IMA Journal of Numerical Analysis (2016) **00**, 1–31
doi: 10.1093/imanum/drw069

Grad-div stabilized discretizations on S-type meshes for the Oseen problem

SEBASTIAN FRANZ, KATHARINA HÖHNE* AND GUNAR MATTHIES
Institut für Numerische Mathematik, Technische Universität Dresden, 01062 Dresden, Germany
*Corresponding author: katharina.hoehne1@tu-dresden.de

[Received on 6 January 2016; revised on 23 November 2016]

We consider discretizations of the singularly perturbed Oseen equations on properly layer-adapted meshes. Using a suitable solution decomposition, we are able to prove optimal convergence orders in the associated energy norm for grad-div stabilized finite element methods in a general setting. Two families of pairs of discrete function spaces, namely $Q_k \times Q_{k-1}$ and $Q_k \times P_{k-1}^{\mathrm{disc}}$, $k \geq 2$, are investigated in detail. The usage of a standard nonstabilized Galerkin method reduces the order by 1 while stabilization outside the layers is enough to regain the full optimal order.

Keywords: Oseen equations; layer-adapted meshes; singular perturbations; grad-div stabilization.

1. Introduction

The numerical solution of the time-dependent Navier–Stokes equations, especially for high Reynolds numbers, is still a great challenge of computational fluid dynamics. Using time discretization and linearization, the Oseen problem arises as an important subproblem that is for high Reynolds numbers singularly perturbed. It is well known that standard finite element methods applied to the Oseen problem may suffer from two difficulties: the generally dominating convection and the incompatibility between the approximation spaces for velocity and pressure.

The streamline-upwind Petrov–Galerkin method (SUPG), introduced in Brooks & Hughes (1982), and the pressure-stabilized Petrov–Galerkin method (PSPG), introduced in Hughes *et al.* (1986), Johnson & Saranen (1986), Tezduyar *et al.* (1992), provide a unique frame to tackle both problems simultaneously. An additional elementwise stabilization of the divergence constraint, abbreviated as grad-div stabilization, increases the robustness; see Franca & Frey (1992), Hansbo & Szepessy (1990), Tobiska & Lube (1991). In order to ensure the strong consistency of these residual-based stabilization methods, various terms have to be added to the weak formulation. The PSPG term can be skipped if inf–sup stable pairs of finite element spaces for approximating velocity and pressure are used. This results in reduced stabilized schemes; see Gelhard *et al.* (2005), Matthies *et al.* (2009). The coupling between velocity and pressure that is caused by the SUPG term complicates the analysis and the grad-div stabilization seems to be even more important; compare Burman & Linke (2008), Gelhard *et al.* (2005), Linke & Rebholz (2013), Jenkins *et al.* (2014), Matthies *et al.* (2009), Olshanskii & Reusken (2004).

On general meshes, the use of stabilization techniques provides semirobust error estimates in the sense of Roos (2012). This means that the constants in the error bounds are independent of the small viscosity parameter while the higher-order norms of the solution that appear strongly increase with decreasing viscosity.

The role of grad-div stabilization has been investigated by several authors. The combination of SUPG and grad-div stabilization applied to the stationary Navier–Stokes equations was considered in Olshanskii

(2002). The grad-div term enhances the accuracy of the numerical solution and the convergence properties of iterative solvers if the parameter of the grad-div stabilization is chosen properly. The influence of grad-div stabilization for the Stokes equations has been studied in Olshanskii & Reusken (2004), Jenkins *et al.* (2014). Within the scope of variational multiscale methods, the grad-div stabilization as a pressure subgrid model has been investigated in Olshanskii *et al.* (2009). The importance of the grad-div stabilization inside local projection stabilizations was stated in Dallmann *et al.* (2016).

