{2}\frac{d}{dt}\|\boldsymbol{\phi}_h\|_0^2 + \mu\|\nabla \cdot \boldsymbol{\phi}_h\|_0^2 \leq \left|\left((\boldsymbol{\phi}_h \cdot \nabla)\boldsymbol{s}_h, \boldsymbol{\phi}_h\right)\right| + \left|\left((\boldsymbol{u}_h \cdot \nabla)\boldsymbol{\phi}_h, \boldsymbol{\phi}_h\right)\right| + \cdots
$$
$$
\leq \|\nabla \boldsymbol{s}_h\|_{L^\infty}\|\boldsymbol{\phi}_h\|_0^2 + \left|\left((\boldsymbol{u}_h \cdot \nabla)\boldsymbol{\phi}_h, \boldsymbol{\phi}_h\right)\right| + \cdots. \tag{63}
$$

Now, the second term on the right-hand side does not vanish. To handle this term, integration by parts, Hölder's and Young's inequality are used to obtain

$$
|(\boldsymbol{u}_h \cdot \nabla\boldsymbol{\phi}_h, \boldsymbol{\phi}_h)| = \frac{1}{2}|((\nabla \cdot \boldsymbol{u}_h)\boldsymbol{\phi}_h, \boldsymbol{\phi}_h)|
$$
$$
\leq \frac{1}{2}\|\nabla \cdot \boldsymbol{s}_h\|_{L^\infty}\|\boldsymbol{\phi}_h\|_0^2 + \frac{1}{2}\|\nabla \cdot \boldsymbol{\phi}_h\|_0\|\boldsymbol{\phi}_h\|_{L^\infty}\|\boldsymbol{\phi}_h\|_0
$$
$$
\leq \frac{\mu}{4}\|\nabla \cdot \boldsymbol{\phi}_h\|_0^2 + \left(\frac{1}{2}\|\nabla \cdot \boldsymbol{s}_h\|_{L^\infty} + \frac{1}{4\mu}\|\boldsymbol{\phi}_h\|_{L^\infty}^2\right)\|\boldsymbol{\phi}_h\|_0^2.
$$

Inserting this estimate in (63) leads to

$$
\frac{1}{2}\frac{d}{dt}\|\boldsymbol{\phi}_h\|_0^2 + \frac{3\mu}{4}\|\nabla \cdot \boldsymbol{\phi}_h\|_0^2 \leq \left(\frac{1}{2}\|\nabla \cdot \boldsymbol{s}_h\|_{L^\infty} + \|\nabla \boldsymbol{s}_h\|_{L^\infty} + \frac{1}{4\mu}\|\boldsymbol{\phi}_h\|_{L^\infty}^2\right)\|\boldsymbol{\phi}_h\|_0^2 + \cdots.
$$

Utilizing now the following threshold condition, whose validity will be discussed below,

$$
\|\boldsymbol{\phi}_h\|_0 \leq C_{\text{thr}}h^{3/2}, \tag{64}
$$

and applying the inverse inequality (8) give

$$
\|\boldsymbol{\phi}_h\|_{L^\infty} \leq c_{\text{inv}}h^{-d/2}\|\boldsymbol{\phi}_h\|_0 \leq c_{\text{inv}}C_{\text{thr}} = C,
$$

so that $\|\boldsymbol{\phi}_h\|_{L^\infty}$ is bounded. For this result, the threshold condition (64) has to be proved. Assuming $\boldsymbol{u}_h(0) = \boldsymbol{s}_h(0)$, so that $\boldsymbol{\phi}_h(0) = \mathbf{0}$, a standard bootstrap argument can be applied for this purpose, see [68, Theorem 3.7]. Note that the threshold condition (64) cannot be proved for low order elements such as the MINI element. Using $\boldsymbol{u}_h(0) = \boldsymbol{s}_h(0)$ and applying the same analysis that proves Theorem 4.10 leads to $\|\boldsymbol{\phi}_h\|_0 = \mathcal{O}(h^r)$ for $r \geq 2$. Finally, the triangle inequality together with (33) shows that $\|\boldsymbol{u}(t) - \boldsymbol{u}_h(t)\|_0 = \mathcal{O}(h^r)$, $r \geq 2$, also in the case that the convective form of the nonlinear term is used for the Galerkin method with grad–div stabilization.
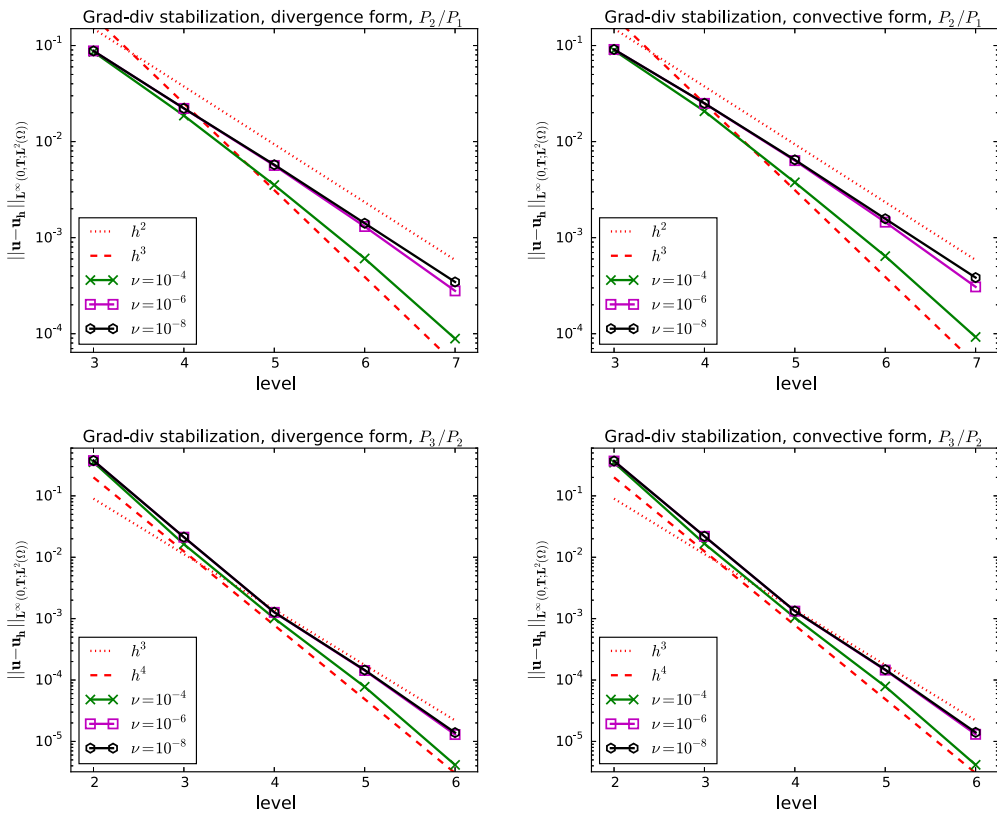
Example 4.11 (*Galerkin Method with grad–div Stabilization*). The same setup as in Example 4.4 is considered, but now the Galerkin method with grad–div stabilization (59) is used. Since the Taylor–Hood pairs of finite element spaces P_2/P_1 and P_3/P_2 were utilized in the simulations, the asymptotic optimal choice of the grad–div stabilization parameter is a constant independent of the mesh width. For the simulations, $\mu = 0.25$ was chosen, which is the same choice as in [50].

Results for the convective form and the divergence form of the nonlinear term are presented in Figs. 10 and 11. First, comparing them with Figs. 6 and 7, the stabilizing effect of the grad–div method can be clearly recognized. And second, there are basically no differences between the results of both discretizations of the nonlinear term, which is in accordance with the analysis discussed in this section. For small viscosity coefficients and on coarse meshes, the predicted second order error reduction for P_2/P_1 and third order error reduction for P_3/P_2 can be observed. □

In [69], the limit $\mu \to \infty$ is studied for Taylor–Hood pairs of spaces P_r/P_{r-1}. Assuming that for the given mesh the solution obtained with the Scott–Vogelius pair $P_r/P_{r-1}^{\text{disk}}$ is unique, i.e., that the Scott–Vogelius pair is discretely inf–sup stable, then it is shown that the sequence of velocity solutions for the Taylor–Hood pair with grad–div stabilization converges to the velocity solution of the Galerkin method for the Scott–Vogelius pair. Hence, [69] establishes a connection between the method of Section 4.2.1 and the grad–div stabilization.
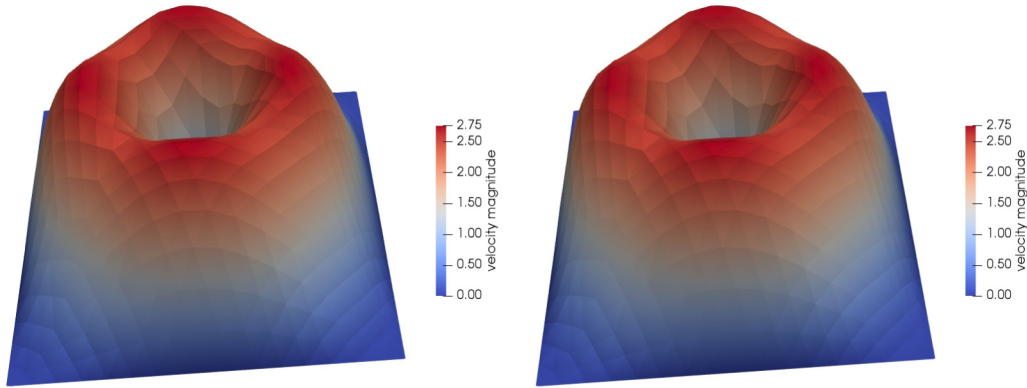
Fig. 10. Example 4.11 Galerkin method with grad–div stabilization for the Navier–Stokes equation, $\nu = 10^{-6}$, level 5, solution at $t = 1.3409$, left: convective form of the convective term, right: divergence form of the convective term.

Fig. 11. Example 4.11 Galerkin method with grad–div stabilization, left: divergence form of the nonlinear term, right: convective form of the nonlinear term.

In [70], a one-level LPS method is considered. For the finite element discretization, a pair of inf–sup stable spaces with spatial order r, $r \geq 2$, is used, in which the velocity space $V_{h,r}$ consists of continuous piecewise polynomials and the pressure space $Q_{h,r-1}$ contains continuous or discontinuous piecewise polynomials.

For any $K \in \mathcal{T}_h$, let $D(K)$ be a finite-dimensional space and $\Pi_K : L^2(K) \rightarrow D(K)$ be the local $L^2(K)$ projection into $D(K)$. The local fluctuation operator $\sigma_{h,K}^* : L^2(K) \rightarrow L^2(K)$ is given by $\sigma_{h,K}^* v = v - \Pi_K v$. The

stabilization term S_h is defined by

$$S_h(\boldsymbol{v}_h, \boldsymbol{w}_h) = \sum_{K \in \mathcal{T}_h} \tau_K \left(\sigma_{h,K}^*(\nabla \boldsymbol{v}_h), \sigma_{h,K}^*(\nabla \boldsymbol{w}_h) \right)_K \quad \text{for } \boldsymbol{v}_h, \boldsymbol{w}_h \in V_{h,r},$$

where $\{\tau_K\}$ are user-chosen non-negative constants. Since the stabilization is a viscous term, the stabilization parameters are viscosity coefficients from the dimensional point of view. Finally, it is set

$$\tau_h^{\min} := \min_{K \in \mathcal{T}_h} \tau_K, \qquad \tau_h^{\max} := \max_{K \in \mathcal{T}_h} \tau_K. \tag{65}$$

The LPS method studied in [70] reads as follows: Find $\boldsymbol{u}_h : (0,T] \to V_{h,r}$, $p_h : (0,T] \to Q_{h,r-1}$ satisfying

$$(\partial_t \boldsymbol{u}_h, \boldsymbol{v}_h) + \nu(\nabla \boldsymbol{u}_h, \nabla \boldsymbol{v}_h) + b(\boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{v}_h) + S_h(\boldsymbol{v}_h, \boldsymbol{w}_h) - (\nabla \cdot \boldsymbol{v}_h, p_h) = (\boldsymbol{f}, \boldsymbol{v}_h) \quad \forall \, \boldsymbol{v}_h \in V_{h,r},$$

$$(\nabla \cdot \boldsymbol{u}_h, q_h) = 0 \quad \forall \, q_h \in Q_{h,r-1}, \tag{66}$$

where the nonlinear term is defined in (44).

Two interpolation operators are defined in [70], where the first one is used to bound the velocity error and the second one to bound the error in the pressure. The operator $j_h : V \to V_{h,r}$ provides optimal approximation properties and preserves the discrete divergence, i.e.,

$$(\nabla \cdot (\boldsymbol{v} - j_h \boldsymbol{v}), q_h) = 0 \quad \forall \, q_h \in Q_{h,r-1}, \ \boldsymbol{v} \in V.$$

Let C_{stab} denote the following stability constant, see [70, (2.18)],

$$\|j_h \boldsymbol{v}\|_{1,p} \le C_{\text{stab}} \|\boldsymbol{v}\|_{1,p}, \quad p \in [2, \infty], \quad \boldsymbol{v} \in W^{1,p}(\Omega)^d. \tag{67}$$

The interpolation operator $i_h : L^2(\Omega) \to Q_{h,r-1}$ satisfies also optimal approximation properties, the stability property

$$\|q - i_h q\|_0 \le C_i \|q\|_0 \quad \forall \, q \in L^2(\Omega), \tag{68}$$

and the property

$$(q - i_h q, r_h)_K = 0 \quad \forall \, q \in L^2(\Omega), \ r_h \in D(K), \ K \in \mathcal{T}_h. \tag{69}$$

The following estimated is proved in [70, Lemma 3.3]. Let $\boldsymbol{v}_h \in V_h^{\text{div}}$, then

$$|(\nabla \cdot \boldsymbol{v}_h, \boldsymbol{\phi} \cdot \boldsymbol{\xi})| \le C_d S_h(\boldsymbol{v}_h, \boldsymbol{v}_h)^{1/2} \|\boldsymbol{\phi}\|_{L^\infty} \|\boldsymbol{\xi}\|_0, \quad \boldsymbol{\phi} \in L^\infty(\Omega)^d, \quad \boldsymbol{\xi} \in L^2(\Omega)^d, \tag{70}$$

holds true where

$$C_d := \frac{C_i \sqrt{d}}{\sqrt{\tau_h^{\min}}}, \tag{71}$$

with C_i being the constant in (68) and $\tau_h^{\min}$ is defined in (65). To prove (70), one uses that $\boldsymbol{v}_h \in V_h^{\text{div}}$ and the property (69) to get

$$(\nabla \cdot \boldsymbol{v}_h, \boldsymbol{\phi} \cdot \boldsymbol{\xi}) = (\nabla \cdot \boldsymbol{v}_h, \boldsymbol{\phi} \cdot \boldsymbol{\xi} - i_h(\boldsymbol{\phi} \cdot \boldsymbol{\xi})) = \sum_{K \in \mathcal{T}_h} \left(\sigma_{h,K}^*(\nabla \cdot \boldsymbol{v}_h), \boldsymbol{\phi} \cdot \boldsymbol{\xi} - i_h(\boldsymbol{\phi} \cdot \boldsymbol{\xi}) \right)_K.$$

Estimate (70) follows now by bounding the norm of the divergence with the norm of the gradient and then applying (68).

The proof of the following theorem can be found in [70, Theorem 3.1].

Theorem 4.12 (*Error Estimate for the LPS Method with Stabilization for the Gradient*). *For $T > 0$, assume for the solution of* (5) *sufficient regularity, let $r \ge 2$, and let $\tau_K \sim 1$ for all $K \in \mathcal{T}_h$. Then, for the LPS approximation defined in* (66) *with initial condition $\boldsymbol{u}_h(0) = j_h \boldsymbol{u}_0$ the following bound holds*

$$\|\boldsymbol{u}(t) - \boldsymbol{u}_h(t)\|_0 \le M_{\exp} M_{\boldsymbol{u},p} h^r, \quad 0 \le t \le T, \tag{72}$$

where

$$M_{\exp} := \exp \left(C_{\text{stab}} \|\boldsymbol{u}\|_{L^1(W^{1,\infty})} + \frac{C_{\text{stab}}^2 C_d^2 \text{diam}(\Omega)^2 \|\boldsymbol{u}\|_{L^2(W^{1,\infty})}^2}{2} + 1 \right),$$

and

$$M_{\boldsymbol{u},p} := C \left(T\|\partial_t \boldsymbol{u}\|_{L^2(H^r)}^2 + \left(\nu + \tau_h^{\max} + T\left(\|\boldsymbol{u}\|_{L^\infty(L^\infty)}^2 + h^2\|\boldsymbol{u}\|_{L^\infty(W^{1,\infty})}^2\right)\right)\|\boldsymbol{u}\|_{L^2(H^{r+1})}^2 + \frac{\|p\|_{L^2(H^r)}^2}{\tau_h^{\min}} \right)^{1/2},$$

with C_{stab} and C_d being the constants in (67) *and* (71), *respectively.*

Proof (*Sketch*). Considering only the important terms in the error equation for $\boldsymbol{\phi}_h = \boldsymbol{u}_h - j_h\boldsymbol{u}$ yields

$$\frac{1}{2}\frac{d}{dt}\|\boldsymbol{\phi}_h\|_0^2 + \nu\|\nabla\boldsymbol{\phi}_h\|_0^2 + S_h(\boldsymbol{\phi}_h, \boldsymbol{\phi}_h) \leq \left|-b(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{\phi}_h) + b(\boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{\phi}_h)\right| + \left|(\nabla \cdot \boldsymbol{\phi}_h, p - i_h p)\right| + \cdots.$$

To bound the last term, property (69), the Cauchy–Schwarz inequality, the definition of the stabilization term, and the bound of the $L^2(\Omega)$ norm of the divergence by the same norm of the gradient are utilized

$$(\nabla \cdot \boldsymbol{\phi}_h, p - i_h p) = \sum_{K \in \mathcal{T}_h} (\sigma_{h,K}^*(\nabla \cdot \boldsymbol{\phi}_h), p - i_h p)_K \leq \frac{\sqrt{d}}{\sqrt{\tau_h^{\min}}}\|p - i_h p\|_0 S_h(\boldsymbol{\phi}_h, \boldsymbol{\phi}_h)^{1/2}.$$

Denoting by $\boldsymbol{\eta} = \boldsymbol{u} - j_h\boldsymbol{u}$ and using the skew-symmetric property (45), the difference of the nonlinear terms is decomposed as

$$b(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{\phi}_h) - b(\boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{\phi}_h) = b(\boldsymbol{\eta}, \boldsymbol{u}, \boldsymbol{\phi}_h) + b(j_h\boldsymbol{u}, \boldsymbol{\eta}, \boldsymbol{\phi}_h) - b(\boldsymbol{\phi}_h, j_h\boldsymbol{u}, \boldsymbol{\phi}_h).$$

Applying the bound (46) and Young's inequality gives for the first term

$$|b(\boldsymbol{\eta}, \boldsymbol{u}, \boldsymbol{\phi}_h)| \leq 2T\|\boldsymbol{\eta}\|_0^2\|\nabla\boldsymbol{u}\|_{L^\infty}^2 + \frac{1}{2}T\|\nabla \cdot \boldsymbol{\eta}\|_0^2\|\boldsymbol{u}\|_{L^\infty}^2 + \frac{1}{4T}\|\boldsymbol{\phi}_h\|_0^2.$$

With a similar bound as (46), one obtains for the second term

$$|b(j_h\boldsymbol{u}, \boldsymbol{\eta}, \boldsymbol{\phi}_h)| \leq 2T\|\nabla\boldsymbol{\eta}\|_0^2\|j_h\boldsymbol{u}\|_{L^\infty}^2 + \frac{1}{2}T\|\nabla \cdot j_h\boldsymbol{u}\|_{L^\infty}^2\|\boldsymbol{\eta}\|_0^2 + \frac{1}{4T}\|\boldsymbol{\phi}_h\|_0^2.$$

For bounding the third term, one applies Hölder's inequality to the definition (44) of the nonlinear term and uses (70) and Young's inequality

$$
\begin{aligned}
|b(\boldsymbol{\phi}_h, j_h\boldsymbol{u}, \boldsymbol{\phi}_h)| &\leq \|\nabla j_h\boldsymbol{u}\|_{L^\infty}\|\boldsymbol{\phi}_h\|_0^2 + \frac{C_d}{2}S_h(\boldsymbol{\phi}_h, \boldsymbol{\phi}_h)^{1/2}\|\boldsymbol{\phi}_h\|_0\|j_h\boldsymbol{u}\|_{L^\infty} \\
&\leq \frac{1}{8}S_h(\boldsymbol{\phi}_h, \phi_h) + \left(\|\nabla j_h\boldsymbol{u}\|_{L^\infty} + \frac{C_d^2}{2}\|j_h\boldsymbol{u}\|_{L^\infty}^2\right)\|\boldsymbol{\phi}_h\|_0^2.
\end{aligned}
$$

The first term is absorbed by the left-hand side and the second term is handled with Gronwall's lemma. Notice that, since $j_h\boldsymbol{u}$ is continuous, piecewise differentiable and vanishes on the boundary, for $\boldsymbol{x} \in \Omega$ one can write

$$j_h\boldsymbol{u}(\boldsymbol{x}) = \int_0^1 \nabla j_h\boldsymbol{u}((1-\xi)\boldsymbol{x}_0 + \xi\boldsymbol{x}) \cdot (\boldsymbol{x} - \boldsymbol{x}_0)\,d\xi,$$

for an appropriate $\boldsymbol{x}_0 \in \partial\Omega$, so that one deduces that $\|j_h\boldsymbol{u}\|_{L^\infty} \leq \text{diam}(\Omega)\|\nabla j_h\boldsymbol{u}\|_{L^\infty}$.

Altogether, all important terms can be bounded without introducing inverse powers of the viscosity coefficient. The order of the interpolation terms in the bounds leads to (72). $\square$

Example 4.13 (*Inf–sup Stable Pair of Finite Element Spaces with LPS Stabilization for the Velocity Gradient*[3]). In [70], finite element spaces that satisfy the assumption of the analysis are mentioned. On quadrilateral or hexahedral meshes, the combination of the mapped $Q_r^{\text{bubble}}/P_{r-1}^{\text{disk}}$, $r \geq 2$, spaces for velocity and pressure and $D(K) = P_{r-1}(K)$ are suitable. In mapped spaces, the basis functions and nodal functionals are defined on a reference mesh cell $\hat{K}$ and the spaces on the actual mesh cells are obtained with the corresponding map. The velocity space on the reference cell $\hat{K} = [-1, 1]^2$ is given by

$$Q_r^{\text{bubble}}(\hat{K}) = Q_r(\hat{K}) + \text{span}\left\{(1 - \hat{x}_1^2)(1 - \hat{x}_2^2)\hat{x}_i^{r-1}, \ i = 1, 2\right\}.$$

---

[3] We like to acknowledge Naveed Ahmed (Gulf University for Science and Technology, Kuwait City) for providing us the implementation of this method.

Table 3
Number of degrees of freedom (including Dirichlet nodes) for Example 4.13.

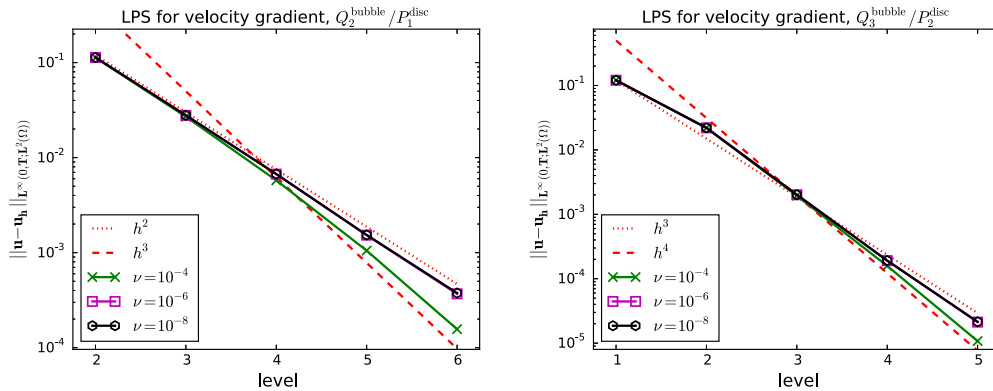|	$Q_2^{\text{bubble}}/P_1^{\text{disk}}$		$Q_3^{\text{bubble}}/P_2^{\text{disk}}$	
Level	Velocity	Pressure	Velocity	Pressure
1			114	24
2	226	48	402	96
3	834	192	1506	384
4	3202	768	5826	1536
5	12546	3072	22914	6144
6	49666	12288		

Fig. 12. Example 4.13 LPS stabilization for the velocity gradient.

The same example as defined in Example 4.4 is considered. For defining the initial grid, level 1, the domain Ω was decomposed into four squares, see Table 3 for information on the numbers of degrees of freedom. The stabilization parameter was set to be $\tau_K = 0.1$, which is the same value as used in [70]. The result of the numerical simulations, presented in Fig. 12, confirms the prediction from Theorem 4.12. □

4.2.3. Non inf–sup stable pairs of finite element spaces

The use of non inf–sup stable pairs of finite element spaces is appealing from the implementation point of view, since the same principal finite element spaces can be used for velocity and pressure, but it requires a stabilization of the discrete continuity equation, a so-called pressure stabilization, for obtaining a well-posed problem. For a recent survey of pressure stabilizations, for the model problem of the Stokes equations, it is referred to [71].

Non inf–sup stable finite element approximations to (5) are studied in [72] and several methods are analyzed. In particular, the case already mentioned above that the polynomial spaces for the finite element velocity and pressure are the same, apart of the boundary condition and the vector-valued character of the velocity space, is considered. For the pressure stabilization, a local projection stabilization (LPS) is used that is the so-called term-by-term stabilization introduced in [73]. Term-by-term stabilization is a special variant of LPS. It is defined just on a single mesh and the projection-based stabilization of standard LPS methods is replaced by an interpolation-based stabilization. The stability and weak convergence for a method, including LPS stabilization for the convective term and for the pressure, in natural norms of the solution of the Navier–Stokes equations are proved in [74].

In [72], fully discrete approximations to (5) with the implicit Euler method in time are analyzed. Here, for consistency of presentation, we introduce the methods for the continuous-in-time case. Clearly, the basic ideas of the fully discrete error analysis from [72] can be carried over to the continuous-in-time case with the appropriate changes.

The first method studied in [72] is an extension of the Galerkin method with grad–div stabilization to non inf–sup stable elements. This method, which includes an LPS term for the pressure, reads as follows: Find $\boldsymbol{u}_h : (0, T] \to V_{h,r}$, $p_h : (0, T] \to Q_{h,r}$ such that

$$(\partial_t \boldsymbol{u}_h, \boldsymbol{v}_h) + \nu(\nabla \boldsymbol{u}_h, \nabla \boldsymbol{v}_h) + b(\boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{v}_h) - (\nabla \cdot \boldsymbol{v}_h, p_h) + \mu(\nabla \cdot \boldsymbol{u}_h, \nabla \cdot \boldsymbol{v}_h) = (\boldsymbol{f}, \boldsymbol{v}_h) \quad \forall \, \boldsymbol{v}_h \in \boldsymbol{V}_{h,r},$$

$$(\nabla \cdot \boldsymbol{u}_h, q_h) + s_{\text{pres}}(p_h, q_h) = 0 \quad \forall \, q_h \in Q_{h,r}, \tag{73}$$

where the nonlinear term is defined in (44), the LPS pressure stabilization is given by

$$s_{\text{pres}}(p_h, q_h) = \sum_{K \in \mathcal{T}_h} \tau_{p,K}(\sigma_h^*(\nabla p_h), \sigma_h^*(\nabla q_h))_K, \tag{74}$$

μ and $\{\tau_{p,K}\}$ are the grad–div and pressure stabilization parameters, respectively, and $\sigma_h^* = \text{Id} - \sigma_h^{r-1}$ is the fluctuation operator, where Id is the identity operator and σ_h^{r-1} is a locally stable projection or interpolation operator from $L^2(\Omega)^d$ onto $\boldsymbol{Y}_{h,r-1}$. This operator can be chosen as the Bernardi–Girault [75] or the Scott–Zhang [76], [77, § 4.8], interpolation operator. The spaces $\boldsymbol{Y}_{h,l}$ are defined as follows:

$$Y_{h,l} = \left\{ v_h \in C^0(\overline{\Omega}) \mid v_{h|_K} \in P_l(K), \quad \forall \, K \in \mathcal{T}_h \right\}, \; l \geq 1, \quad \boldsymbol{Y}_{h,l} = (Y_{h,l})^d,$$

where $P_l(K)$ is the space of polynomials on K of degree less than or equal to l. The following approximation property holds for the fluctuations

$$\|\sigma_h^*(\boldsymbol{v})\|_0 = \|(\text{Id} - \sigma_h^j)(\boldsymbol{v})\|_0 \leq C h^s |\boldsymbol{v}|_s, \quad 1 \leq s \leq j + 1. \tag{75}$$

The physical unit of the pressure stabilization parameters is second, hence it is a time scale.

The proof of the following theorem can be obtained arguing as in [72, Theorem 3.5].

Theorem 4.14 (*Non inf–sup Stable Pairs of Finite Element Spaces with grad–div Stabilization and LPS Pressure Stabilization*). *Consider $T > 0$ and assume sufficient regularity for the solution of (5). Let $r \geq 2$ and let*

$$\alpha_1 h_K^2 \leq \tau_{p,K} \leq \alpha_2 h_K^2, \tag{76}$$

for some positive constants α_1, α_2 independent of h and ν. Then, there exists a positive constant $M_{\boldsymbol{u},p}$ depending on $\|\boldsymbol{u}_0\|_r$ and

$$\left(\int_0^T \left((\alpha_2 + \mu^{-1}) \|p(s)\|_{H^r/\mathbb{R}}^2 + T \|\partial_t \boldsymbol{u}(s)\|_r^2 \right. \right.$$
$$\left. \left. + \left(\nu + \mu + \alpha_1^{-1} + T |\Omega|^{\frac{4-d}{d}} \|\boldsymbol{u}(s)\|_2^2 \right) \|\boldsymbol{u}(s)\|_{r+1}^2 \right) ds \right)^{1/2}$$

such that for $t \in [0, T]$ the following bound holds

$$\|\boldsymbol{u}(t) - \boldsymbol{u}_h(t)\|_0 \leq M_{\boldsymbol{u},p} \exp\left(\frac{L(T)}{2} \right) h^r + C h^{r+1} \|\boldsymbol{u}\|_{L^\infty(H^{r+1})}, \tag{77}$$

where $L(T)$ is

$$L(T) = 1 + C \int_0^T |\Omega|^{\frac{4-d}{2d}} \left(2 \|\boldsymbol{u}(s)\|_3 + |\Omega|^{\frac{4-d}{2d}} \frac{\|\boldsymbol{u}(s)\|_2^2}{2\mu} \right) ds.$$

Proof (*Sketch*). This sketch stresses only the main new aspect if non inf–sup stable pairs of finite element spaces are studied instead of inf–sup stable pairs.

For the proof of Theorem 4.14, see [72, Theorem 3.5], one compares the approximation $\boldsymbol{u}_h$ with $\hat{\boldsymbol{u}}_h = R_h \boldsymbol{u} \in \boldsymbol{V}_{h,r}$ satisfying

$$(\boldsymbol{u} - \hat{\boldsymbol{u}}_h, \boldsymbol{v}_h) = 0, \quad \forall \, \boldsymbol{v}_h \in \boldsymbol{Y}_{h,r-1}. \tag{78}$$

The interpolation operator $R_h \boldsymbol{u}$ exists and satisfies optimal approximation properties, see [73]. In the proof of Theorem 4.14, the nonlinear term is bounded as in (62) with s_h replaced by $\hat{\boldsymbol{u}}_h$. In [72, (3.27), (3.28)] it is shown that $\|\hat{\boldsymbol{u}}_h\|_{L^\infty}$ and $\|\nabla \hat{\boldsymbol{u}}_h\|_{L^\infty}$ can be bounded in terms of $\|\boldsymbol{u}\|_2$ and $\|\boldsymbol{u}\|_3$, respectively, but with constants that are not scale-invariant. Careful dimensional analysis shows that

$$\|\hat{\boldsymbol{u}}_h\|_{L^\infty} \leq C |\Omega|^{\frac{4-d}{2d}} \|\boldsymbol{u}\|_2, \quad \|\nabla \hat{\boldsymbol{u}}_h\|_{L^\infty} \leq C |\Omega|^{\frac{4-d}{2d}} \|\boldsymbol{u}\|_3.$$

A new aspect for non inf–sup stable pairs of finite element spaces is that one cannot apply test functions that are discretely divergence-free, since such finite element functions do not necessarily exist. Consequently, the error equation contains a term that couples the discrete pressure p_h, which does not appear in the error equation for inf–sup stable pairs of finite element spaces like (38), and the divergence of the velocity test function. The finite element pressure p_h is compared with $\hat{p}_h$, which is defined as the Lagrange interpolant of the continuous pressure p to which one subtracts its mean. Property (78) is required for bounding the term $(\nabla \cdot \hat{\boldsymbol{u}}_h, \lambda_h)$, which appears in the error equation, where $\lambda_h = \hat{p}_h - p_h$. Using that $\boldsymbol{u}$ is divergence-free, integration by parts, the range of σ_h^{r-1}, and (78) yields

$$|(\nabla \cdot \hat{\boldsymbol{u}}_h, \lambda_h)| = |(\nabla \cdot (\hat{\boldsymbol{u}}_h - \boldsymbol{u}), \lambda_h)| = |(\hat{\boldsymbol{u}}_h - \boldsymbol{u}, \nabla \lambda_h)| = |(\hat{\boldsymbol{u}}_h - \boldsymbol{u}, \sigma_h^*(\nabla \lambda_h))|.$$

From this equation, one obtains with (76) and the approximation properties of $\hat{\boldsymbol{u}}_h$

$$|(\nabla \cdot \hat{\boldsymbol{u}}_h, \lambda_h)| \le C\alpha_1^{-1} h^{2r} \|\boldsymbol{u}\|_{r+1}^2 + \frac{1}{4} s_{\text{pres}}(\lambda_h, \lambda_h),$$

such that the second term on the right-hand side can be absorbed into the left-hand side of the error equation. □

From the error bound (77), one observes the order of convergence r, $r \ge 2$, for the $L^\infty(L^2)$ error of the velocity. The error bound (77) is the analog to (61) for the Galerkin method with grad–div stabilization and inf–sup stable elements. As for inf–sup stable pairs of finite element spaces, an asymptotic optimal choice of the grad–div stabilization parameter with respect to the mesh width is $\mu \sim 1$, due to the appearance of both μ and μ^{-1} on the right-hand side of (77). It is noted in [72] that under assuming additional regularity of the pressure, the choice $\mu \sim h$ is likewise asymptotically optimal.

From the point of view of the physical dimension, the parameters $\{\tau_{p,K}\}$ are time scales. Hence, the inverse of α_1 and α_2 are viscosity parameters. In fact, for the steady-state Stokes equations with viscosity, the parameter for LPS pressure stabilizations, and also several other pressure stabilizations, has the form Ch_K^2/ν, e.g., compare the review on pressure stabilizations [71]. But for these equations, one cannot derive robust error bounds, because the dependency on inverse powers of the viscosity enters from estimating the pressure term. Inspecting the analysis from [72], which leads to the robust estimate (77), shows that the choice $\tau_{p,K} = Ch_K^2/\nu$ does not enable to derive a robust estimate since $\tau_{p,K}$ has to be bounded from above without having the factor ν available in the corresponding expression.

Another method analyzed in [72] uses LPS stabilization of the velocity gradient in the discrete momentum equation instead of grad–div stabilization: Find $\boldsymbol{u}_h : (0, T] \to \boldsymbol{V}_{h,r}$, $p_h : (0, T] \to Q_{h,r}$ such that

$$
\begin{aligned}
(\partial_t \boldsymbol{u}_h, \boldsymbol{v}_h) &+ \nu(\nabla \boldsymbol{u}_h, \nabla \boldsymbol{v}_h) + b(\boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{v}_h) - (\nabla \cdot \boldsymbol{v}_h, p_h) \\
&+ \sum_{K \in \mathcal{T}_h} \tau_{\nu,K} \left(\sigma_h^*(\nabla \boldsymbol{u}_h), \sigma_h^*(\nabla \boldsymbol{v}_h) \right)_K = (\boldsymbol{f}, \boldsymbol{v}_h) \quad \forall \, \boldsymbol{v}_h \in \boldsymbol{V}_{h,r}, \\
&\quad (\nabla \cdot \boldsymbol{u}_h, q_h) + s_{\text{pres}}(p_h, q_h) = 0 \quad \forall \, q_h \in Q_{h,r},
\end{aligned}
\tag{79}
$$

where the nonlinear term is defined in (44) and the LPS pressure term in (74). The LPS velocity term is a viscous term, hence the parameters $\{\tau_{\nu,K}\}$ represent viscosity parameters.

Theorem 4.15 (*Non inf–sup Stable Pairs of Finite Element Spaces with Velocity Gradient LPS Term*). *Let $T > 0$, assume that the solution of (5) is sufficiently smooth, let $r \ge 2$, and let the stabilization parameters satisfy*

$$\alpha_1 h_K^2 \le \tau_{p,K} \le \alpha_2 h_K^2, \qquad \tau_1 \le \tau_{\nu,K} \le \tau_2, \tag{80}$$

for some positive constants $\alpha_1, \alpha_2, \tau_1$, and τ_2, independent of h. Then, there is a positive constant $M_{\boldsymbol{u},p}$, which depends on $\|\boldsymbol{u}_0\|_r$ and

$$\left(\int_0^T \left((\alpha_2 + \tau_1^{-1}) \|p(s)\|_{H^r/\mathbb{R}}^2 + T \|\partial_t \boldsymbol{u}(s)\|_r^2 + \left(\nu + \tau_1 + T |\Omega|^{\frac{4-d}{d}} \|\boldsymbol{u}(s)\|_2^2 \right) \|\boldsymbol{u}(s)\|_{r+1}^2 \right) ds \right)^{1/2},$$

such that for $t \in [0, T]$ the error bound

$$\|\boldsymbol{u}(t) - \boldsymbol{u}_h(t)\|_0 \le M_{\boldsymbol{u},p} \exp\left(\frac{L(T)}{2} \right) h^r + Ch^{r+1} \|\boldsymbol{u}\|_{L^\infty(H^{r+1})}$$

holds, where $L(T)$ is given by

$$L(T) = 1 + C \int_0^T |\Omega|^{\frac{4-d}{2d}} \left(\|\boldsymbol{u}(s)\|_3 + |\Omega|^{\frac{4-d}{2d}} \frac{\|\boldsymbol{u}(s)\|_2^2}{\tau_2} \right) ds.$$

Proof (*Sketch*)**.** Again, the presentation will be restricted to the main issues in the proof of Theorem 4.15, see the proof of [72, Theorem 4.5] for more details.