Standard methods for stationary singularly perturbed convection–diffusion problems do also lack stability; see, e.g., the book Roos *et al.* (2008). For these problems, knowledge of the layer structure of their solutions allows creation and usage of properly layer-adapted meshes; see the survey book Linß (2010). Such special meshes have quite a long history. First ideas were proposed by Bakhvalov (1969) using exponentially fitted boundary meshes. The piecewise uniform Shishkin meshes (Miller *et al.*, 1996; Stynes & O'Riordan, 1997) are easier in structure and in the analysis of methods based on them, but their approximation quality is inferior to Bakhvalov meshes. The combination of Bakhvalov's graded fine meshes with Shishkin's choice of the transition point (Linß, 1999, 2000) led to general S-type meshes (Roos & Linß, 1999) where the mesh inside the layers is defined with the help of a mesh-generating function. The use of layer-adapted meshes allows for robust or uniform error estimates in the sense of Roos (2012). Hence, the error bounds are completely independent of the viscosity parameter.

On such adapted meshes, the oscillations induced by boundary conditions are strongly reduced. However, the lack of stability is still there. Therefore, stabilization methods like SUPG (also called SDFEM), local projection stabilization or interior penalty stabilization are used in addition. Some results for those combinations can be found in Roos *et al.* (2008), Franz *et al.* (2008), Franz (2008), Franz & Matthies (2010), Stynes & Tobiska (2008).

This article now combines these two approaches. We apply a grad-div stabilization on layer-adapted meshes. Using suitable pairs of function spaces for approximating velocity and pressure, we prove error bounds that are uniform in ε and yield an optimal convergence rate. Stabilized finite element methods for the Oseen problem on anisotropic meshes have been analysed in Apel *et al.* (2008). The authors considered equal-order interpolation for velocity and pressure, but S-type meshes were explicitly excluded.

We consider the stationary Oseen equations

$$\begin{aligned} -\varepsilon\Delta\boldsymbol{u} - (\boldsymbol{b}\cdot\nabla)\boldsymbol{u} + c\boldsymbol{u} + \nabla p &= \boldsymbol{f} \qquad \text{in } \Omega = (0,1)^2, \\ \operatorname{div}\boldsymbol{u} &= 0 \qquad \text{in } \Omega, \\ \boldsymbol{u} &= 0 \qquad \text{on } \Gamma = \partial\Omega, \end{aligned} \tag{1.1}$$

where the perturbation parameter ε with $0 < \varepsilon \ll 1$ corresponds to the viscosity, the convection field $\boldsymbol{b} = (b_1, b_2) \in W^{1,\infty}(\Omega)^2$ with $\operatorname{div}\boldsymbol{b} = 0$ fulfils $b_1 \geq \beta_1 > 0,\ b_2 \geq \beta_2 > 0$ with some positive constants β_1, β_2 and $0 < c_0 \leq c \in L^\infty(\Omega)$ is assumed. Then $\boldsymbol{u}$ describes the velocity of our problem and p the pressure.

Since $0 < \varepsilon \ll 1$, we expect layers at $x = 0$ and $y = 0$. A typical solution of (1.1) can be seen in Fig. 1.

As the solution exhibits boundary layers, the use of quasi-uniform meshes gives accurate approximations of (1.1) only if the mesh size is of the order of the perturbation parameter ε. Since this gives a prohibitive restriction, layer-adapted meshes will be used for the discretizations.

We are aware that the considered problem class is not a prototype for Oseen problems coming from applications as we utilize a solution decomposition (see Assumption 2.1) which is usually unavailable in

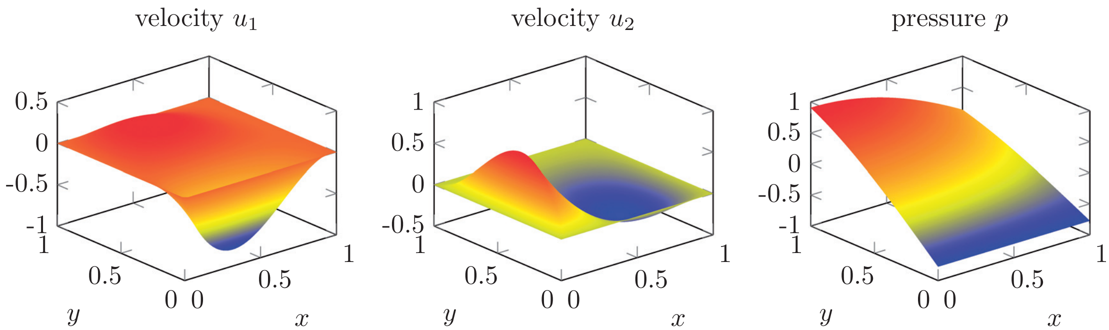

FIG. 1. Typical solution of (1.1) with exponential layers in the velocity.

real-world problems. However, this article is intended as a starting point for the robust numerical analysis of such problems on layer-adapted meshes.