As in the proof of Theorem 4.14, $\boldsymbol{u}_h$ will be compared with $\hat{\boldsymbol{u}}_h$ satisfying (78). To bound error terms containing the pressure appropriately, p_h is compared with $\hat{p}_h = R_h p$, where $\hat{p}_h$ satisfies a property analogous to (78) for scalar functions

$$(p - \hat{p}_h, q_h) = 0, \quad \forall\, q_h \in Y_{h,r-1}. \tag{81}$$

Denote by $\boldsymbol{\phi}_h = \hat{\boldsymbol{u}}_h - \boldsymbol{u}_h$ and by $\lambda_h = \hat{p}_h - p_h$. The main step in the proof is the bound of the nonlinear convective term in the error equation. It will be shown that this term can be bounded with a constant independent of inverse powers of the viscosity using only the pressure stabilization. Arguing as in (46), the first step of bounding the nonlinear term is

$$|b(\boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{\phi}_h) - b(\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{u}}_h, \boldsymbol{\phi}_h)| \leq \|\nabla \hat{\boldsymbol{u}}_h\|_{L^\infty} \|\boldsymbol{\phi}_h\|_0^2 + \frac{1}{2}((\nabla \cdot \boldsymbol{\phi}_h)\hat{\boldsymbol{u}}_h, \boldsymbol{\phi}_h).$$

For the second term on the right-hand side, the key observation is that it has a structure that appears in the discrete continuity equation. A decomposition of this term will be applied that allows to utilize the error equation with respect to the continuity equation, introducing in this way $s_{\mathrm{pres}}(\cdot, \cdot)$. More precisely, the decomposition has the form

$$((\nabla \cdot \boldsymbol{\phi}_h)\hat{\boldsymbol{u}}_h, \boldsymbol{\phi}_h) = \left(\nabla \cdot \boldsymbol{\phi}_h, \sigma_h^r(\hat{\boldsymbol{u}}_h \cdot \boldsymbol{\phi}_h) \right) + \left(\nabla \cdot \boldsymbol{\phi}_h, (I - \sigma_h^r)(\hat{\boldsymbol{u}}_h \cdot \boldsymbol{\phi}_h) \right). \tag{82}$$

For the first term on the right-hand side of (82), one can use $\left(\boldsymbol{0}, \sigma_h^r(\hat{\boldsymbol{u}}_h \cdot \boldsymbol{\phi}_h) \right)$ as test function in the error equation, since by the definition of the projection operator, it is an admissible choice. In this way, a link between the pressure stabilization and the estimate of the nonlinear convective term is established

$$\left(\nabla \cdot \boldsymbol{\phi}_h, \sigma_h^r(\hat{\boldsymbol{u}}_h \cdot \boldsymbol{\phi}_h) \right) = s_{\mathrm{pres}}(p_h, \sigma_h^r(\hat{\boldsymbol{u}}_h \cdot \boldsymbol{\phi}_h)) + (\nabla \cdot \hat{\boldsymbol{u}}_h, \sigma_h^r(\hat{\boldsymbol{u}}_h^{n+1} \cdot \boldsymbol{\phi}_h)). \tag{83}$$

Using the approximation properties of $\hat{\boldsymbol{u}}_h$ and the $L^2(\Omega)$ stability of the operator σ_h^r, one obtains for the second term on the right-hand side of (83)

$$(\nabla \cdot \hat{\boldsymbol{u}}_h, \sigma_h^r(\hat{\boldsymbol{u}}_h^{n+1} \cdot \boldsymbol{\phi}_h)) \leq C h^{2r} \tau_2 \|\boldsymbol{u}\|_{r+1}^2 + C \tau_2^{-1} \|\hat{\boldsymbol{u}}_h\|_{L^\infty}^2 \|\boldsymbol{\phi}_h\|_0^2.$$

For the first term on the right-hand side of (83), one adds and subtracts $\hat{p}_h$ in the first argument of $s_{\mathrm{pres}}(\cdot, \cdot)$. Since $s_{\mathrm{pres}}(\cdot, \cdot)$ is a continuous, symmetric, positive semi-definite bilinear form, compare Assumptions A1 and A2 for a more general pressure stabilization, there holds in particular a Cauchy–Schwarz inequality. Utilizing this Cauchy–Schwarz inequality, the bound of the pressure stabilization parameters (80), the $L^2(\Omega)$ stability of σ_h^r, the inverse inequality (8), the approximation properties of $\hat{p}_h$, and the approximation properties (75) of the fluctuation operator gives

$$s_{\mathrm{pres}}(p_h, \sigma_h^r(\hat{\boldsymbol{u}}_h \cdot \boldsymbol{\phi}_h))$$

$$\leq \frac{1}{4} s_{\mathrm{pres}}(\hat{p}^h, \hat{p}^h) + 3 s_{\mathrm{pres}}(\sigma_h^r(\hat{\boldsymbol{u}}_h \cdot \boldsymbol{\phi}_h), \sigma_h^r(\hat{\boldsymbol{u}}_h \cdot \boldsymbol{\phi}_h)) + \frac{1}{8} s_{\mathrm{pres}}(\lambda_h, \lambda_h)$$

$$\leq C \alpha_2 h^2 \left(\|\sigma_h^*(\nabla(p - \hat{p}^h))\|_0^2 + \|\sigma_h^*(\nabla p)\|_0^2 + \|\sigma_h^*(\nabla(\sigma_h^r(\hat{\boldsymbol{u}}_h \cdot \boldsymbol{\phi}_h)))\|_0^2 \right) + \frac{1}{8} s_{\mathrm{pres}}(\lambda_h, \lambda_h)$$

$$\leq C \alpha_2 h^{2r} \|p\|_r^2 + C \alpha_2 \|\hat{\boldsymbol{u}}_h\|_{L^\infty}^2 \|\boldsymbol{\phi}_h\|_0^2 + \frac{1}{8} s_{\mathrm{pres}}(\lambda_h, \lambda_h).$$

To conclude the estimate of the nonlinear term, one needs to bound the second term on the right-hand side of (82). To this end, the following result from [78, Theorem 2.2] is applied: for $\boldsymbol{v} \in W^{1,\infty}(\Omega)^d$ and $\boldsymbol{v}_h \in \boldsymbol{Y}_{h,r}$, it holds

$$\|(I - \sigma_h^r)(\boldsymbol{v} \cdot \boldsymbol{v}_h)\|_0 \leq C h \|\nabla \boldsymbol{v}\|_{L^\infty} \|\boldsymbol{v}_h\|_0. \tag{84}$$

Applying (84) to $\boldsymbol{v} = \hat{\boldsymbol{u}}_h$ and $\boldsymbol{v}_h = \boldsymbol{\phi}_h$ together with the inverse inequality (8) yields

$$\left((\nabla \cdot \boldsymbol{\phi}_h), (I - \sigma_h^r)(\hat{\boldsymbol{u}}_h \cdot \boldsymbol{\phi}_h) \right) \leq C \|\nabla \cdot \boldsymbol{\phi}_h\|_0 h \|\nabla \hat{\boldsymbol{u}}_h\|_{L^\infty} \|\boldsymbol{\phi}_h\|_0 \leq C \|\nabla \hat{\boldsymbol{u}}_h\|_{L^\infty} \|\boldsymbol{\phi}_h\|_0^2.$$

Collecting all estimates leads to

$$
\begin{aligned}
\left| b(\boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{\phi}_h) - b(\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{u}}_h, \boldsymbol{\phi}_h) \right| &\leq C \left(\|\nabla \hat{\boldsymbol{u}}_h\|_\infty + (\alpha_2 + \tau_2^{-1}) \|\hat{\boldsymbol{u}}_h\|_\infty^2 \right) \|\boldsymbol{\phi}_h\|_0^2. \\
&\quad + Ch^{2r} \left(\alpha_2 \|p\|_r^2 + \tau_2 \|\boldsymbol{u}\|_{r+1}^2 \right) + \frac{1}{8} s_{\mathrm{pres}}(\lambda_h, \lambda_h).
\end{aligned}
\tag{85}
$$

The first term on the right-hand side is handled as in the previous theorems by applying Gronwall's lemma, the second term becomes part of the error bound, and the last term is absorbed into the left-hand side of the error equation. We want to emphasize that in the above proof, it is essential to have velocity and pressure finite element spaces containing polynomials of the same degree. This fact comes from the application of (84) to a function in the velocity space ($\boldsymbol{v}_h \in Y_{h,r}$) together with the use of $\sigma_h^r(\hat{\boldsymbol{u}}_h \cdot \boldsymbol{\phi}_h)$ as a test function for the pressure.

To finish with the sketch of the proof of Theorem 4.15, it shall be indicated in which step property (81) is required. Since there is no grad–div stabilization in the method, there is no direct control on the divergence of the error and, as a consequence, the following term, which appears in the error equation, cannot be bounded in the same way as in Theorem 4.14. Using the orthogonality property (81), $\|\nabla \cdot \boldsymbol{\phi}_h\|_{0,K} \leq \sqrt{d} \|\nabla \boldsymbol{\phi}_h\|_{0,K}$, $\tau_{v,K} > \tau_1$, and the approximation properties of $\hat{p}_h$ gives

$$
\begin{aligned}
(p - \hat{p}_h, \nabla \cdot \boldsymbol{\phi}_h) &= (p - \hat{p}_h, \sigma_h^*(\nabla \cdot \boldsymbol{\phi}_h)) \\
&\leq \left(\sum_{K \in \mathcal{T}_h} \tau_{v,K}^{-1} \|p - \hat{p}_h\|_{0,K}^2 \right)^{1/2} \sqrt{d} \left(\sum_{K \in \mathcal{T}_h} \tau_{v,K} \|\sigma_h^*(\nabla \boldsymbol{\phi}_h)\|_{0,K}^2 \right)^{1/2} \\
&\leq Ch^{2r} \tau_1^{-1} \|p\|_r^2 + \frac{1}{4} \sum_{K \in \mathcal{T}_h} \tau_{v,K} \|\sigma_h^*(\nabla \boldsymbol{\phi}_h)\|_{0,K}^2.
\end{aligned}
\tag{86}
$$

The second term on the right-hand side is now absorbed into the left-hand side of the error equation. $\quad\square$

Hence, also method (79) provides convergence of order r for the velocity in $L^\infty(L^2)$ with constants independent of inverse powers of v.

We like to mention that velocity fluctuations are also used to define turbulence models for the simulation of incompressible turbulent flows. In so-called three-scale algebraic VMS-multigrid methods, proposed in [79,80], the deformation tensor of the velocity fluctuations is the main part of the turbulence model and in the three-scale coarse space projection-based VMS method, proposed in [81], fluctuations of the deformation tensor are the main part, e.g., see [82] for a survey of VMS methods. The parameter used in these methods depends typically on the numerical solution and therefore it is nonlinear.

Inspecting the proof of Theorem 4.15, it can be observed that the LPS term with the velocity gradient can be replaced with a respective LPS term for the divergence

$$
\sum_{K \in \mathcal{T}_h} \tau_{\mu,K} \left(\sigma_h^r(\nabla \cdot \boldsymbol{u}_h), \sigma_h^r(\nabla \cdot \boldsymbol{v}_h) \right)_K,
$$

with $\tau_{\mu,K} \sim 1$, and the same order of convergence as in Theorem 4.15 can be proved, compare [72, Theorem 4.6].

There is another LPS method for equal order pairs of finite element spaces for which a finite element error analysis can be found in the literature. This method includes term-by-term LPS stabilization for the convective term, the divergence and the pressure. However, the available error bound for this method is not robust. For this reason, its discussion is postponed to Section 4.4.2.

It was already emphasized in the proof of Theorem 4.15 that the convective term in the error equation can be bounded with a constant independent of inverse powers of the viscosity using only the pressure stabilization. Now, the natural question is whether using only an appropriate pressure stabilization can lead to a method where error bounds of this kind can be derived. This question is answered positively in [83]. While finishing the present review, we became aware of Ref. [84], where equal order pairs of finite elements with symmetric pressure stabilization are considered for discretizing the evolutionary Navier–Stokes equations. Robust error bounds of order r are proved for the $L^\infty(L^2)$ error of the velocity. We found that essentially the same idea explained in the proof of Theorem 4.15 is applied to obtain the bounds in [84].

A fully discrete version of the following method with symmetric pressure stabilization, see Assumption A1 below for justifying this notion, using the implicit Euler scheme, is studied in [83]: Find $\boldsymbol{u}_h : (0, T] \rightarrow V_{h,r}$,

$p_h \ : \ (0, T] \to Q_{h,r}$ such that

$$(\partial_t \boldsymbol{u}_h, \boldsymbol{v}_h) + \nu(\nabla \boldsymbol{u}_h, \nabla \boldsymbol{v}_h) + b(\boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{v}_h) - (\nabla \cdot \boldsymbol{v}_h, p_h) = (\boldsymbol{f}, \boldsymbol{v}_h),$$
$$(\nabla \cdot \boldsymbol{u}_h, q_h) + \alpha s_{\text{pres}}(p_h, q_h) = 0, \tag{87}$$

where the nonlinear term is defined in (44), α is a parameter whose units are time over length squared (i.e., the units of the inverse of the kinematic viscosity) and the pressure stabilization term should satisfy a number of assumptions.

Remark 4.16 (*Assumptions on the Pressure Stabilization Term*).

A1   The pressure stabilization $s_{\text{pres}} : Q_{h,r} \times Q_{h,r} \to \mathbb{R}$ is a symmetric, positive semi-definite bilinear form.
A2   Continuity. For all $q_h \in Q_{h,r}$, it holds $s_{\text{pres}}(q_h, q_h) \le C\|q_h\|_0^2$ with a constant $C > 0$ independent of h and ν.
A3   Weak consistency. There exists a projection $\Pi_h : Q \to Q_{h,r}$ such that

$$\|q - \Pi_h q\|_0 \le Ch^s \|q\|_s, \quad q \in Q \cap H^s, \quad 0 \le s \le r+1,$$

and

$$s_{\text{pres}}(\Pi_h q, \Pi_h q)^{1/2} \le Ch^s \|q\|_s, \quad q \in Q \cap H^s, \quad 1 \le s \le r.$$

Typically, Π_h is the $L^2(\Omega)$ projection or some interpolant for non-smooth functions like the Scott–Zhang interpolant.

A4   There is an interpolation operator $\mathcal{I}_h^r \ : \ V \to V_{h,r}$ such that for $1 \le s \le r$ and for all $\boldsymbol{v} \in V$ and $q_h \in Q_{h,r}$

$$(\nabla \cdot (\boldsymbol{v} - \mathcal{I}_h^r \boldsymbol{v}), q_h) \le C \left(\sum_{K \in \mathcal{T}_h} h_K^{-2} \|\boldsymbol{v} - \mathcal{I}_h^l \boldsymbol{v}\|_{0,K}^2 + \|\boldsymbol{v} - \mathcal{I}_h^r \boldsymbol{v}\|_1^2 \right)^{\frac{1}{2}} s_{\text{pres}}(q_h, q_h)^{\frac{1}{2}},$$

and

$$\|\boldsymbol{v} - \mathcal{I}_h^r \boldsymbol{v}\|_m \le Ch^{s+1-m} |\boldsymbol{v}|_{s+1} \quad \forall \ \boldsymbol{v} \in H^s(\Omega)^d,$$

for $m \in \{0, 1\}$, $1 \le s \le r$, with $C > 0$ independent of h and ν.

Assumptions A1–A4 are quite similar to those in [85], where finite element methods for the time-dependent Stokes equations are analyzed. Pressure stabilizations can be divided into two classes, e.g., see the survey [71]: residual-based stabilizations and stabilizations that use only the pressure. It turns out that practically all stabilizations of the latter class satisfy Assumptions A1–A4: the stabilization of Brezzi and Pitkäranta [86], the stabilization of Dohrmann and Bochev [87], the orthogonal subscale stabilization proposed in [88], the classical LPS stabilization [89], the term by term local projection stabilization from [73], and the continuous interior penalty (CIP) stabilization proposed in [85,90].

The following error estimate is derived in [83, Theorem 2].

Theorem 4.17 (*Non inf–sup Stable Pairs of Finite Element Spaces with Symmetric Pressure Stabilization*). *Let, for $T > 0$, the solution of (5) be sufficiently regular and let $r \ge 2$. Let Assumptions A1–A4 for the pressure stabilizations be satisfied, and let $\alpha \sim 1$. Then, there holds for $t \in [0, T]$ that*

$$\|\boldsymbol{u}(t) - \boldsymbol{u}_h(t)\|_0 \le M_{\boldsymbol{u},p} \exp\left(\frac{L(T)}{2} \right) h^r + Ch^{r+1} \|\boldsymbol{u}\|_{L^\infty(H^{r+1})},$$

where $M_{\boldsymbol{u},p} > 0$ depends on

$$\left(\|\boldsymbol{u}_0\|_r^2 + \int_0^T \left[\alpha \|p(s)\|_{H^r/\mathbb{R}}^2 + T \|p(s)\|_{H^{r+1}/\mathbb{R}}^2 + T \|\partial_t \boldsymbol{u}(s)\|_r^2 \right. \right.$$
$$\left. \left. + \left(\nu + \alpha^{-1} + |\Omega|^{\frac{4-d}{d}} \|\boldsymbol{u}(s)\|_2^2 \right) \|\boldsymbol{u}(s)\|_{r+1}^2 \right] ds \right)^{1/2}$$

and $L(T)$ is defined by

$$L(T) = 1 + C \int_0^T |\Omega|^{\frac{4-d}{2d}} \left(\|\boldsymbol{u}(s)\|_3 + \alpha |\Omega|^{\frac{4-d}{2d}} \|\boldsymbol{u}(s)\|_2^2 \right) ds.$$

Table 4
Numerical examples for the Navier–Stokes equations, number of degrees of freedom (including Dirichlet nodes) for non inf–sup stable pairs of finite element spaces on triangular grids.

	P_2/P_2	
Level	Velocity	Pressure
3	290	145
4	1090	545
5	4226	2113
6	16642	8321
7	66050	33025

Proof (*Sketch*). For proving Theorem 4.17, u_h is compared with s_h defined as the velocity component of the pair $(s_h, l_h) \in (V_{h,r}, Q_{h,r})$ solving

$$(\nabla s_h, \nabla v_h) - \alpha(l_h, \nabla \cdot v_h) = (\nabla u, \nabla v_h) \quad \forall\, v_h \in V_{h,r},$$
$$(\nabla \cdot s_h, q_h) + \alpha s_{\mathrm{pres}}(l_h, q_h) = 0 \quad \forall\, q_h \in Q_{h,r}.$$

A standard finite element error analysis shows that $\|s_h - u\|_0 + h\|\nabla(s_h - u)\|_0 + \alpha h s_{\mathrm{pres}}(l_h, l_h)^{1/2} \le Ch^{r+1}\|u\|_{r+1}$.

The finite element pressure solution p_h is compared with $\hat{p}_h = \Pi_h p$, where Π_h is the projection in Assumption A3. The same arguments as in the proof of Theorem 4.15 are applied to bound the nonlinear term. Estimating the term $(p - \hat{p}_h, \nabla \cdot \boldsymbol{\phi}_h)$, with $\boldsymbol{\phi}_h = s_h - u_h$, is different than in (86) and it requires to increase the regularity assumptions on the pressure from $H^r(\Omega)$ to $H^{r+1}(\Omega)$, compared with Theorem 4.15. Integrating by parts yields

$$(p - \hat{p}_h, \nabla \cdot \boldsymbol{\phi}_h) = -(\nabla(p - \hat{p}_h), \boldsymbol{\phi}_h) \le Ch^{2r}T\|p\|_{r+1} + \frac{1}{T}\|\boldsymbol{\phi}_h\|_0^2,$$

such that the second term on the right-hand side can be handled with Gronwall's lemma. □

Theorem 4.17 is formulated for quadratic and higher order finite elements. However, as stated in [83, Remark 3], an analogous error bound can be proved also for linear elements. Concerning low order methods, in [91] a weakly divergence-free method using continuous linear elements for the velocity and piecewise constants for the pressure is considered. A stabilization term for the pressure is added to the method that is enhanced with a discrete divergence-free convective field. A priori bounds of order h are obtained for the $L^\infty(L^2)$ error of the velocity with constants independent of inverse powers of the viscosity.

Example 4.18 (*Non inf–sup Stable Pairs of Finite Element Spaces*). Numerical results for method (73) with grad–div and LPS pressure stabilization and for method (87) with only LPS pressure stabilization are presented in Fig. 13. Exactly the same problem and the same meshes are considered as in Example 4.4. Simulations were performed for the P_2/P_2 pair of finite element spaces, see Table 4 for the corresponding numbers of degrees of freedom. For method (73), the grad–div stabilization parameter was set to be $\mu = 0.1$ and the pressure stabilization parameter to be $\tau_{p,K} = 1.0$. The pressure stabilization parameter in method (87) was chosen to be $\tau_{p,K} = 0.1$.

The numerical results in Fig. 13 show for small viscosity coefficients and on coarse grids second order error reduction for the velocity error in $L^\infty(L^2)$, as predicted by numerical analysis. If the spatial grids become sufficiently fine with respect to the viscosity coefficient, an increase of the order can be observed in this example for both methods. □

4.3. Methods of order $r + 1/2$ in $L^\infty(L^2)$

This section surveys methods where the $L^\infty(L^2)$ error of the velocity can be proved to be of order $r + 1/2$ with a robust error bound. These methods are a special realization of the SUPG method, a CIP method, a modification of the method with LPS stabilization of the velocity gradient (79), a similar method with a different type of LPS stabilization, and a method based on $H(\mathrm{div})$-conforming discontinuous Galerkin elements. Except for the last one, all the methods are formulated for equal order pairs of finite element spaces.
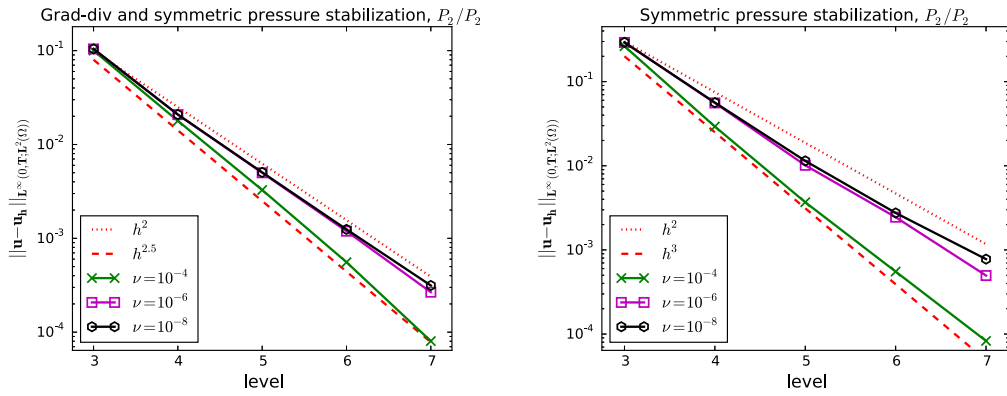
Fig. 13. Non inf–sup stable pair of finite element spaces, method (73) (left) and method (87) (right).

Since the SUPG or streamline diffusion method is one of the most popular finite element discretizations for scalar convection–diffusion equations, it is natural that it is utilized also in the context of the incompressible Navier–Stokes equations (5). In fact, probably the first paper that analyzes a finite element method for the time-dependent Navier–Stokes equations in the standard velocity–pressure formulation for high Reynolds number flows (small ν), paper [13], uses the SUPG method as stabilization.

A main topic of the analysis of the SUPG method for scalar convection–diffusion equations presented in [24] is the choice of the stabilization parameter. By construction, this parameter is a time scale, see Section 3.2. However, the major difficulties of the numerical solution of these equations arise from very small spatial structures, the layers, such that it seems to be reasonable that the stabilization parameter should depend on the local mesh width and not on the length of the time step. The first numerical analysis for stabilization parameters depending on the mesh width was provided in [24], e.g., by considering the continuous-in-time situation, where a time step is not present. However, for the SUPG method in the framework of the incompressible Navier–Stokes equations, an analysis of the continuous-in-time situation does not yet seem to be available, to the best of our knowledge. For this reason, unlike to the other methods discussed in this paper, a fully discrete version of the SUPG stabilization will be discussed here.

In [13], a SUPG method is analyzed for space–time finite elements, where velocity and pressure spaces consist of continuous piecewise linear functions (P_1) in space and discontinuous piecewise linear functions in time. To the best of our knowledge, there seems to be no extension of this analysis available to higher order finite elements in space. Hence, the presentation below will be restricted to this case. A main assumption of the analysis in [13] is that $\Delta t \sim h$.

Since it suffices for the discussion of the estimate of the velocity error in $L^\infty(L^2)$ and it simplifies the presentation considerably, the SUPG method will be considered here only in a form which was analyzed in [13] for $|\nabla \cdot \boldsymbol{u}_h| \geq C > 0$, where C is an arbitrary constant not depending on ν and the mesh width. Let $0 = t^0 < t^1 < \cdots < t^N = T$, denote by $\Omega^n = \Omega \times (t^{n-1}, t^n)$, $\boldsymbol{V}_{h,1}^n = \boldsymbol{V}_{h,1} \times P_1(t^{n-1}, t^n)$, and $Q_{h,1}^n = Q_{h,1} \times P_1(t^{n-1}, t^n)$. Then, the SUPG method reads as follows: Find $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_{h,1}^n \times Q_{h,1}^n$ such that for $n = 0, 1, \ldots, N$ and all $(\boldsymbol{v}_h, q_h) \in \boldsymbol{V}_{h,1}^n \times Q_{h,1}^n$

$$
\begin{aligned}
&(\partial_t \boldsymbol{u}_h, \boldsymbol{v}_h)_{\Omega^n} + a_{\text{SUPG}}(\boldsymbol{u}_h; (\boldsymbol{u}_h, p_h), (\boldsymbol{v}_h, q_h)) + \mu(\nabla \cdot \boldsymbol{u}_h, \nabla \cdot \boldsymbol{v}_h)_{\Omega^n} \\
&\quad + \sum_{K \in \mathcal{T}_h} \delta_K (\partial_t \boldsymbol{u}_h, \partial_t \boldsymbol{v}_h + (\boldsymbol{u}_h \cdot \nabla)\boldsymbol{v}_h + \nabla q_h)_{K^n} + (\boldsymbol{u}_{h,+}^n - \boldsymbol{u}_{h,-}^n, \boldsymbol{v}_{h,+}^n) \\
&= (\boldsymbol{f}, \boldsymbol{v}_h)_{\Omega^n} + \sum_{K \in \mathcal{T}_h} \delta_K (\boldsymbol{f}, \partial_t \boldsymbol{v}_h + (\boldsymbol{u}_h \cdot \nabla)\boldsymbol{v}_h + \nabla q_h)_{K^n},
\end{aligned}
\tag{88}
$$

where $K^n = K \times (t^{n-1}, t^n)$,

$$a_{\text{SUPG}}(\boldsymbol{w}_h; (\boldsymbol{u}_h, p_h), (\boldsymbol{v}_h, q_h))$$

$$= \nu(\nabla \boldsymbol{u}_h, \nabla \boldsymbol{v}_h)_{\Omega^n} + ((\boldsymbol{w}_h \cdot \nabla)\boldsymbol{u}_h, \boldsymbol{v}_h)_{\Omega^n} + \frac{1}{2}(\nabla \cdot \boldsymbol{w}_h, \boldsymbol{u}_h \cdot \boldsymbol{v}_h)_{\Omega^n} - (\nabla \cdot \boldsymbol{v}_h, p_h)_{\Omega^n} + (\nabla \cdot \boldsymbol{u}_h, q_h)_{\Omega^n}$$

$$+ \sum_{K \in \mathcal{T}_h} \delta_K((\boldsymbol{w}_h \cdot \nabla)\boldsymbol{u}_h + \nabla p_h, \partial_t \boldsymbol{v}_h + (\boldsymbol{w}_h \cdot \nabla)\boldsymbol{v}_h + \nabla q_h)_{K^n},$$

and

$$\boldsymbol{v}_{\pm}^n = \lim_{s \to 0\pm} \boldsymbol{v}(t^n + s), \quad \boldsymbol{u}_{-}^0 = \boldsymbol{u}_0.$$

Note that in contrast to formulation (18) for the scalar equations, the temporal derivative of the velocity test function appears. The following error estimate is proved in [13].

Theorem 4.19 (*Error Estimate for the SUPG Method*). *Assume that the solution of the Navier–Stokes equations* (5) *is sufficiently smooth for $T > 0$. Let $\delta_K = \delta \sim h$, $\Delta t \sim h$, $\mu \sim h$, and $\nu \leq h$. Then, for sufficiently small mesh width h, there is a constant C that does not depend on h and ν such that the velocity solution of the SUPG method* (88) *satisfies*

$$\|\boldsymbol{u}(t^n) - \boldsymbol{u}_h^n\|_0 \leq C h^{3/2} \tag{89}$$

for all time instants t^n, $n = 0, \ldots, N$.

Proof (*Sketch*). The error analysis in [13] is performed in the norm

$$\|\|(\boldsymbol{v}, q)\|\|_{S^n}^2 = \frac{1}{2}\left(\|\boldsymbol{v}_{-}^n\|_0^2 + \sum_{l=1}^{n-1}\|\boldsymbol{v}_{+}^l - \boldsymbol{v}_{-}^l\|_0^2 + \|\boldsymbol{v}_{+}^0\|_0^2\right) + \nu\|\nabla \boldsymbol{v}\|_{0,S^n}^2$$

$$+ \delta\|\partial_t \boldsymbol{v} + (\boldsymbol{u}_h \cdot \nabla)\boldsymbol{v} + \nabla q\|_{0,S^n}^2 + \mu\|\nabla \cdot \boldsymbol{v}\|_{0,S^n}^2, \tag{90}$$

with $S^n = \Omega \times (0, t^n)$. It is clear that the bound (89) is obtained as a consequence of bounding the error in the norm (90).

First, a stability estimate for the discrete solution is proved in [13]. This estimate is derived in the usual way by utilizing the discrete solution as test function. As final step, a discrete Gronwall inequality is applied, which requires that the length of the time step, and since $\Delta t \sim h$ is assumed, the mesh width, is sufficiently small.

The proof of the error bound starts with considering the error equation, which is obtained by subtracting the discrete equation (88) from the continuous equation. Splitting the errors $\boldsymbol{u} - \boldsymbol{u}_h = (\boldsymbol{u} - I_h\boldsymbol{u}) - (I_h\boldsymbol{u} - \boldsymbol{u}_h) = \boldsymbol{\eta} - \boldsymbol{\phi}_h$ and $p - p_h = (p - J_h p) - (J_h p - p_h) = (p - J_h p) - \lambda_h$, where I_h and J_h are interpolation operators with the usual interpolation properties, and rearranging terms of the error equation yields

$$\|\|(\boldsymbol{\phi}_h, \lambda_h)\|\|_{S^n}^2 = (\partial_t \boldsymbol{\eta} + (\boldsymbol{u}_h \cdot \nabla)\boldsymbol{\eta}, \boldsymbol{\phi}_h)_{S^n} + (\nabla \cdot \boldsymbol{\eta}, \lambda_h)_{S^n} + \sum_{l=0}^{n-1}\sum_{K \in \mathcal{T}_h} \delta\nu(\Delta\boldsymbol{\eta}, \partial_t\boldsymbol{\phi}_h + (\boldsymbol{u}_h \cdot \nabla)\boldsymbol{\phi}_h + \nabla\lambda_h)_{K^l} + \cdots.$$

Applying integration by parts gives

$$\|\|(\boldsymbol{\phi}_h, \lambda_h)\|\|_{S^n}^2 = -(\boldsymbol{\eta}, \partial_t\boldsymbol{\phi}_h)_{S^n} - \sum_{l=0}^{n-1}(\boldsymbol{\eta}_{-}^{l+1}, \boldsymbol{\phi}_{h,-}^{l+1}) - (\boldsymbol{\eta}_{+}^l, \boldsymbol{\phi}_{h,+}^l)$$

$$- (\boldsymbol{\eta}, (\boldsymbol{u}_h \cdot \nabla)\boldsymbol{\phi}_h)_{S^n} - (\nabla \cdot \boldsymbol{u}_h, \boldsymbol{\eta} \cdot \boldsymbol{\phi}_h)_{S^n} - (\boldsymbol{\eta}, \nabla\lambda_h)_{S^n}$$

$$+ \sum_{l=0}^{n-1}\sum_{K \in \mathcal{T}_h} \delta\nu(\Delta\boldsymbol{\eta}, \partial_t\boldsymbol{\phi}_h + (\boldsymbol{u}_h \cdot \nabla)\boldsymbol{\phi}_h + \nabla\lambda_h)_{K^l} + \cdots.$$

Concentrating on some important terms, using the Cauchy–Schwarz inequality and Young's inequality lead to

$$
\begin{aligned}
\|(\boldsymbol{\phi}_h, \lambda_h)\|_{S^n}^2 \leq{} & \frac{2}{\delta} \|\boldsymbol{\eta}\|_{0,S^n}^2 + \frac{\delta}{8} \|\partial_t \boldsymbol{\phi}_h + (\boldsymbol{u}_h \cdot \nabla)\boldsymbol{\phi}_h + \nabla \lambda_h\|_{0,S^n}^2 \\
& + 2\delta \nu^2 \sum_{l=0}^{n-1} \sum_{K \in \mathcal{T}_h} \|\Delta \boldsymbol{\eta}\|_{0,K^l}^2 + \frac{\delta}{8} \|\partial_t \boldsymbol{\phi}_h + (\boldsymbol{u}_h \cdot \nabla)\boldsymbol{\phi}_h + \nabla \lambda_h\|_{0,S^n}^2 + \cdots .
\end{aligned}
\tag{91}
$$

The second and fourth terms on the right-hand side of (91) are absorbed by the contribution of the SUPG term on the left-hand side. Using the interpolation property of $\|\boldsymbol{\eta}\|_{0,S^n}$ and the choice of δ gives that the first term is of order h^3, with respect to the square of the error. The asymptotic of δ and the condition that $\nu \leq h$ lead also for the third term of (91) to third order.

There are even more terms leading to order $3/2$, but in this sketch of the proof, it is concentrated to terms that are related to the SUPG stabilization. The final step of the proof is again the application of a discrete Gronwall lemma. $\square$

An error analysis of the SUPG method applied to the non-conforming Crouzeix–Raviart finite element of lowest order and combined with the backward Euler scheme as temporal discretization is presented in [92]. The SUPG stabilization parameter depends on the length of the time step. Robust estimates are derived. For a certain relation of the mesh width, the length of the time step, and a stabilization parameter for a velocity jump term, the velocity error in $L^\infty(L^2)$ is proved to be of order $4/3$.

The stabilization terms of the SUPG method are part of a two-scale residual-based variational multiscale (VMS) method for the simulation of turbulent incompressible flows proposed in [14]. This VMS method possesses two more stabilization terms. Recent studies of a two-dimensional flow at high Reynolds number and at three-dimensional turbulent channel flows in [93,94] show that the results obtained with the SUPG method and the residual-based VMS method are similarly accurate.

In [90], a CIP finite element method to approximate the solution of the Navier–Stokes equations (5) is considered. Equal order approximations of the velocity and pressure ($\boldsymbol{u}_h, p_h$), belonging to the spaces $V_{h,r}$ and $Q_{h,r}$, respectively, are taken and the boundary conditions are weakly imposed in the method. Incompressibility, convective dominance, and pressure are stabilized by adding the following stabilization terms to the Galerkin formulation:

$$
\begin{aligned}
j(\boldsymbol{u}_h, \boldsymbol{v}_h) &:= \sum_{K \in \mathcal{T}_h} \sum_{F \in \partial K} \int_F h_F^2 \, [\![\nabla \boldsymbol{u}_h]\!]_F : [\![\nabla \boldsymbol{v}_h]\!]_F \; d\boldsymbol{s}, \\
j_{\boldsymbol{u}_h}(\boldsymbol{u}_h, \boldsymbol{v}_h) &:= \sum_{K \in \mathcal{T}_h} \sum_{F \in \partial K} \int_F h_F^2 |I_h^1 \boldsymbol{u}_h \cdot \boldsymbol{n}_F|^2 \, [\![\nabla \boldsymbol{u}_h]\!]_F : [\![\nabla \boldsymbol{v}_h]\!]_F \; d\boldsymbol{s}, \\
\iota(p_h, q_h) &= \sum_{K \in \mathcal{T}_h} \sum_{F \in \partial K} \int_F h_F^2 \, [\![\nabla p_h]\!]_F \cdot [\![\nabla q_h]\!]_F \; d\boldsymbol{s},
\end{aligned}
\tag{92}
$$

where the same notations as for convection–diffusion equations are used. In addition, $I_h^1 \boldsymbol{u}_h$ denotes the interpolation of $\boldsymbol{u}_h$ onto the space $V_{h,1}$ of continuous piecewise linear polynomials.

The robust, with respect to ν, stability of the method relies on the fact that the stabilization terms (92) control some interpolation errors of the streamline derivative, the divergence, and the pressure gradient. More precisely, let $r_h \boldsymbol{v}$ denote the Oswald quasi-interpolant, compare [90, Definition 1], then the following bounds hold for any $\boldsymbol{w}_h \in V_{h,1}$, $\boldsymbol{v}_h \in V_{h,r}$, $q_h \in Q_{h,r}$, see [90, Lemma 3])

$$
\begin{aligned}
\|h^{1/2}((\boldsymbol{w}_h \cdot \nabla)\boldsymbol{v}_h - r_h((\boldsymbol{w}_h \cdot \nabla)\boldsymbol{v}_h))\|_0^2 &\leq C_1 j_{\boldsymbol{w}_h}(\boldsymbol{v}_h, \boldsymbol{v}_h), \\
\|h^{1/2}(\nabla \cdot \boldsymbol{v}_h - r_h(\nabla \cdot \boldsymbol{v}_h))\|_0^2 &\leq C_2 j(\boldsymbol{v}_h, \boldsymbol{v}_h), \\
\|h^{1/2}(\nabla q_h - r_h(\nabla q_h))\|_0^2 &\leq C_3 \iota(q_h, q_h),
\end{aligned}
\tag{93}
$$

where C_1, C_2, C_3 are positive constants.[4]

---

[4] We remark that the error analysis of the CIP method presented here follows closely [90], although, contrary to most methods in the present paper, the analysis and the corresponding estimates are not necessarily dimensionally correct.