The article is organized as follows. Section 2 introduces a weak formulation, a solution decomposition and layer-adapted meshes. In Section 3, we discretize and analyse our problem in an abstract setting without specifying the finite element spaces. We prove optimal convergence orders for the fully grad-div stabilized method and for stabilization only outside the layers while the order is reduced by 1 for the standard Galerkin method. We present in Section 4, two pairs of spaces that will suit our purpose. Finally, Section 5 presents numerical simulations confirming the theoretical results.

Notation: On a domain D, we denote the L^2-norm by $\|\cdot\|_D$ and the L^∞-norm by $\|\cdot\|_{\infty,D}$. In the case of $D = \Omega$, we drop the reference to the domain from the notation. Moreover, C will denote a generic constant independent of the perturbation parameter ε and the mesh parameter N.

2. Solution decomposition and layer-adapted meshes

Let us start with a weak formulation of (1.1). We define $V := H_0^1(\Omega)^2$ as the space of vector functions with components in the Sobolev space $H_0^1(\Omega)$ of $L^2(\Omega)$-functions with weak derivatives in $L^2(\Omega)$ and vanishing boundary traces, and $Q := L_0^2(\Omega)$ as the space of $L^2(\Omega)$-functions with vanishing integral mean value.

A weak formulation then reads as follows:

Find $(\boldsymbol{u}, p) \in V \times Q$ such that for all $(\boldsymbol{v}, q) \in V \times Q$,

$$\begin{aligned} \varepsilon(\nabla \boldsymbol{u}, \nabla \boldsymbol{v}) - \big((\boldsymbol{b}\cdot\nabla)\boldsymbol{u}, \boldsymbol{v}\big) + (c\boldsymbol{u}, \boldsymbol{v}) - (p, \operatorname{div} \boldsymbol{v}) &= (\boldsymbol{f}, \boldsymbol{v}), \\ (q, \operatorname{div} \boldsymbol{u}) &= 0. \end{aligned} \tag{2.1}$$

Note that $\operatorname{div} \boldsymbol{b} = 0$ implies

$$\big((\boldsymbol{b}\cdot\nabla)\boldsymbol{u}, \boldsymbol{v}\big) = -\big((\boldsymbol{b}\cdot\nabla)\boldsymbol{v}, \boldsymbol{u}\big) \qquad \forall\, \boldsymbol{u}, \boldsymbol{v} \in V, \tag{2.2}$$

in particular,

$$\big((\boldsymbol{b}\cdot\nabla)\boldsymbol{v}, \boldsymbol{v}\big) = 0 \qquad \forall\, \boldsymbol{v} \in V. \tag{2.3}$$

Thus, applying the Lax–Milgram lemma in the subspace of divergence-free functions, we establish the unique velocity field $\boldsymbol{u}$. The unique pressure $p \in Q$ such that $(\boldsymbol{u}, p)$ solves (2.1) follows from the Babuška–Brezzi condition for the pair (V, Q); see Girault & Raviart (1986, Ch. I).

We will make some assumptions about a solution decomposition of $\boldsymbol{u} = (u_1, u_2)$ and p in order to carry out our analysis. We assume that u_1 and u_2 will consist of layer and regular components while the pressure p will not show any layer behaviour. To be more precise, we assume the following bounds.