Concerning the viscosity-independent control of the divergence, the following result is applied in the error analysis of the method, see [90, Lemma 7],

$$C\|h^{1/2}\nabla \cdot \boldsymbol{u}_h\|_0^2 \le j(\boldsymbol{u}_h, \boldsymbol{u}_h) + \iota(p_h, p_h) + \|\boldsymbol{u}_h \cdot \boldsymbol{n}\|_{0,\partial\Omega}^2. \tag{94}$$

Let $\varPi_h^r$ denote both, the $L^2(\Omega)$ orthogonal projection onto $V_{h,r}$ and $Q_{h,r}$, then the following bound, analogous to (84), is also essential in the error analysis, see [90, Corollary 1] and [78, Theorem 2.2],

$$\|\varPi_h^r(\boldsymbol{u} \cdot \boldsymbol{v}_h) - \boldsymbol{u} \cdot \boldsymbol{v}_h\|_0 \le Ch\|\nabla\boldsymbol{u}\|_{L^\infty}\|\boldsymbol{v}_h\|_0, \quad \forall\, \boldsymbol{u} \in W^{1,\infty}(\Omega)^d,\ \boldsymbol{v}_h \in V_{h,r}. \tag{95}$$

In the following theorem, whose proof can be found in [90, Theorem 2, Corollary 3], the rate of convergence of the velocity is stated.

Theorem 4.20 (*Error Estimate for the CIP Method*). *Let the solution of* (5) *be sufficiently smooth and assume $v < h$. Then, the following bound holds for the velocity approximation $\boldsymbol{u}_h$ of the CIP method with stabilization terms* (92)

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^\infty(L^2)} \le C(T, \boldsymbol{u}, p)h^{r+1/2},$$

where $C(T, \boldsymbol{u}, p)$ does not depend on inverse powers of v.

Proof (*Sketch*). For proving the statement of Theorem 4.20, $\boldsymbol{u}_h$ is compared with $\varPi_h^r\boldsymbol{u}$ and p_h with $\varPi_h^r p$. Denote $\boldsymbol{\phi}_h = \varPi_h^r\boldsymbol{u} - \boldsymbol{u}_h$, $\boldsymbol{\eta} = \boldsymbol{u} - \varPi_h^r\boldsymbol{u}$, and $\lambda_h = \varPi_h^r p - p_h$. Some details concerning the bounds of the more involved terms appearing in the error equation will be presented.

First, terms that come from a decomposition of the nonlinear terms in the error equations are considered. One of them is $(\boldsymbol{\eta}, (\boldsymbol{u}_h \cdot \nabla)\boldsymbol{\phi}_h)$. Adding and subtracting $I_h^1\boldsymbol{u}_h$ and using the orthogonality properties of the $L^2(\Omega)$ projection with respect to the range of r_h yield

$$(\boldsymbol{\eta}, (\boldsymbol{u}_h \cdot \nabla)\boldsymbol{\phi}_h) = (\boldsymbol{\eta}, ((\boldsymbol{u}_h - I_h^1\boldsymbol{u}_h) \cdot \nabla)\boldsymbol{\phi}_h) + (\boldsymbol{\eta}, (I_h^1\boldsymbol{u}_h \cdot \nabla)\boldsymbol{\phi}_h - r_h(I_h^1\boldsymbol{u}_h \cdot \nabla\boldsymbol{\phi}_h)). \tag{96}$$

For the first term on the right-hand side, one obtains after some manipulations, including inverse inequalities and the $L^\infty(\Omega)$ stability of the $L^2(\Omega)$ projection,

$$(\boldsymbol{\eta}, ((\boldsymbol{u}_h - I_h^1\boldsymbol{u}_h) \cdot \nabla)\boldsymbol{\phi}_h) \le C\|\nabla\boldsymbol{u}\|_{L^\infty}\left(\|\boldsymbol{\phi}_h\|_0^2 + h^{-1/2}\|\boldsymbol{\eta}\|_0^2\right).$$

The second term on the right-hand side of (96) is a truncation error, even of higher order than $2r+1$. For this term, applying (93), one can prove

$$(\boldsymbol{\eta}, I_h^1\boldsymbol{u}_h \cdot \nabla\boldsymbol{\phi}_h - r_h(I_h^1\boldsymbol{u}_h \cdot \nabla\boldsymbol{\phi}_h)) \le C\varepsilon_1^{-1}h^{-1}\|\boldsymbol{\eta}\|_0^2 + \varepsilon_1 C_1 j_{\boldsymbol{u}_h}(\boldsymbol{\phi}_h, \boldsymbol{\phi}_h),$$

where ε_1 is any positive parameter. Note that the square root of the first term on the right-hand side is of order $r + 1/2$.

For the term $(\nabla \cdot \boldsymbol{u}_h, \boldsymbol{\eta} \cdot \boldsymbol{\phi}_h)$, appearing also in the error equation, the control of the divergence given by (94) is applied.

Bounding the term $(\nabla \cdot \boldsymbol{u}_h, \boldsymbol{u} \cdot \boldsymbol{\phi}_h)$ covers three pages in the proof of [90, Theorem 2]. The first observation is that one can take $\boldsymbol{v}_h = \boldsymbol{0}$ and $q_h = \varPi_h^r(\boldsymbol{u} \cdot \boldsymbol{\phi}_h)$ in the error equation, which gives

$$(\nabla \cdot \boldsymbol{u}_h, \varPi_h^r(\boldsymbol{u} \cdot \boldsymbol{\phi}_h)) = \left(\boldsymbol{u}_h \cdot \boldsymbol{n}, \varPi_h^r(\boldsymbol{u} \cdot \boldsymbol{\phi}_h)\right)_{\partial\Omega} - \iota(p_h, \varPi_h^r(\boldsymbol{u} \cdot \boldsymbol{\phi}_h)).$$

Then, adding and subtracting $\varPi_h^r(\boldsymbol{u} \cdot \boldsymbol{\phi}_h)$ and using this equality yield

$$(\nabla \cdot \boldsymbol{u}_h, \boldsymbol{u} \cdot \boldsymbol{\phi}_h) = (\nabla \cdot \boldsymbol{u}_h, \boldsymbol{u} \cdot \boldsymbol{\phi}_h - \varPi_h^r(\boldsymbol{u} \cdot \boldsymbol{\phi}_h)) + \left(\boldsymbol{u}_h \cdot \boldsymbol{n}, \varPi_h^r(\boldsymbol{u} \cdot \boldsymbol{\phi}_h)\right)_{\partial\Omega} - \iota(p_h, \varPi_h^r(\boldsymbol{u} \cdot \boldsymbol{\phi}_h)).$$

After some work and with the help of (95), one finally arrives at

$$(\nabla \cdot \boldsymbol{u}_h, \boldsymbol{u} \cdot \boldsymbol{\phi}_h) \le C\varepsilon_2 \left(\iota(\varPi_h^r p, \varPi_h^r p) + \iota(\lambda_h, \lambda_h)\right) + C\varepsilon_2^{-1}h\|\nabla\boldsymbol{u}\|_{L^\infty}^2\|\boldsymbol{\phi}_h\|_0^2 + C\varepsilon_2^{-1}\|\boldsymbol{u}\|_{L^\infty}^2 j(\boldsymbol{\phi}_h, \boldsymbol{\phi}_h),$$

for any positive parameter ε_2. Then, it can be observed that, apart from applying (95), the stabilization terms included to achieve the control over the divergence and the gradient of the pressure are essential to finish the estimate of this term.

Finally, another term appearing in the error equation in the proof of [90, Theorem 2] can be bounded as follows

$$
\sum_{K \in \mathcal{T}_h} \sum_{F \in \partial K} \int_{\partial F} h_F^2 |I_h^1 \boldsymbol{u}_h \cdot \boldsymbol{n}_F|^2 \left[\left|\nabla \Pi_h^r \boldsymbol{u}\right|\right]_F : \left[\left|\nabla \boldsymbol{\phi}_h\right|\right]_F \; d\boldsymbol{s}
$$
$$
\leq \varepsilon_3 j_{\boldsymbol{u}_h}(\boldsymbol{\phi}_h, \boldsymbol{\phi}_h) + C \varepsilon_3^{-1} \left(h \|\nabla \boldsymbol{u}\|_{L^\infty}^2 \|\boldsymbol{\phi}_h\|_0^2 + \left(\|\boldsymbol{u}\|_{L^\infty}^2 + h^2 \|\nabla \boldsymbol{u}\|_{L^\infty}^2 \right) \iota(\Pi_h^r \boldsymbol{u}, \Pi_h^r \boldsymbol{u}) \right),
$$

for any positive parameter ε_3. In this inequality, one can observe the application of the stabilization term added to control the streamline derivative. $\quad\square$

Fully discretized schemes with a CIP method as stabilization are analyzed in [95] for the evolutionary Oseen equations (29). Concretely, the implicit Euler scheme, BDF2, and a pressure-projection method based on the implicit Euler scheme are studied. Robust error estimates for the velocity error in $L^\infty(L^2)$ of order $r + 1/2$ are proved.

Method (79), which is of order r, see Theorem 4.15, can be modified such that the order of error reduction is increased by $1/2$ in the convection-dominated case. This modification consists simply in using different scalings of the LPS stabilization parameters than for the method from Theorem 4.15. The analysis of the modified method requires a little bit higher regularity assumptions on the solution of (5) than for the error bound presented in Theorem 4.15. A fully discrete version of the modified method is analyzed in [72].

Theorem 4.21 (*Non inf–sup Stable Pairs of Finite Element Spaces with Velocity Gradient LPS Term and Modified Stabilization Parameters*). *Let $T > 0$, assume sufficient regularity of the solution of (5), and let $r \geq 2$. Furthermore, assume that the LPS pressure stabilization parameters in method (79) satisfy*

$$
\alpha_1 h_K \leq \tau_{p,K} \leq \alpha_2 h_K,
$$

and the LPS velocity gradient stabilization parameters satisfy

$$
\beta_1 h_K \leq \tau_{v,K} \leq \beta_2 h_K, \tag{97}
$$

with non-negative constants $\alpha_1, \alpha_2, \beta_1$, and β_2. Assume that $\nu \leq \|\boldsymbol{u}\|_{L^\infty} h$, and

$$
8(1 + c_s^2) \|\hat{\boldsymbol{u}}_h^{n+1}\|_{L^\infty}^2 \frac{\alpha_1}{\beta_1} \leq \frac{1}{16}, \tag{98}
$$

where c_s is the stability constant of σ_h^{r-1} in $L^2(\Omega)$. Then, there exists a positive constant $M_{\boldsymbol{u},p}$ depending on $\|\boldsymbol{u}_0\|_r$ h and

$$
\left(\int_0^T \left(\left(\alpha_2 + \beta_1^{-1}\right) \|p(s)\|_{H^{r+1}/\mathbb{R}}^2 + T h \|\partial_t \boldsymbol{u}(s)\|_{r+1}^2 \right. \right.
$$
$$
\left. \left. + \left(\|\boldsymbol{u}\|_{L^\infty(L^\infty)} + |\Omega|^{\frac{4-d}{d}} \left(\beta_1^{-1} \|\boldsymbol{u}(s)\|_2^2 + T h \|\boldsymbol{u}(s)\|_3^2 \right) \right) \|\boldsymbol{u}(s)\|_{r+1}^2 \right) ds \right)^{1/2}
$$

such that for $t \in [0, T]$ the error estimate

$$
\|\boldsymbol{u}(t) - \boldsymbol{u}_h(t)\|_0 \leq M_{\boldsymbol{u},p} \exp\left(\frac{L(T)}{2}\right) h^{r+1/2} + C h^{r+1} \|\boldsymbol{u}\|_{L^\infty(H^{r+1})},
$$

holds with $L(T)$

$$
L(T) = 1 + C \int_0^T |\Omega|^{\frac{4-d}{2d}} \left(\|\boldsymbol{u}(s)\|_3 \left(1 + \alpha_2 h \, |\Omega|^{\frac{4-d}{2d}} \|\boldsymbol{u}(s)\|_3 \right) \right) ds.
$$

Proof (*Sketch*). Very similar arguments can be applied as in the proof for the fully discrete scheme from [72, Theorem 6.1]. Here, only the main differences with respect to the proof of Theorem 4.15 will be mentioned, which allow to prove order $r + 1/2$ for the $L^\infty(L^2)$ error of the velocity. The same notations are used as in the proof of Theorem 4.15.

The major new aspects occur in bounding the difference of the nonlinear terms arising in the error equation. For the method from Theorem 4.15, this bound was presented in some detail in the sketch of the proof. One can observe that only the LPS pressure stabilization was utilized for obtaining (85) but not LPS velocity gradient
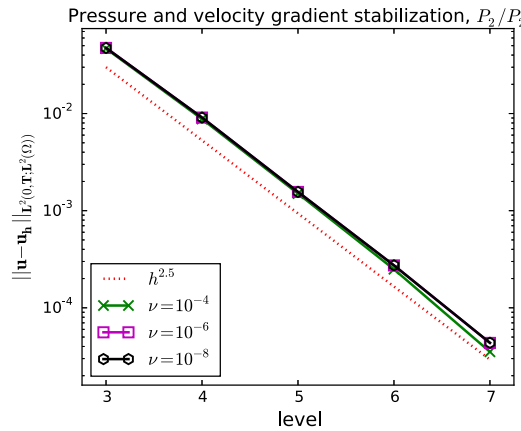
Fig. 14. LPS stabilization of the velocity gradient with modified stabilization parameters.

stabilization. For the present method, the whole analysis for the difference of the nonlinear terms in [72] covers more than four pages. This difference is decomposed into several terms. For one of them, one can derive the estimate

$$\left| b(\boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{\phi}_h) - b(\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{u}}_h, \boldsymbol{\phi}_h) \right|$$
$$\leq \left(\frac{1}{4T} + C \left(\|\nabla \hat{\boldsymbol{u}}_h\|_{L^\infty} + \alpha_2 h \|\hat{\boldsymbol{u}}_h\|_{W^{1,\infty}}^2 \right) \right) \|\boldsymbol{\phi}_h\|_0^2 + \frac{1}{8} s_{\text{pres}}(\lambda_h, \lambda_h) + \frac{1}{8} \sum_{K \in \mathcal{T}_h} \tau_{\nu,K} \|\sigma_h^*(\nabla \boldsymbol{\phi}_h)\|_{0,K}^2$$
$$+ C h^{2r+1} \left(\alpha_2 \|p\|_{r+1}^2 + |\Omega|^{\frac{4-d}{d}} \left(\beta_1^{-1} \|\boldsymbol{u}(s)\|_2^2 + T h \|\boldsymbol{u}(s)\|_3^2 \right) \|\boldsymbol{u}\|_{r+1}^2 \right).$$

Note that the increased regularity for the pressure is needed at this step in order to gain half an order in the fourth term of the error bound. The second term of the bound contains the pressure stabilization and the third term the velocity stabilization. The first term is handled as usual by applying a Gronwall lemma. Also for bounding another term of the decomposition of the difference of the nonlinear terms, one needs both stabilizations included in the method to obtain

$$|b(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{\phi}_h) - b(\hat{\boldsymbol{u}}_h, \hat{\boldsymbol{u}}_h, \boldsymbol{\phi}_h)| \leq C \, |\Omega|^{\frac{4-d}{d}} \left((\beta_1 h)^{-1} \|\boldsymbol{u}(s)\|_2^2 + T \|\boldsymbol{u}(s)\|_3^2 \right) \|\boldsymbol{u} - \hat{\boldsymbol{u}}_h\|_0^2$$
$$+ \frac{1}{4T} \|\boldsymbol{\phi}_h\|_0^2 + \frac{1}{4} \sum_{K \in \mathcal{T}_h} \tau_{\nu,K} \|\sigma_h^*(\nabla \boldsymbol{\phi}_h)\|_{0,K}^2.$$

The first term on the right-hand side of is of order $\mathcal{O}(h^{2r+1})$ under assumption (97) for the stabilization parameters $\tau_{\nu,K}$. The last term on the right-hand side shows that the stabilization of the velocity gradient is again utilized for this estimate.

For bounding the truncation error, an increased regularity of the temporal derivative of the velocity is needed to get $\|\partial_t \boldsymbol{u} - \partial_t \hat{\boldsymbol{u}}_h\|_0 \leq C h^{r+1} \|\partial_t \boldsymbol{u}\|_{r+1}$.

Condition (98) is needed to bound the term $2 s_{\text{pres}}(\sigma_h^r(\hat{\boldsymbol{u}} \cdot \boldsymbol{\phi}_h), \sigma_h^r(\hat{\boldsymbol{u}} \cdot \boldsymbol{\phi}_h))$ which appears in the proof, and, in particular, to absorb part of this bound in the left-hand side of the error equation. □

Example 4.22 (*LPS Method with Velocity Gradient Fluctuations and Modified Stabilization Parameters*). Again, the same example and setup for the numerical studies are considered as in Example 4.4.

Numerical results are presented in Fig. 14. For the LPS method from Theorem 4.21, the P_2/P_2 pair of finite element spaces was utilized as spatial discretization, see Table 2 for the corresponding numbers of degrees of freedom. The parameters were chosen to be $\tau_{p,K} = 0.001 h_K$ for the pressure stabilization and $\tau_{\nu,K} = 0.01 h_K$ for the velocity gradient stabilization. The linear systems of equations were solved iteratively with the flexible GMRES method and a coupled multigrid preconditioner. It can be clearly seen from the results that the method behaves robustly and that the order of reduction of the velocity error in $L^\infty(L^2)$ is 2.5. □

In [96], a different kind of LPS stabilization is analyzed using, as in [72], equal order conforming elements for velocity and pressure. LPS stabilizations for the velocity gradient and for the pressure are added to the standard

Galerkin formulation. The LPS stabilization utilizes a mesh of macro elements for the projection (two-level LPS method). For this two-level method, the error bounds in [96] depend both on the cell diameters of the original and the macro element meshes, which are assumed to be comparable. Following the error analysis in [96], one can deduce that, chosen the stabilization parameters to be both of the size of a macro element cell, a rate of error reduction of order $r + 1/2$ for piecewise polynomials of degree r can be proved. The key points in the error analysis are similar to those sketched in the proof of Theorem 4.21, with some differences corresponding to the different kinds of LPS stabilization.[5]

The last method that will be discussed in this section is a $H(\mathrm{div})$-conforming discontinuous Galerkin method with upwind stabilization.[6] An error analysis of this method for the semi-discrete time-dependent Navier–Stokes equations is presented in [64]. The finite element velocity space $V_{h,r}$ is the space of functions in $H(\mathrm{div}, \Omega)$ that restricted to a mesh cell $K \in \mathcal{T}_h$ belong to the space $RT_r(K)$ (Raviart–Thomas) or $BDM_r(K)$ (Brezzi–Douglas–Marini), $r \geq 1$. The corresponding pressure space, $Q_{h,\tilde{r}}$ consists of discontinuous polynomials of degree $\tilde{r} = r$ or $\tilde{r} = r - 1$, respectively. For these spaces, it holds $\nabla \cdot V_{h,r} \subseteq Q_{h,\tilde{r}}$. Denote by ∇_h the broken gradient, by $\mathcal{F}_h$ the set of all facets, and by $\mathcal{F}_h^i$ and $\mathcal{F}_h^\partial$ the subsets of interior and boundary facets, respectively. The jump $[\![\phi]\!]_F$ of a function ϕ and its average $\{\phi\}_F$ across the interior facet $F \in \mathcal{F}_h^i$ are defined by

$$[\![\phi]\!]_F = \phi^+ - \phi^-, \quad \{\phi\}_F = \frac{\phi^+ + \phi^-}{2}.$$

For the boundary facet $F \in \mathcal{F}_h^\partial$, one sets

$$[\![\phi]\!]_F = \{\phi\}_F = \phi.$$

The method analyzed in [64] reads as follows: Find $u_h : (0, T] \to V_{h,r}$, $p_h : (0, T] \to Q_{h,\tilde{r}}$ satisfying for all $(v_h, q_h) \in (V_{h,r}, Q_{h,\tilde{r}})$

$$(\partial_t u_h, v_h) + v a_h(u_h, v_h) + b_h(v_h, p_h) - b_h(u_h, q_h) + c_h(u_h, u_h, v_h) = (f, v_h), \tag{99}$$

with

$$a_h(u_h, v_h) = (\nabla_h u_h, \nabla_h v_h) - \sum_{F \in \mathcal{F}_h} \int_F \left[(\{\nabla_h u_h\}_F \, n_K \cdot [\![v_h]\!]_F) + ([\![u_h]\!]_F \cdot \{\nabla_h v_h\}_F \, n_K) \right] ds$$

$$+ \sum_{F \in \mathcal{F}_h} \int_F \left(\frac{\sigma}{h_K} [\![u_h]\!]_F \cdot [\![v_h]\!]_F \right) ds,$$

$$b_h(u_h, q_h) = -(q_h, \nabla_h \cdot u_h),$$

$$c_h(u_h, u_h, v_h) = ((u_h \cdot \nabla_h) u_h, v_h) - \sum_{F \in \mathcal{F}_h^i} \int_F (u_h \cdot n_K) [\![u_h]\!]_F \{v_h\}_F \, ds$$

$$+ \sum_{F \in \mathcal{F}_h^i} \int_K \frac{1}{2} |(u_K \cdot n_K)| [\![u_h]\!]_F [\![v_h]\!]_F \, ds,$$

where σ is a positive parameter that has to be sufficiently large. The discretization of the viscous term is the standard symmetric interior penalty form. In the discretization of the convective term, an upwind flux is utilized in the last term.

The proof of the following theorem can be found in [64].

Theorem 4.23 ($H(\mathrm{div})$-*conforming Discontinuous Galerkin Method with Upwind*). *Let $T > 0$, assume sufficient regularity of the solution of* (5), *and assume $v < h$. Then, the following bound holds for the velocity approximation*

---

[5] In [96], also the case of a one-level LPS method is considered, in which the projection is defined on the same mesh and the velocity space has to contain appropriate bubble functions. However, the presented error analysis does not fit in this case. The reason is that the application of [96, Lemma 2.4] requires to have equal velocity and pressures spaces. The same kind of argument was applied in the proof of Theorem 4.15, see the comment after Eq. (85) about the application of (84), which is valid only in case of equal velocity and pressures spaces. Eq. (84) is the analog to [96, (2.17), Lemma 2.4].

[6] The presentation of this method follows closely [64] and it is, in contrast to most of the other methods in the present paper, not necessarily dimensionally correct.

$\boldsymbol{u}_h$ computed with method (99)

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^\infty(L^2)} \le (E(\boldsymbol{u})F(\boldsymbol{u}) + G(\boldsymbol{u}))^{1/2} h^{r+1/2},$$

where

$$E(\boldsymbol{u}) = C e^{T + \|\nabla \boldsymbol{u}\|_{L^1(L^\infty)}},$$

$$F(\boldsymbol{u}) = h\|\partial_t \boldsymbol{u}\|_{L^2(H^{r+1})}^2 + \left(T + (h+1)\|\boldsymbol{u}\|_{L^1(W^{1,\infty})}\right)\|\boldsymbol{u}\|_{L^\infty(H^{r+1})}^2,$$

$$G(\boldsymbol{u}) = C(h+T)\|\boldsymbol{u}\|_{L^\infty(H^{r+1})}^2.$$

Proof (*Sketch*)**.** For proving the statement of Theorem 4.23, the finite element velocity $\boldsymbol{u}_h$ is compared with the Raviart–Thomas interpolation $I_{\mathrm{RT}}\boldsymbol{u} \in \boldsymbol{V}_{h,r}$. Thanks to the orthogonality properties of the Raviart–Thomas interpolation operator, the following bound can be obtained, see [64, Lemma 5.3],

$$((\boldsymbol{u} \cdot \nabla_h)\boldsymbol{e}_h, \boldsymbol{\eta}) \le C\|\nabla \boldsymbol{u}\|_{L^\infty}\left(\|\boldsymbol{\eta}\|_0^2 + \|\boldsymbol{e}_h\|_0^2\right), \tag{100}$$

where $\boldsymbol{e}_h = \boldsymbol{u}_h - I_{\mathrm{RT}}\boldsymbol{u}$ and $\boldsymbol{\eta} = \boldsymbol{u} - I_{\mathrm{RT}}\boldsymbol{u}$. More precisely, for the proof of (100) the mean value $\langle \boldsymbol{u} \rangle_K$ of $\boldsymbol{u}$ at each mesh cell $K \in \mathcal{T}_h$ is added and subtracted to the velocity. The following bound holds

$$\|\boldsymbol{u} - \langle \boldsymbol{u} \rangle_K\|_{L^\infty(K)} \le Ch_K\|\nabla \boldsymbol{u}\|_{L^\infty(K)}. \tag{101}$$

Due to the properties of $I_{\mathrm{RT}}\boldsymbol{u}$, one has for any $K \in \mathcal{T}_h$

$$(((\langle \boldsymbol{u} \rangle_K \cdot \nabla_h)\boldsymbol{e}_h, \boldsymbol{\eta})_K = 0.$$

Then, applying (101) and the inverse inequality (8) yields

$$((\boldsymbol{u} \cdot \nabla_h)\boldsymbol{e}_h, \boldsymbol{\eta}) = \sum_{K \in \mathcal{T}_h} \int_K ((\boldsymbol{u} - \langle \boldsymbol{u} \rangle_K) \cdot \nabla_h)\boldsymbol{e}_h \cdot \boldsymbol{\eta}\ d\boldsymbol{x} \le C\|\nabla \boldsymbol{u}\|_\infty \|\boldsymbol{\eta}\|_0 \|\boldsymbol{e}_h\|_0,$$

from which (100) is reached by applying Young's inequality.

For the estimation of the nonlinear term, (100) is used to achieve the following result, see [64, Lemma 5.4],

$$c_h(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{e}_h) - c_h(\boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{e}_h) \le C(1 + h^{-1})\|\nabla \boldsymbol{u}\|_{L^\infty}\left(\|\boldsymbol{\eta}\|_0^2 + h^2\|\nabla_h \boldsymbol{\eta}\|_0^2\right) + C\|\nabla \boldsymbol{u}\|_{L^\infty}\|\boldsymbol{e}_h\|_0^2. \quad \square$$

Numerical studies with a smooth solution presented in [64] used the BDM_1/P_0 pair of spaces. In the case of small viscosity coefficients and coarse grids, orders of error reduction for the velocity error in $L^2(\Omega)$ at the final time are reported that are larger than 1.5 but notably smaller than 2.

4.4. Methods with non-robust error analysis that behave numerically in a robust way

So far, only methods have been discussed for which robust finite element error estimates can be proved. However, there are much more methods that behave in practice in a robust way, in particular, all methods which are utilized for the simulation of incompressible turbulent flows. This section provides a brief survey on such methods for which error estimates are available that are, however, not robust. Since there are no robust error estimates for these methods, there is no analytic information on the order of error reduction with respect to the kinetic energy for high Reynolds number flows simulated on coarse grids.

4.4.1. Large Eddy Simulation and other methods based on spatial averaging

Large Eddy Simulation (LES) is a widely used approach for simulating turbulent incompressible flows. LES performs a scale separation in resolved (large) and unresolved (small, subgrid) scales by an averaging (filtering) in space.

The simplest and one of the most popular LES models is the Smagorinsky model [97], which adds to the momentum equation of the Navier–Stokes equations (5) the nonlinear viscous term

$$\left(C_S \delta^2 \|\mathbb{D}\boldsymbol{u}\|_F \mathbb{D}\boldsymbol{u}, \mathbb{D}\boldsymbol{v}\right), \tag{102}$$

where $C_S > 0$ is a user-chosen constant, $\delta > 0$ is the filter width, which is in practice connected to the local mesh size, $\mathbb{D}\boldsymbol{u} = (\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^T)/2$ is the velocity deformation tensor, and $\|\cdot\|_F$ denotes the Frobenius norm of a tensor.

The continuous Smagorinsky model is mathematically well understood. Based on the property that the nonlinear viscous operator is strongly monotone, it is shown in [98] that the variational form of the Smagorinsky model admits a stable unique solution in two and three dimensions for arbitrarily long time intervals. Concerning the finite element error analysis, [99] studies the error of the solution of a discrete Smagorinsky-type model to the solution of the variational Smagorinsky model. The discrete model adds another viscous term of the form $(\nu_0(\delta)\mathbb{D}\boldsymbol{u}, \mathbb{D}\boldsymbol{v})$, $\nu_0(\delta) > 0$, to the Smagorinsky model. Then, a robust estimate for the velocity is derived, where the error bound depends on inverse powers of $\nu_0(\delta)$. For sufficiently smooth solutions of the Smagorinsky model, also an estimate for $\nu_0(\delta) = 0$ is proved, where the error bounds, however, depend on inverse powers of δ. For constant δ, even for small viscosity coefficients and on coarse grids, an optimal order of convergence of the velocity error in $L^\infty(L^2)$ to the solution of the continuous Smagorinsky model is observed. In [100, Chapter 10], a Smagorinsky-type model is studied, in which $\|\mathbb{D}\boldsymbol{u}\|_F$ in (102) is replaced by $\|\mathbb{D}\overline{\boldsymbol{u}}\|_F$, where $\overline{\boldsymbol{u}} = g * \boldsymbol{u}$ is a filtered (smoothed) velocity field with an appropriate filter g. In the finite element analysis, the error with respect to the solution of the Navier–Stokes equations is considered and non-robust error estimates are derived, see [100, Thm. 10.3]. Recently, a robust estimate of the error between the numerical solution with the Smagorinsky model and the solution of the Navier–Stokes equations of order $3/2$ was proved in [101] under the assumptions that $\boldsymbol{V}_{h,r}$ and $Q_{h,l}$ is an inf–sup stable pair of spaces with $\nabla \cdot \boldsymbol{V}_{h,r} \subseteq Q_{h,l}$. For the Scott–Vogelius pair of spaces $P_r/P_{r-1}^{\mathrm{disk}}$, $r \geq d$, on barycentric-refined grids, the reduction to rate $3/2$ compared with the optimal order $r + 1$ is $r + 1 - 3/2 = r - 1/2$. In [101], it is referred to a lowest order pair that satisfies the assumptions, which is proposed in [102]. In this pair, the pressure is piecewise constant and for defining the velocity space, a barycentric refinement of the grid for the pressure is applied. Then, the velocity space is piecewise linear on the barycentric-refined grid, i.e., in each mesh cell of the original grid there are $(d + 1)^2$ degrees of freedom of the velocity. In addition, some degrees of freedom correspond to normal derivatives on facets. It turns out that it is not possible to impose Dirichlet conditions in the usual way, such that the introduction of a penalty term for the tangential component is necessary.

A method that applies the spatial filtering as a post-processing step is studied in [103]. In the first step of this method, an intermediate velocity field $\boldsymbol{w}_h$ is computed with the Galerkin finite element method. Then, spatial filtering is applied to $\boldsymbol{w}_h$ via the discrete Stokes filter: Find $G_h \boldsymbol{w}_h = \overline{\boldsymbol{w}_h} \in \boldsymbol{V}_{h,r}$ and $\overline{r_h} \in Q_{h,l}$ such that

$$
\begin{aligned}
\delta^2(\nabla \overline{\boldsymbol{w}_h}, \nabla \boldsymbol{v}_h) + (\overline{\boldsymbol{w}_h}, \boldsymbol{v}_h) - (\nabla \cdot \boldsymbol{v}_h, \overline{r_h}) &= (\boldsymbol{w}_h, \boldsymbol{v}_h) \quad &\forall \, \boldsymbol{v}_h \in \boldsymbol{V}_{h,r}, \\
(\nabla \cdot \overline{\boldsymbol{w}_h}, q_h) &= 0 \quad &\forall \, q_h \in Q_{h,l},
\end{aligned}
$$

where δ is again the filter width, which is a multiple (larger than or equal to 1) of the local mesh width. It is stated in [103] that this step introduces often too much numerical viscosity. For this reason, an approximate discrete deconvolution operator is applied, leading to $D_h \overline{\boldsymbol{w}_h}$. A standard choice for approximate deconvolution operators is the discrete van Cittert operator of order N, given by

$$
D_h \overline{\boldsymbol{w}_h} = \sum_{n=0}^{N} (\mathrm{Id} - G_h)^n \overline{\boldsymbol{w}_h}.
$$

In practice, usually $N \in \{0, 1, 2\}$ is chosen. Finally, a relaxation step is applied to compute the velocity field for the current time instant

$$
\boldsymbol{u}_h = (1 - \chi)\boldsymbol{w}_h + \chi D_h \overline{\boldsymbol{w}_h}, \quad \chi \in [0, 1].
$$

The numerical analysis in [103] studies inf–sup stable pairs of finite element spaces. In the error analysis, instead of the Stokes filter, the simpler differential filter

$$
\delta^2(\nabla \overline{\boldsymbol{w}_h}, \nabla \boldsymbol{v}_h) + (\overline{\boldsymbol{w}_h}, \boldsymbol{v}_h) = (\boldsymbol{w}_h, \boldsymbol{v}_h) \quad \forall \, \boldsymbol{v}_h \in \boldsymbol{V}_{h,r}, \tag{103}
$$

is considered, to reduce technical details. For the velocity error in $L^\infty(L^2)$, a non-robust error bound of order r is derived. Variants of the method from [103], e.g., with a nonlinear filter or without approximate deconvolution are analyzed in [104,105]. In these papers, the same principal result with respect to the velocity error in $L^\infty(L^2)$ is proved as in [103].

4.4.2. Variational multiscale methods

Variational Multiscale (VMS) methods for turbulent flow simulations are based on the variational form (5) of the Navier–Stokes equations. The scale separation in resolved and unresolved scales is defined by projections into

appropriate function spaces, which is an essential difference to LES. The basic principles of VMS methods, which can be applied also to others than flow problems, are developed in [33,106]. There are quite different realizations of these methods for incompressible flow problems, compare the recent reviews [82,107,108]. In particular, there are VMS methods where the terms that model the impact of the unresolved scales on the resolved scales are derived solely by mathematical arguments, like the residual-based VMS method from [14]. Other VMS methods employ a second scale separation of the resolved scales and they utilize eddy viscosity models, like the Smagorinsky model. It is well known that also LPS methods fit into the framework of VMS methods [109]. Because of the mainly mathematically based derivation of VMS methods, there is the expectation that they are accessible to numerical analysis.

Several papers analyze variants of the so-called three-scale projection-based VMS method from [81], which is based on ideas from [110]. This method is designed for inf–sup stable pairs of finite element spaces. The finite element momentum equation includes the additional term

$$(\nu_{\mathrm{T}} \left(\mathbb{D}\boldsymbol{u}_h - \mathbb{G}_H \right), \mathbb{D}\boldsymbol{v}_h), \tag{104}$$

with the turbulent viscosity $\nu_{\mathrm{T}} \geq 0$ and the symmetric tensor-valued function $\mathbb{G}_H$ is defined by the $L^2(\Omega)$ projection

$$(\mathbb{D}\boldsymbol{u}_h - \mathbb{G}_H, \mathbb{L}_H) = 0 \quad \forall \, \mathbb{L}_H \in L_H. \tag{105}$$

The space L_H contains symmetric tensor-valued functions, e.g., piecewise constant tensors on the given mesh. The motivation behind this method consists in restricting the direct impact of the turbulent viscosity to the so-called large resolved scales $\mathbb{D}\boldsymbol{u}_h - \mathbb{G}_H$. A numerical example for a smooth solution in three dimensions and a small viscosity coefficient is presented in [81] that shows that the projection-based VMS method is more accurate on coarse grids, in particular with respect to the velocity error in $L^\infty(L^2)$, than the Smagorinsky model. In [111], the case that ν_{T} is a positive constant is analyzed. Then, (104) and (105) can be reformulated to give the additional term of the form of a LPS stabilization, compare (26),

$$\left(\nu_{\mathrm{T}}(\mathrm{Id} - \Pi_{L_H})\mathbb{D}\boldsymbol{u}_h, (\mathrm{Id} - \Pi_{L_H})\mathbb{D}\boldsymbol{v}_h \right),$$

in the momentum equation, where Π_{L_H} is the $L^2(\Omega)$ projection on L_H. Thus, the additional term has the same principal form as the LPS term of the method from [70], which is discussed in some detail in Section 4.2.2. However, the projection operators are different. Two estimates for velocity errors are presented in [111]. An additional viscosity is defined, which is called multiscale viscosity in [112], and error bounds containing inverse powers of the sum of ν and the multiscale viscosity are derived. However, the bounds are not robust. First, there are still terms with inverse powers of ν and second, it cannot be excluded that the multiscale viscosity vanishes. The analysis is extended to a projection-based VMS method with piecewise constant turbulent viscosity and grad–div term in [112]. Convergence order r for the velocity error in $L^\infty(L^2)$ is proved. Thanks to the grad–div term, inverse powers of ν do not appear in the error bound. However, there are inverse powers of the sum of ν and the multiscale viscosity and again, one cannot exclude that in the worst case the multiscale viscosity vanishes. Altogether, the bound derived in [112] can be considered to be almost robust. In [111,112], the continuous-in-time case is analyzed. Fully discrete schemes, with piecewise constant ν_{T} and without grad–div stabilization are studied in [113,114]. In [113], an uncoupled version of the projection-based VMS method is investigated, where the projection is computed in a separate step, and in [114], an explicit treatment of the projection term is considered. Both papers do not take into account the multiscale viscosity and non-robust bounds of order r for the velocity error in $L^\infty(L^2)$ are derived. A numerical study from [113] for a smooth solution, $\nu = 2 \cdot 10^{-3}$, and the Taylor–Hood pair P_2/P_1 shows that the order of reduction of the velocity error in $L^\infty(L^2)$ on coarse grids is smaller than optimal.

As mentioned before, the so-called term-by-term stabilization method is a special variant of a LPS method. A finite element error analysis of a method including term-by-term LPS stabilization is provided in [115]. It is for a fully discrete scheme with the implicit Euler time stepping scheme, the IMEX discretization of the convective term, and equal order pairs of finite element spaces. Term-by-term stabilizations for the convective term, the grad–div term, and the pressure gradient are utilized. The velocity error in $L^\infty(L^2)$ is proved to be of order r. However, the error bound is not robust since the constant scales like $\nu^{-1/2}$.