ASSUMPTION 2.1 The solution $\boldsymbol{u} = (u_1, u_2)$ can be decomposed as

$$u_1 = g_1 + v_1 + w_1 + z_1 \quad \text{and} \quad u_2 = g_2 + v_2 + w_2 + z_2,$$

where we have for fixed $k \in \mathbb{N}$, all $x, y \in [0, 1]$ and $0 \le i + j \le k + 1$ the pointwise estimates

$$\begin{aligned}
\left|\partial_x^i \partial_y^j g_1(x,y)\right| &\le C, & \left|\partial_x^i \partial_y^j g_2(x,y)\right| &\le C,\\
\left|\partial_x^i \partial_y^j v_1(x,y)\right| &\le C\varepsilon^{1-i}\, e^{-\beta_1 x/\varepsilon}, & \left|\partial_x^i \partial_y^j v_2(x,y)\right| &\le C\varepsilon^{-i}\, e^{-\beta_1 x/\varepsilon},\\
\left|\partial_x^i \partial_y^j w_1(x,y)\right| &\le C\varepsilon^{-i}\, e^{-\beta_2 y/\varepsilon}, & \left|\partial_x^i \partial_y^j w_2(x,y)\right| &\le C\varepsilon^{1-i}\, e^{-\beta_2 y/\varepsilon},\\
\left|\partial_x^i \partial_y^j z_1(x,y)\right| &\le C\varepsilon^{1-(i+j)}\, e^{-\beta_1 x/\varepsilon}\, e^{-\beta_2 y/\varepsilon}, & \left|\partial_x^i \partial_y^j z_2(x,y)\right| &\le C\varepsilon^{1-(i+j)}\, e^{-\beta_1 x/\varepsilon}\, e^{-\beta_2 y/\varepsilon},
\end{aligned}$$

where g_1, g_2 are the regular parts, v_1, v_2 the exponential boundary layers in the x-direction, w_1, w_2 the exponential boundary layers in the y-direction and z_1, z_2 the corner layers. The bounds

$$\left|\partial_x^i \partial_y^j p(x,y)\right| \le C, \quad 0 \le i + j \le k,$$

are assumed for the pressure p.

REMARK 2.2 The divergence constraint guarantees the existence of a scalar stream function ψ such that $\boldsymbol{u} = (\partial_y \psi, -\partial_x \psi)$. Inserting this into the Oseen equations and applying the (two-dimensional) curl operator leads to a fourth-order problem. In Franz *et al.* (2016), it was established that its solution has a principle structure that allows us to derive the assumptions stated above providing that enough compatibility conditions on the data of our problem are fulfilled. The nature of these additional assumptions is similar to those of convection–diffusion problems (Kellogg & Stynes, 2005, 2007). Furthermore, our numerical studies confirm the assumed decomposition.

Concerning the pressure, our analysis does also work if the pressure has strong boundary layers similar to those of u_1 or u_2. Then we would estimate p similarly to $\boldsymbol{u}$ which is possible if we use the same layer-adapted mesh for all solution components.

With Assumption 2.1 in mind, we are able to construct *a priori* adapted meshes. The solution $\boldsymbol{u}$ has weak and strong layers. In order to resolve the strong layers, we use fine meshes along $y = 0$ for u_1 and along $x = 0$ for u_2. Weak layers are not visible at first sight. In the process of a numerical analysis, their unpleasant ε-dependence comes into play via derivatives. Thus, the numerical analysis works best if the weak layers are also resolved. Therefore, we will use layer-adapted meshes that are fine at $x = 0$ and at $y = 0$ for both components u_1 and u_2. For simplicity, these meshes are also used for p although we did assume that p has no boundary layers.

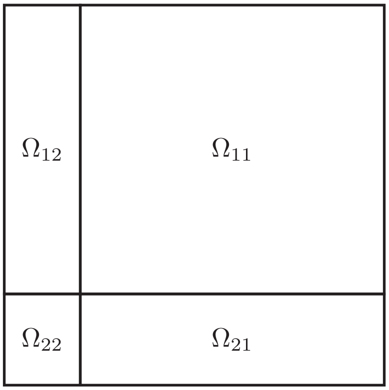

FIG. 2. Decomposition of Ω into subregions.