4.4.3. Other methods

The grad–div stabilization introduced in Section 4.2.2 results in a mutual coupling between all components of the velocity. There are much more non-zero matrix entries than for other terms, e.g., the standard gradient form of the viscous term or the convective, skew-symmetric, and divergence formulations of the convective term. In [116], a sparse grad–div stabilization is proposed and analyzed that neglects several couplings of the standard grad–div stabilization. The natural expectation is that the sparse approach possesses weaker stability properties than the standard method. In [116], only the splitting error for a projection scheme is analyzed and the corresponding error bound is not robust. Another method to facilitate the application of the grad–div stabilization in practice is proposed in [117]. This modular method computes in the first step an intermediate solution $(\hat{\boldsymbol{u}}_h, p_h)$ by solving the Galerkin discretization of the Navier–Stokes equations. In the second step, a kind of regularization of the intermediate velocity field is applied, which is based on the grad–div operator. Several variants of this step are investigated in [117], where the simplest one is the solution of the variational problem

$$\mu \Delta t (\nabla \cdot \boldsymbol{u}_h, \nabla \cdot \boldsymbol{v}_h) + (\boldsymbol{u}_h, \boldsymbol{v}_h) = (\hat{\boldsymbol{u}}_h, \boldsymbol{v}_h) \quad \forall \, \boldsymbol{v}_h \in V_{h,r},$$

with an appropriate stabilization parameter μ. For a fully discretized problem, a non-robust estimate for the velocity error in $L^\infty(L^2)$ of order r is derived for different variants of the modular grad–div stabilization. An extension of this idea to the BDF2 temporal discretization is studied in [118], where a non-robust velocity estimate in $L^\infty(L^2)$ of order r is proved.

Inserting the differential filter (103) in the momentum equation gives the additional term $(\boldsymbol{w}_h - \overline{\boldsymbol{w}_h}, \boldsymbol{v}_h)/(\Delta t)$. This term has in some sense a similar structure like (104): it is the difference of some term containing the finite element solution and a smoothed term with this solution, where smooth has to be understood in this context in the sense of less oscillations or variations of values. Approaches of this kind has been explored also with respect to the temporal evolution of the solution, leading to so-called temporal regularizations, which add, e.g., terms of the form

$$\gamma (\boldsymbol{u} - \overline{\boldsymbol{u}}) \quad \text{or} \quad -\gamma \Delta (\boldsymbol{u} - \overline{\boldsymbol{u}}) \tag{106}$$

to the momentum equation. If $\boldsymbol{f}$ and $\overline{\boldsymbol{u}}$ are smooth and γ is sufficiently large, then both terms in (106) enforce that $\boldsymbol{u}$ is also smooth. In temporal regularizations, $\overline{\boldsymbol{u}}$ is chosen as the solution from a previous time instant, or a combination of solutions from previous time instants, starting with an already smooth or smoothed initial condition. In [119], a temporal regularization with

$$-\gamma h \Delta (\boldsymbol{u}_{n+1} - \boldsymbol{u}_n)$$

is analyzed for inf–sup stable pairs of finite element spaces. A non-robust bound of the velocity error in $L^\infty(L^2)$ of order r is proved. A so-called stabilization of the curvature is considered in [120], which adds a temporal regularization of the form

$$\frac{\gamma}{\nu} (-\nu \Delta U_{n+1} + (\boldsymbol{u}_{n+1} \cdot \nabla) U_{n+1} + \nabla P_{n+1})$$

with $U_{n+1} = \boldsymbol{u}_{n+1} - 2\boldsymbol{u}_n + \boldsymbol{u}_{n-1}$ and $P_{n+1} = p_{n+1} - 2p_n + p_{n-1}$. It is shown in [120], for inf–sup stable pairs of finite element spaces, that the velocity error in $L^\infty(L^2)$ is bounded of order r in a non-robust way.

5. Summary

In the convergence theory of finite element methods for evolutionary convection–diffusion equations and incompressible Navier–Stokes equations, one can find two types of estimates: robust ones, where the constant in the error bound does not depend on inverse powers of the diffusion or viscosity, and non-robust ones. The classical convergence theory for the Galerkin discretization of the Navier–Stokes equations with inf–sup stable pairs of finite element spaces predicts an optimal order of convergence for the velocity error in $L^\infty(L^2)$, and with that for the kinetic energy, in a non-robust way. However, numerical simulations with non weakly divergence-free pairs of finite element spaces reveal that in the situation of small viscosity coefficients and coarse grids, the solution obtained with the Galerkin method is polluted with spurious oscillations, even if the analytic solution is smooth, compare Example 4.4. This is not the case for methods for which robust estimates can be derived. For such methods, one obtains meaningful numerical solutions in this situation and even a certain order of error reduction, which is the

Table 5

Discretizations of the incompressible Navier–Stokes equations for which robust estimates for the velocity error in $L^\infty(L^2)$ are proved.

Method	Finite element	Order	Analysis	Section
Galerkin	inf–sup stable, EMAC form of the nonlinear term	$r - 1$	[53] + Theorem 4.1	Section 4.1.3
Galerkin	inf–sup stable, weakly divergence-free	r	[57]	Section 4.2.1
Galerkin	H(div)-conforming, inf–sup stable, weakly divergence-free	r	[62,63]	Section 4.2.1
grad–div stabilization	inf–sup stable	r	[50]	Section 4.2.2
LPS for velocity gradient	inf–sup stable	r	[70]	Section 4.2.2
grad–div stabilization, LPS for pressure	non inf–sup stable	r	[72]	Section 4.2.3
LPS for velocity gradient, LPS for pressure	non inf–sup stable	r	[72]	Section 4.2.3
symmetric pressure stabilization	non inf–sup stable	r	[83,84]	Section 4.2.3
SUPG	P_1/P_1, space–time finite elements	1.5	[13]	Section 4.3
CIP	non inf–sup stable	$r + 1/2$	[90]	Section 4.3
LPS for velocity gradient, LPS for pressure, modified parameter (two types of methods)	non inf–sup stable	$r + 1/2$	[72,96]	Section 4.3
H(div)-conforming DG with upwind	inf–sup stable	$r + 1/2$	[64]	Section 4.3

notion used in this survey for the behavior of a method on coarse grids, can be observed. Thus, in the case of small viscosity coefficients and coarse grids, only robust estimates provide useful information about the behavior of a numerical method on coarse grids if the analytic solution is smooth. In this respect, robust error bounds can be considered as the necessary condition (although perhaps not sufficient condition) for a good practical performance in more realistic problems, where the solutions are unlikely to have bounded derivatives as $\nu \to 0$.

For scalar linear convection–diffusion equations, robust estimates of order $r+1/2$ can be proved for many popular stabilized methods. Sometimes, there are restrictions in the analysis, e.g., that the coefficients of the equation do not depend on time. There are special situations, like in Example 3.4 for the SUPG method and P_1 finite elements, where an optimal order of convergence can be observed in numerical simulations.

The situation is more complex for the incompressible Navier–Stokes equations due to the nonlinear convective term and the interaction of velocity and pressure. Robust estimates could be proved so far only for methods whose structure is relatively simple compared with methods that are actually used for turbulent flow simulations. These more complicated methods might contain stabilization mechanisms already used in the simple methods, like the grad–div and SUPG stabilizations in the residual-based VMS method. An overview of discretizations for which robust estimates for the velocity error in $L^\infty(L^2)$ are known is provided in Table 5.

For the Galerkin method and inf–sup stable pairs of finite element spaces that are not weakly divergence-free, only the EMAC form of the nonlinear term allows a robust estimate whose order is $r - 1$. The analysis carried out in the present paper reveals how the different terms in the Navier–Stokes equations affect this order reduction. In particular, the interaction between velocity and pressure is responsible for the order reduction to $r - 1$, unless stabilization terms are added, while the nonlinear term is responsible for the lack of convergence, unless the EMAC form is used or stabilization terms are added.

For other methods using inf–sup stable pairs of finite element spaces, robust estimates of order r for the velocity error in $L^\infty(L^2)$ are proved. Note that for many pairs of such finite element spaces, the velocity space contains piecewise polynomials of degree r and the pressure space of degree $r - 1$. From the point of view of physical consistency (mass conservation, fundamental invariance property, pressure robustness), the use of weakly divergence-free pairs is attractive. However, their implementation might be somewhat involved. The question about the possibility of yielding $r + 1/2$ convergence for these methods adding some kind of stabilization remains open. For non weakly divergence-free pairs, the use of the grad–div stabilization is an appealing approach, because grad–div stabilization is easily to implement. Implementing the LPS method for the velocity gradient is certainly more

involved. Altogether, the best order of convergence of the velocity error in $L^\infty(L^2)$ that could be proved so far with a robust estimate for a method with inf–sup stable pairs of finite element spaces and using continuous velocity elements is r. This is one order less than the best approximation error for the velocity with respect to $L^2(\Omega)^d$. Consequently, the convergence order of the finite element error with respect to the kinetic energy is reduced by one compared with the optimal order.

Concerning inf–sup stable pairs of finite element spaces with $H(\mathrm{div})$-conforming discontinuous velocity finite elements, a robust estimate of order $r + 1/2$ could be proved for an upwind discretization of the convective term.

There is a larger variety of methods where robust error bounds have been proved for non inf–sup stable pairs of finite element spaces. Such pairs are often so-called equal order pairs, where both the finite element velocity and pressure consist of piecewise polynomials of degree r. The stabilization mechanisms that lead for inf–sup stable pairs to methods of order r give also methods of at least order r for equal order pairs. The parameter choice in the LPS method for the velocity gradient can be modified such that even order $r + 1/2$ is reached, which can be proven even for two different types of LPS approaches. There are two more methods with this order, the CIP method and the SUPG method, where for the latter only analysis for $r = 1$ is available. Thus, with equal order pairs of finite element spaces it is possible to achieve order $r + 1/2$ for the convergence of the velocity error in $L^\infty(L^2)$ with robust error bounds, which is half an order less than the optimal order, and so it is the case of the kinetic energy.

By far most of the finite element error analysis that can be found in the literature is for H^1-conforming methods. Nevertheless, the analysis is not complete, e.g., for the SUPG method only a special case is covered and there are many methods, discussed in Section 4.4, where robust error bounds are not yet available. A generalization of $H(\mathrm{div})$-conforming discontinuous methods are so-called hybrid or hybridizable discontinuous Galerkin (HDG) methods. Different formulations for such methods have been derived. However, a finite element analysis for the time-dependent Navier–Stokes equations does not seem to be available, although these methods are used for the simulation of turbulent flows, e.g., in [121]. Numerical analysis of HDG methods for the Stokes equations can be found, e.g., in [61,122].

The state of the art is that there is no method for which robust estimates for the velocity error in $L^\infty(L^2)$, and with that for the error in the kinetic energy, of optimal order can be proved so far. This survey presented numerical results for several methods from Table 5 and all of them showed that the analytic prediction of the order of error reduction is sharp. An optimal order could not be observed for any of these methods. We performed with these methods also simulations on a family of triangulations of Friedrichs–Keller type, like for the SUPG method for convection–diffusion equations in Fig. 4. For brevity, the results are not presented here. Even on these regular grids, none of the methods showed optimal error reduction of the velocity error in $L^\infty(L^2)$. Thus, one can conclude that among the methods for which numerical studies are presented in this survey, there is no method with optimal error reduction, even in special situations like regular grids. In summary, whether or not there is a method with an optimal robust error bound for the velocity error in $L^\infty(L^2)$, perhaps at least in special situations, remains an open problem.

Another widely open problem is the question of local error estimates, that is, to obtain (robust) error bounds in subsets of the domain away from layers. For time-dependent problems, only in [4,5] such estimates can be found, for the SUPG method applied to linear evolutionary convection–reaction–diffusion equations. Further open problems, like the analysis of the considered equations on anisotropic grids or numerical analysis for temporal or spatio-temporal averages of errors are already stated in [17].

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

References

[1] Hermann Sohr, The navier-stokes equations, in: Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks], Birkhäuser Verlag, Basel, 2001, p. x+367, An elementary functional analytic approach.

[2] Volker John, Finite Element Methods for Incompressible Flow Problems, in: Springer Series in Computational Mathematics, vol. 51, Springer, Cham, 2016, p. xiii+812.

[3] Eberhard Hopf, Über die Anfangswertaufgabe für die hydrodynamischen grundgleichungen, Math. Nachr. 4 (1951) 213–231.

[4] U. Nävert, A Finite Element Method for Convection-Diffusion Problems (Ph.D. thesis), Chalmers University of Technology, Göteborg, 1982.

[5] Javier de Frutos, Bosco García-Archilla, Julia Novo, Local error estimates for the SUPG method applied to evolutionary convection-reaction-diffusion equations, J. Sci. Comput. 66 (2) (2016) 528–554.

[6] Clemens Bartsch, Viktoria Wiedmeyer, Zahra Lakdawala, Robert I.A. Patterson, Andreas Voigt, Kai Sundmacher, Volker John, Stochastic-deterministic population balance modeling and simulation of a fluidized bed crystallizer experiment, Chem. Eng. Sci. 208 (2019) 115102.

[7] T.J.R. Hughes, A. Brooks, A multidimensional upwind scheme with no crosswind diffusion, in: Finite Element Methods for Convection Dominated Flows (Papers, Winter Ann. Meeting Amer. Soc. Mech. Engrs., New York, 1979), in: AMD, vol. 34, Amer. Soc. Mech. Engrs. (ASME), New York, 1979, pp. 19–35.

[8] Alexander N. Brooks, Thomas J.R. Hughes, Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible navier-Stokes equations, in: FENOMECH '81, Part I (Stuttgart, 1981), Comput. Methods Appl. Mech. Engrg. 32 (1–3) (1982) 199–259.

[9] Claes Johnson, Uno Nävert, An analysis of some finite element methods for advection-diffusion problems, in: Analytical and Numerical Approaches to Asymptotic Problems in Analysis (Proc. Conf., Univ. Nijmegen, Nijmegen, 1980), in: North-Holland Math. Stud, vol. 47, North-Holland, Amsterdam-New York, 1981, pp. 99–116.

[10] G. Lube, L. Tobiska, A nonconforming finite element method of streamline diffusion type for the incompressible Navier-Stokes equations, J. Comput. Math. 8 (2) (1990) 147–158.

[11] Lutz Tobiska, Gert Lube, A modified streamline diffusion method for solving the stationary Navier-Stokes equation, Numer. Math. 59 (1) (1991) 13–29.

[12] Claes Johnson, Jukka Saranen, Streamline diffusion methods for the incompressible euler and Navier-Stokes equations, Math. Comp. 47 (175) (1986) 1–18.

[13] Peter Hansbo, Anders Szepessy, A velocity-pressure streamline diffusion finite element method for the incompressible Navier-Stokes equations, Comput. Methods Appl. Mech. Engrg. 84 (2) (1990) 175–192.

[14] Y. Bazilevs, V.M. Calo, J.A. Cottrell, T.J.R. Hughes, A. Reali, G. Scovazzi, Variational multiscale residual-based turbulence modeling for large eddy simulation of incompressible flows, Comput. Methods Appl. Mech. Engrg. 197 (1–4) (2007) 173–201.

[15] Hans-Görg Roos, Martin Stynes, Lutz Tobiska, Robust Numerical Methods for Singularly Perturbed Differential Equations, second ed., in: Springer Series in Computational Mathematics, vol. 24, Springer-Verlag, Berlin, 2008, p. xiv+604, Convection-diffusion-reaction and flow problems.

[16] Stephen B. Pope, Turbulent Flows, Cambridge University Press, Cambridge, 2000, p. xxxiv+771.

[17] Volker John, Petr Knobloch, Julia Novo, Finite elements for scalar convection-dominated equations and incompressible flow problems: A never ending story? Comput. Vis. Sci. 19 (5–6) (2018) 47–63.

[18] Philippe G. Ciarlet, The Finite Element Method for Elliptic Problems, in: Studies in Mathematics and its Applications, vol. 4, North-Holland Publishing Co., Amsterdam-New York-Oxford, 1978, p. xix+530.

[19] Daniel Arndt, Helene Dallmann, Gert Lube, Local projection FEM stabilization for the time-dependent incompressible Navier-Stokes problem, Numer. Methods Partial Differential Equations 31 (4) (2015) 1224–1250.

[20] Javier de Frutos, Bosco García-Archilla, Julia Novo, Stabilization of Galerkin finite element approximations to transient convection-diffusion problems, SIAM J. Numer. Anal. 48 (3) (2010) 953–979.

[21] Timothy A. Davis, Algorithm 832: UMFPACK V4.3—An unsymmetric-pattern multifrontal method, ACM Trans. Math. Softw. 30 (2) (2004) 196–199.

[22] Volker John, Ellen Schmeyer, Finite element methods for time-dependent convection-diffusion-reaction equations with small diffusion, Comput. Methods Appl. Mech. Engrg. 198 (3–4) (2008) 475–494.

[23] Volker John, Julia Novo, On (essentially) non-oscillatory discretizations of evolutionary convection-diffusion equations, J. Comput. Phys. 231 (4) (2012) 1570–1586.

[24] Volker John, Julia Novo, Error analysis of the SUPG finite element discretization of evolutionary convection-diffusion-reaction equations, SIAM J. Numer. Anal. 49 (3) (2011) 1149–1176.

[25] Erik Burman, Giles Smith, Analysis of the space semi-discretized SUPG method for transient convection-diffusion equations, Math. Models Methods Appl. Sci. 21 (10) (2011) 2049–2068.

[26] G. Lube, D. Weiss, Stabilized finite element methods for singularly perturbed parabolic problems, Appl. Numer. Math. 17 (4) (1995) 431–459.

[27] Ramon Codina, Jordi Blasco, Analysis of a stabilized finite element approximation of the transient convection-diffusion-reaction equation using orthogonal subscales, Comput. Vis. Sci. 4 (3) (2002) 167–174.

[28] N. Ahmed, G. Matthies, L. Tobiska, H. Xie, Discontinuous Galerkin time stepping with local projection stabilization for transient convection-diffusion-reaction problems, Comput. Methods Appl. Mech. Engrg. 200 (21–22) (2011) 1747–1756.

[29] Naveed Ahmed, Gunar Matthies, Higher order continuous Galerkin-Petrov time stepping schemes for transient convection-diffusion-reaction equations, ESAIM Math. Model. Numer. Anal. 49 (5) (2015) 1429–1450.

[30] Gabriel R. Barrenechea, Volker John, Petr Knobloch, A local projection stabilization finite element method with nonlinear crosswind diffusion for convection-diffusion-reaction equations, ESAIM Math. Model. Numer. Anal. 47 (5) (2013) 1335–1366.

[31] Erik Burman, Miguel A. Fernández, Finite element methods with symmetric stabilization for the transient convection-diffusion-reaction equation, Comput. Methods Appl. Mech. Engrg. 198 (33–36) (2009) 2508–2519.

[32] Jean-Luc Guermond, Subgrid stabilization of Galerkin approximations of linear contraction semi-groups of class C^0 in Hilbert spaces, Numer. Methods Partial Differential Equations 17 (1) (2001) 1–25.

[33] Thomas J.R. Hughes, Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods, Comput. Methods Appl. Mech. Engrg. 127 (1–4) (1995) 387–401.