We will use so-called S-type meshes; see Roos & Linß (1999). They are constructed using transition points where the mesh changes between fine and coarse. To be more precise, let us assume the number N of mesh intervals in each coordinate direction is even. Then we define the transition parameters

$$\lambda_x := \min\left\{\frac{1}{2}, \frac{\sigma\varepsilon}{\beta_1}\ln N\right\} \quad \text{and} \quad \lambda_y := \min\left\{\frac{1}{2}, \frac{\sigma\varepsilon}{\beta_2}\ln N\right\}$$

with some user-chosen positive parameter σ. For the sake of mere simplicity in our subsequent analysis, we assume that

$$\varepsilon < \frac{\min\{\beta_1, \beta_2\}}{2\sigma \ln N} \quad \Leftrightarrow \quad N < \exp\frac{\min\{\beta_1, \beta_2\}}{2\sigma\varepsilon}$$

such that λ_x and λ_y are smaller than $1/2$. This is the typical case for (1.1) as otherwise a standard analysis on an equidistant mesh suffices. Using the transition parameters λ_x and λ_y, the domain Ω is divided into the subdomains $\Omega_{11}, \Omega_{21}, \Omega_{12}$ and Ω_{22} according to

$$\begin{aligned} &\Omega_{11} := [\lambda_x, 1] \times [\lambda_y, 1], \quad \Omega_{12} := [0, \lambda_x] \times [\lambda_y, 1], \\ &\Omega_{21} := [\lambda_x, 1] \times [0, \lambda_y] \quad \text{and} \quad \Omega_{22} := [0, \lambda_x] \times [0, \lambda_y]; \end{aligned}$$

see Fig. 2.

Note that Ω_{11} covers the nonlayer region, Ω_{22} the corner layers and Ω_{12} and Ω_{21} the two boundary layer regions.

With the help of a *mesh-generating function* ϕ which is monotonically increasing with $\phi(0) = 0$ and $\phi(1/2) = \ln N$, we construct the layer-adapted mesh as a tensor product mesh with nodes according to

$$x_i := \begin{cases} \frac{\sigma\varepsilon}{\beta_1}\phi\left(\frac{i}{N}\right), & i = 0, \ldots, N/2, \\ 1 - 2(1-\lambda_x)(1-\frac{i}{N}), & i = N/2, \ldots, N, \end{cases}$$

TABLE 1 *Some examples of mesh-generating and mesh-characterizing functions of S-type meshes*

Name	$\phi(t)$	$\max\phi'$	$\psi(t)$	$\max\lvert\psi'\rvert$
Shishkin mesh	$2t\ln N$	$2\ln N$	N^{-2t}	$2\ln N$
Bakhvalov–Shishkin mesh	$-\ln(1-2t(1-N^{-1}))$	$2N$	$1-2t(1-N^{-1})$	2
Modified Bakhvalov–Shishkin mesh	$\frac{t}{q-t}$, $q=\frac{1}{2}(1+\frac{1}{\ln N})$	$3(\ln N)^2$	$e^{-t/q-t}$	$3/(2q)\le 3$

and

$$y_j := \begin{cases} \frac{\sigma\varepsilon}{\beta_2}\phi\left(\frac{j}{N}\right), & j=0,\ldots,N/2,\\ 1-2(1-\lambda_y)(1-\frac{j}{N}), & j=N/2,\ldots,N. \end{cases}$$

Thus, the dimensions of cells in the layer regions depend on ϕ while the mesh is uniform outside. Related to the mesh-generating function ϕ, we define the *mesh-characterizing function* ψ by

$$\phi = -\ln\psi \text{ which implies } \psi = e^{-\phi}.$$

The derivative of this function will be used to characterize the convergence behaviour on general S-type meshes. Some examples of S-type meshes from Roos & Linß (1999) are presented in Table 1.

In order to analyse the convergence of numerical methods on these meshes, the mesh-generating function ϕ has to fulfil an additional assumption; see Roos & Linß (1999).

ASSUMPTION 2.3 Let the mesh-generating function ϕ be piecewise differentiable such that

$$\max_{t\in[0,1/2]}\phi'(t)\le CN \text{ or equivalently } \max_{t\in[0,1/2]}\frac{|\psi'(t)|}{\psi(t)}\le CN \tag{2.4}$$

is fulfilled.

Note that (2.4) is satisfied for all meshes given in Table 1.

We will now use (2.4) to bound the mesh width from above inside the layers. Let $h_i := x_i - x_{i-1}$ and $t_i = i/N$ for $i=1,\ldots,N/2$. Then it holds for $t\in[t_{i-1},t_i]$ (with $\max\phi'$ taken over $t\in[t_{i-1},t_i]$),

$$\psi(t_i) = e^{-\phi(t_i)} = e^{-(\phi(t_i)-\phi(t))}e^{-\phi(t)} \ge e^{-(\phi(t_i)-\phi(t_{i-1}))}\psi(t) \ge e^{-N^{-1}\max\phi'}\psi(t)\ge C\psi(t), \tag{2.5}$$

where we used (2.4) for the last estimate. Furthermore, we have

$$x = \frac{\sigma\varepsilon}{\beta_1}\phi(t) = -\frac{\sigma\varepsilon}{\beta_1}\ln\psi(t) \quad \text{which gives} \quad \psi(t) = e^{-\beta_1 x/(\sigma\varepsilon)}.$$

Using this, the monotonicity of ψ and (2.5), we obtain for $i = 1,\dots,N/2$ and $x \in [x_{i-1}, x_i]$,

$$\begin{aligned} h_i = \frac{\sigma\varepsilon}{\beta_1}(\phi(t_i) - \phi(t_{i-1})) &\le \frac{\sigma}{\beta_1}\varepsilon N^{-1} \max_{t\in[t_{i-1},t_i]} \phi'(t) \le \frac{\sigma}{\beta_1}\varepsilon N^{-1}\left(\max_{t\in[t_{i-1},t_i]} |\psi'(t)|\right)/\psi(t_i) \\ &\le C\varepsilon N^{-1}\left(\max_{t\in[t_{i-1},t_i]} |\psi'(t)|\right)/\psi(t) \le C\varepsilon N^{-1} \max|\psi'|\, e^{\beta_1 x/(\sigma\varepsilon)} \end{aligned} \tag{2.6}$$

with $\max|\psi'| := \max_{t\in[0,1/2]} |\psi'(t)|$. Similarly, we get for $j = 1,\dots,N/2$,

$$\ell_j := y_j - y_{j-1} \le C\varepsilon N^{-1} \max|\psi'|\, e^{\beta_2 y/(\sigma\varepsilon)} \tag{2.7}$$

with $y \in [y_{j-1}, y_j]$. Of course, the simpler bounds

$$\begin{aligned} h_i &\le C\varepsilon N^{-1} \max\phi' \le C\varepsilon, \qquad i = 1,\dots,N/2, \\ \ell_j &\le C\varepsilon N^{-1} \max\phi' \le C\varepsilon, \qquad j = 1,\dots,N/2, \end{aligned}$$

follow also from (2.4). Let us denote by

$$h := \max_{i=1,\dots,N/2} h_i \quad \text{and} \quad \ell := \max_{j=1,\dots,N/2} \ell_j$$

the maximal dimensions of cells in the layer region perpendicular to the boundaries. We will see later on that the errors will be bounded by powers of the quantity

$$E(N,\varepsilon) := h + \ell + N^{-1}\max|\psi'|, \tag{2.8}$$

which depends on h, ℓ, N and ψ. By the above considerations, we have

$$E(N,\varepsilon) \le \begin{cases} C\varepsilon + N^{-1}\max|\psi'| & \text{for a general S-type mesh,} \\ CN^{-1}\ln N, & \text{for a Shishkin mesh.} \end{cases}$$

Thus, for $\varepsilon \le CN^{-1}$ this quantity is (almost) of order 1 and especially for the Bakhvalov–Shishkin and the modified Bakhvalov–Shishkin meshes we obtain

$$E(N,\varepsilon) \le CN^{-1}.$$

We denote in the following by $\tau_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j]$ a specific element with dimensions h_i and ℓ_j and by τ a generic mesh rectangle. Note that the mesh cells are assumed to be closed.

3. Abstract problem and its convergence analysis

We are interested in conforming discretizations of the Oseen problem, i.e., we choose finite element spaces $V_N \subset V$ and $Q_N \subset Q$. For this section, we will not specify these discrete spaces and consider examples in Section 4.