[34] M. Feistauer, K. Švadlenka, Discontinuous Galerkin method of lines for solving nonstationary singularly perturbed linear problems, J. Numer. Math. 12 (2) (2004) 97–117.

[35] Vít Dolejší, Miloslav Feistauer, Discontinuous Galerkin Method, in: Springer Series in Computational Mathematics, vol. 48, Springer, Cham, 2015, p. xiv+572, Analysis and applications to compressible flow.

[36] J. Gopalakrishnan, G. Kanschat, A multilevel discontinuous Galerkin method, Numer. Math. 95 (3) (2003) 527–550.

[37] Erik Burman, Consistent SUPG-method for transient transport problems: Stability and convergence, Comput. Methods Appl. Mech. Engrg. 199 (17–20) (2010) 1114–1123.

[38] Santiago Badia, Alba Hierro, On monotonicity-preserving stabilized finite element approximations of transport problems, SIAM J. Sci. Comput. 36 (6) (2014) A2673–A2697.

[39] Hans-Görg Roos, Martin Stynes, Some open questions in the numerical analysis of singularly perturbed differential equations, Comput. Methods Appl. Math. 15 (4) (2015) 531–550.

[40] Guohui Zhou, How accurate is the streamline diffusion finite element method? Math. Comp. 66 (217) (1997) 31–44.

[41] John G. Heywood, Rolf 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 (2) (1982) 275–311.

[42] John G. Heywood, Rolf Rannacher, Finite element approximation of the nonstationary Navier-Stokes problem. II. Stability of solutions and error estimates uniform in time, SIAM J. Numer. Anal. 23 (4) (1986) 750–777.

[43] John G. Heywood, Rolf 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 (3) (1988) 489–512.

[44] John G. Heywood, Rolf Rannacher, Finite-element approximation of the nonstationary Navier-Stokes problem. IV. Error analysis for second-order time discretization, SIAM J. Numer. Anal. 27 (2) (1990) 353–384.

[45] D.N. Arnold, F. Brezzi, M. Fortin, A stable finite element for the Stokes equations, Calcolo 21 (4) (1984) 337–344 (1985).

[46] P. Hood, C. Taylor, Navier–Stokes equations using mixed interpolation, in: J.T. Oden, R.H. Gallagher, O.C. Zienkiewicz, C. Taylor (Eds.), Finite Element Methods in Flow Problems, University of Alabama in Huntsville Press, 1974, pp. 121–132.

[47] Volker John, Alexander Linke, Christian Merdon, Michael Neilan, Leo G. Rebholz, On the divergence constraint in mixed finite element methods for incompressible flows, SIAM Rev. 59 (3) (2017) 492–544.

[48] Javier de Frutos, Bosco García-Archilla, Volker John, Julia Novo, Grad-div stabilization for the evolutionary Oseen problem with inf-sup stable finite elements, J. Sci. Comput. 66 (3) (2016) 991–1024.

[49] Bosco García-Archilla, Julia Novo, Edriss S. Titi, Uniform in time error estimates for a finite element method applied to a downscaling data assimilation algorithm for the Navier-Stokes equations, SIAM J. Numer. Anal. 58 (1) (2020) 410–429.

[50] Javier de Frutos, Bosco García-Archilla, Volker John, Julia Novo, Analysis of the grad-div stabilization for the time-dependent Navier-Stokes equations with inf-sup stable finite elements, Adv. Comput. Math. 44 (1) (2018) 195–225.

[51] Hongsen Chen, Pointwise error estimates for finite element solutions of the Stokes problem, SIAM J. Numer. Anal. 44 (1) (2006) 1–28.

[52] Sergey Charnyi, Timo Heister, Maxim A. Olshanskii, Leo G. Rebholz, On conservation laws of Navier-Stokes Galerkin discretizations, J. Comput. Phys. 337 (2017) 289–308.

[53] Maxim A. Olshanskii, Leo G. Rebholz, Longer time accuracy for incompressible Navier-Stokes simulations with the EMAC formulation, Comput. Methods Appl. Mech. Engrg. 372 (2020) 113369, 17.

[54] L.R. Scott, M. Vogelius, Conforming finite element methods for incompressible and nearly incompressible continua, in: Large-Scale Computations in Fluid Mechanics, Part 2 (La Jolla, Calif., 1983), in: Lectures in Appl. Math, vol. 22, Amer. Math. Soc., Providence, RI, 1985, pp. 221–244.

[55] Jinshui Qin, On the Convergence of Some Low Order Mixed Finite Elements for Incompressible Fluids (Ph.D. thesis), Department of Mathematics, Pennsylvania State University, 1994.

[56] Shangyou Zhang, A new family of stable mixed finite elements for the 3D Stokes equations, Math. Comp. 74 (250) (2005) 543–554.

[57] Philipp W. Schroeder, Gert Lube, Pressure-robust analysis of divergence-free and conforming FEM for evolutionary incompressible Navier-Stokes flows, J. Numer. Math. 25 (4) (2017) 249–276.

[58] Benjamin R. Cousins, Leo G. Rebholz, Nicholas E. Wilson, Enforcing energy, helicity and strong mass conservation in finite element computations for incompressible Navier-Stokes simulations, Appl. Math. Comput. 218 (4) (2011) 1208–1221.

[59] Philipp W. Schroeder, Volker John, Philip L. Lederer, Christoph Lehrenfeld, Gert Lube, Joachim Schöberl, On reference solutions and the sensitivity of the 2D Kelvin-Helmholtz instability problem, Comput. Math. Appl. 77 (4) (2019) 1010–1028.

[60] Bernardo Cockburn, Guido Kanschat, Dominik Schötzau, A note on discontinuous Galerkin divergence-free solutions of the Navier-Stokes equations, J. Sci. Comput. 31 (1–2) (2007) 61–73.

[61] Philip L. Lederer, Christoph Lehrenfeld, Joachim Schöberl, Hybrid discontinuous Galerkin methods with relaxed $H(\mathrm{div})$-conformity for incompressible flows. Part II, ESAIM Math. Model. Numer. Anal. 53 (2) (2019) 503–522.

[62] Philipp W. Schroeder, Gert Lube, Divergence-free $H(\mathrm{div})$-FEM for time-dependent incompressible flows with applications to high Reynolds number vortex dynamics, J. Sci. Comput. 75 (2) (2018) 830–858.

[63] Philipp W. Schroeder, Christoph Lehrenfeld, Alexander Linke, Gert Lube, Towards computable flows and robust estimates for inf-sup stable FEM applied to the time-dependent incompressible Navier-Stokes equations, SeMA J. 75 (4) (2018) 629–653.

[64] Yongbin Han, Yanren Hou, Robust error analysis of H(div)-conforming DG method for the time-dependent incompressible Navier-Stokes equations, J. Comput. Appl. Math. 390 (2021) 113365, 13.

[65] Eleanor W. Jenkins, Volker John, Alexander Linke, Leo G. Rebholz, On the parameter choice in grad-div stabilization for the Stokes equations, Adv. Comput. Math. 40 (2) (2014) 491–516.

[66] Leopoldo P. Franca, Thomas J.R. Hughes, Two classes of mixed finite element methods, Comput. Methods Appl. Mech. Engrg. 69 (1) (1988) 89–129.

[67] Maxim Olshanskii, Gert Lube, Timo Heister, Johannes Löwe, Grad-div stabilization and subgrid pressure models for the incompressible Navier-Stokes equations, Comput. Methods Appl. Mech. Engrg. 198 (49–52) (2009) 3975–3988.

[68] Blanca Ayuso, Bosco García-Archilla, Julia Novo, The postprocessed mixed finite-element method for the Navier-Stokes equations, SIAM J. Numer. Anal. 43 (3) (2005) 1091–1111.

[69] Michael A. Case, Vincent J. Ervin, Alexander Linke, Leo G. Rebholz, A connection between Scott-Vogelius and grad-div stabilized Taylor-Hood FE approximations of the Navier-Stokes equations, SIAM J. Numer. Anal. 49 (4) (2011) 1461–1481.

[70] Naveed Ahmed, Gunar Matthies, Higher-order discontinuous Galerkin time discretizations for the evolutionary Navier–Stokes equations, IMA J. Numer. Anal. (2021) in press.

[71] Volker John, Petr Knobloch, Ulrich Wilbrandt, Finite element pressure stabilizations for incompressible flow problems, in: T. Bodnár, G.P. Galdi, Š. Nečasová (Eds.), Fluids Under Pressure, in: Adv. Math. Fluid Mech, Birkhäuser, 2020, pp. 483–573.

[72] Javier de Frutos, Bosco García-Archilla, Volker John, Julia Novo, Error analysis of non inf-sup stable discretizations of the time-dependent Navier-Stokes equations with local projection stabilization, IMA J. Numer. Anal. 39 (4) (2019) 1747–1786.

[73] T. Chacón Rebollo, M. Gómez Mármol, V. Girault, I. Sánchez Muñoz, A high order term-by-term stabilization solver for incompressible flow problems, IMA J. Numer. Anal. 33 (3) (2013) 974–1007.

[74] T. Chacón Rebollo, M. Gómez Mármol, M. Restelli, Numerical analysis of penalty stabilized finite element discretizations of evolution Navier-Stokes equations, J. Sci. Comput. 63 (3) (2015) 885–912.

[75] C. Bernardi, V. Girault, A local regularization operator for triangular and quadrilateral finite elements, SIAM J. Numer. Anal. 35 (5) (1998) 1893–1916.

[76] L. Ridgway Scott, Shangyou Zhang, Finite element interpolation of nonsmooth functions satisfying boundary conditions, Math. Comp. 54 (190) (1990) 483–493.

[77] Susanne C. Brenner, L. Ridgway Scott, The Mathematical Theory of Finite Element Methods, third ed., Texts in Applied Mathematics, vol. 15, Springer, New York, 2008, p. xviii+397.

[78] Silvia Bertoluzza, The discrete commutator property of approximation spaces, C. R. Acad. Sci., Paris 329 (12) (1999) 1097–1102.

[79] Volker Gravemeier, Michael W. Gee, Wolfgang A. Wall, An algebraic variational multiscale-multigrid method based on plain aggregation for convection-diffusion problems, Comput. Methods Appl. Mech. Engrg. 198 (47–48) (2009) 3821–3835.

[80] Volker Gravemeier, Michael W. Gee, Martin Kronbichler, Wolfgang A. Wall, An algebraic variational multiscale-multigrid method for large eddy simulation of turbulent flow, Comput. Methods Appl. Mech. Engrg. 199 (13–16) (2010) 853–864.

[81] Volker John, Songul Kaya, A finite element variational multiscale method for the Navier-Stokes equations, SIAM J. Sci. Comput. 26 (5) (2005) 1485–1503 (electronic).

[82] Naveed Ahmed, Tomás Chacón Rebollo, Volker John, Samuele Rubino, A review of variational multiscale methods for the simulation of turbulent incompressible flows, Arch. Comput. Methods Eng. 24 (1) (2017) 115–164.

[83] Bosco García-Archilla, Volker John, Julia Novo, Symmetric pressure stabilization for equal-order finite element approximations to the time-dependent Navier–Stokes equations, IMA J. Numer. Anal. 41 (2) (2021) 1093–1129.

[84] Gang Chen, Minfu Feng, Analysis of solving Galerkin finite element methods with symmetric pressure stabilization for the unsteady Navier-Stokes equations using conforming equal order interpolation, Adv. Appl. Math. Mech. 9 (2) (2017) 362–377.

[85] Erik Burman, Miguel A. Fernández, Galerkin finite element methods with symmetric pressure stabilization for the transient stokes equations: Stability and convergence analysis, SIAM J. Numer. Anal. 47 (1) (2008/09) 409–439.

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

[87] Clark R. Dohrmann, Pavel B. Bochev, A stabilized finite element method for the Stokes problem based on polynomial pressure projections, Internat. J. Numer. Methods Fluids 46 (2) (2004) 183–201.

[88] Ramon Codina, Jordi Blasco, A finite element formulation for the Stokes problem allowing equal velocity-pressure interpolation, Comput. Methods Appl. Mech. Engrg. 143 (3–4) (1997) 373–391.

[89] R. Becker, M. Braack, A finite element pressure gradient stabilization for the Stokes equations based on local projections, Calcolo 38 (4) (2001) 173–199.

[90] Erik Burman, Miguel A. Fernández, Continuous interior penalty finite element method for the time-dependent Navier-stokes equations: Space discretization and convergence, Numer. Math. 107 (1) (2007) 39–77.

[91] Alejandro Allendes, Gabriel Barrenechea, Julia Novo, A divergence-free low-order stabilized finite element method for the evolutionary Navier-Stokes equations, SIAM J. Sci. Comput. (2021) in press.

[92] Gang Chen, Min-fu Feng, Yin-nian He, Finite difference streamline diffusion method using nonconforming space for incompressible time-dependent Navier-Stokes equations, Appl. Math. Mech. (English Ed.) 34 (9) (2013) 1083–1096.

[93] Naveed Ahmed, Samuele Rubino, Numerical comparisons of finite element stabilized methods for a 2D vortex dynamics simulation at high Reynolds number, Comput. Methods Appl. Mech. Engrg. 349 (2019) 191–212.

[94] Naveed Ahmed, Volker John, An assessment of two classes of variational multiscale methods for the simulation of incompressible turbulent flows, Comput. Methods Appl. Mech. Engrg. 365 (2020) 112997.

[95]   Erik Burman, Alexandre Ern, Miguel A. Fernández, Fractional-step methods and finite elements with symmetric stabilization for the transient Oseen problem, ESAIM Math. Model. Numer. Anal. 51 (2) (2017) 487–507.

[96]   Gang Chen, Minfu Feng, Hong Zhou, Local projection stabilized method on unsteady Navier-Stokes equations with high Reynolds number using equal order interpolation, Appl. Math. Comput. 243 (2014) 465–481.

[97]   J. Smagorinsky, General circulation experiments with the primitive equations, Mon. Weather Rev. 91 (1963) 99–164.

[98]   O.A. Ladyženskaja, New equations for the description of the motions of viscous incompressible fluids, and global solvability for their boundary value problems, Tr. Mat. Inst. Steklova 102 (1967) 85–104.

[99]   V. John, W.J. Layton, Analysis of numerical errors in large eddy simulation, SIAM J. Numer. Anal. 40 (3) (2002) 995–1020.

[100]   Tomás Chacón Rebollo, Roger Lewandowski, Mathematical and numerical foundations of turbulence models and applications, in: Modeling and Simulation in Science, Engineering and Technology, Birkhäuser/Springer, New York, 2014, p. xviii+517.

[101]   Erik Burman, Peter Hansbo, Mats G. Larson, Error estimates for the Smagorinsky turbulence model: Enhanced stability through scale separation and numerical stabilization, 2021, arXiv:2102.00043v1.

[102]   Snorre H. Christiansen, Kaibo Hu, Generalized finite element systems for smooth differential forms and Stokes' problem, Numer. Math. 140 (2) (2018) 327–371.

[103]   Vincent J. Ervin, William J. Layton, Monika Neda, Numerical analysis of filter-based stabilization for evolution equations, SIAM J. Numer. Anal. 50 (5) (2012) 2307–2335.

[104]   William Layton, Leo G. Rebholz, Catalin Trenchea, Modular nonlinear filter stabilization of methods for higher Reynolds numbers flow, J. Math. Fluid Mech. 14 (2) (2012) 325–354.

[105]   William Layton, Nathaniel Mays, Monika Neda, Catalin Trenchea, Numerical analysis of modular regularization methods for the BDF2 time discretization of the Navier-Stokes equations, ESAIM Math. Model. Numer. Anal. 48 (3) (2014) 765–793.

[106]   Jean-Luc Guermond, Stabilization of Galerkin approximations of transport equations by subgrid modeling, M2AN Math. Model. Numer. Anal. 33 (6) (1999) 1293–1316.

[107]   Ramon Codina, Santiago Badia, Joan Baiges, Javier Principe, Variational multiscale methods in computational fluid dynamics, in: E. Stein, R. Borst, T.J.R. Hughes (Eds.), Encyclopedia of Computational Mechanics, second ed., American Cancer Society, 2017, pp. 1–28.

[108]   Ursula Rasthofer, Volker Gravemeier, Recent developments in variational multiscale methods for large-eddy simulation of turbulent flow, Arch. Comput. Methods Eng. 25 (3) (2018) 647–690.

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

[110]   W. Layton, A connection between subgrid scale eddy viscosity and mixed methods, Appl. Math. Comput. 133 (1) (2002) 147–157.

[111]   Volker John, Songul Kaya, Finite element error analysis of a variational multiscale method for the Navier-Stokes equations, Adv. Comput. Math. 28 (1) (2008) 43–61.

[112]   Lars Röhe, Gert Lube, Analysis of a variational multiscale method for large-eddy simulation and its application to homogeneous isotropic turbulence, Comput. Methods Appl. Mech. Engrg. 199 (37–40) (2010) 2331–2342.

[113]   William Layton, Lars Röhe, Hoang Tran, Explicitly uncoupled VMS stabilization of fluid flow, Comput. Methods Appl. Mech. Engrg. 200 (45–46) (2011) 3183–3199.

[114]   Li Shan, Yanren Hou, Haibiao Zheng, Variational multiscale method based on the Crank-Nicolson extrapolation scheme for the non-stationary Navier-Stokes equations, Int. J. Comput. Math. 89 (16) (2012) 2198–2223.

[115]   Naveed Ahmed, Tomás Chacón Rebollo, Volker John, Samuele Rubino, Analysis of a full space-time discretization of the navier-Stokes equations by a local projection stabilization method, IMA J. Numer. Anal. 37 (3) (2017) 1437–1467.

[116]   Abigail L. Bowers, Sabine Le Borne, Leo G. Rebholz, Error analysis and iterative solvers for Navier-Stokes projection methods with standard and sparse grad-div stabilization, Comput. Methods Appl. Mech. Engrg. 275 (2014) 1–19.

[117]   J.A. Fiordilino, W. Layton, Y. Rong, An efficient and modular grad-div stabilization, Comput. Methods Appl. Mech. Engrg. 335 (2018) 327–346.

[118]   Y. Rong, J.A. Fiordilino, Numerical analysis of a BDF2 modular grad–div stabilization method for the Navier–Stokes equations, J. Sci. Comput. 82 (3) (2020).

[119]   Alexandr Labovsky, William J. Layton, Carolina C. Manica, Monika Neda, Leo G. Rebholz, The stabilized extrapolated trapezoidal finite-element method for the Navier-Stokes equations, Comput. Methods Appl. Mech. Engrg. 198 (9–12) (2009) 958–974.

[120]   Nan Jiang, Muhammad Mohebujjaman, Leo G. Rebholz, Catalin Trenchea, An optimally accurate discrete regularization for second order timestepping methods for Navier-Stokes equations, Comput. Methods Appl. Mech. Engrg. 310 (2016) 388–405.

[121]   Niklas Fehn, Martin Kronbichler, Christoph Lehrenfeld, Gert Lube, Philipp W. Schroeder, High-order DG solvers for underresolved turbulent incompressible flows: A comparison of L^2 and H(div) methods, Internat. J. Numer. Methods Fluids 91 (11) (2019) 533–556.

[122]   Sander Rhebergen, Garth N. Wells, Analysis of a hybridized/interface stabilized finite element method for the Stokes equations, SIAM J. Numer. Anal. 55 (4) (2017) 1982–2003.