The standard Galerkin formulation of (2.1) reads as follows:
Find $(\boldsymbol{u}_N, p_N) \in V_N \times Q_N$ such that

$$\begin{aligned}\varepsilon(\nabla \boldsymbol{u}_N, \nabla \boldsymbol{v}_N) - ((\boldsymbol{b} \cdot \nabla)\boldsymbol{u}_N, \boldsymbol{v}_N) + (c\boldsymbol{u}_N, \boldsymbol{v}_N) - (p_N, \operatorname{div} \boldsymbol{v}_N) &= (\boldsymbol{f}, \boldsymbol{v}_N),\\ (q_N, \operatorname{div} \boldsymbol{u}_N) &= 0\end{aligned} \tag{3.1}$$

for all $(\boldsymbol{v}_N, q_N) \in V_N \times Q_N$.

In order to ensure the unique solvability of the discrete problem (3.1), the discrete spaces V_N and Q_N have to fulfil a compatibility condition.

ASSUMPTION 3.1 For any given N and ε, there exists a positive number $\beta_0(N, \varepsilon)$ such that

$$\inf_{q_N \in Q_N} \sup_{\boldsymbol{v}_N \in V_N} \frac{(q_N, \operatorname{div} \boldsymbol{v}_N)}{\|\nabla \boldsymbol{v}_N\| \|q_N\|} \geq \beta_0(N, \varepsilon) > 0 \tag{3.2}$$

holds true.

Note that in the assumption above, we assume only a lower positive bound for fixed N and ε, not a uniform bound in N and ε that would be an inf–sup or Babuška–Brezzi condition. The ε-dependence of $\beta_0(N, \varepsilon)$ can be caused only via the underlying mesh which depends on ε.

If $\boldsymbol{f}$ contains a large irrotational part in the sense of the Helmholtz decomposition, the numerical accuracy of (3.1) is known to be poor; see Linke (2014). In order to overcome this, the grad-div stabilization term $(\gamma \operatorname{div} \boldsymbol{u}_N, \gamma \operatorname{div} \boldsymbol{v}_N)$ will be added where $\gamma \in L^\infty(\Omega)$ fulfils $\gamma \geq \gamma_0 > 0$ in $\overline{\Omega}$ unless specified otherwise. Our stabilized discrete formulation looks as follows:
Find $(\boldsymbol{u}_N, p_N) \in V_N \times Q_N$ such that

$$\begin{aligned}\varepsilon(\nabla \boldsymbol{u}_N, \nabla \boldsymbol{v}_N) - ((\boldsymbol{b} \cdot \nabla)\boldsymbol{u}_N, \boldsymbol{v}_N) + (c\boldsymbol{u}_N, \boldsymbol{v}_N) + (\gamma \operatorname{div} \boldsymbol{u}_N, \gamma \operatorname{div} \boldsymbol{v}_N) - (p_N, \operatorname{div} \boldsymbol{v}_N) &= (\boldsymbol{f}, \boldsymbol{v}_N),\\ (q_N, \operatorname{div} \boldsymbol{u}_N) &= 0\end{aligned} \tag{3.3}$$

for all $(\boldsymbol{v}_N, q_N) \in V_N \times Q_N$.

We define the bilinear form

$$A((\boldsymbol{u}, p), (\boldsymbol{v}, q)) := \varepsilon(\nabla \boldsymbol{u}, \nabla \boldsymbol{v}) - ((\boldsymbol{b} \cdot \nabla)\boldsymbol{u}, \boldsymbol{v}) + (c\boldsymbol{u}, \boldsymbol{v}) + (\gamma \operatorname{div} \boldsymbol{u}, \gamma \operatorname{div} \boldsymbol{v}) - (p, \operatorname{div} \boldsymbol{v}) + (q, \operatorname{div} \boldsymbol{u}) \tag{3.4}$$

on the product space $V \times Q$ and the norms

$$|||\boldsymbol{v}||| := \left(\varepsilon \|\nabla \boldsymbol{v}\|^2 + c_0 \|\boldsymbol{v}\|^2 + \|\gamma \operatorname{div} \boldsymbol{v}\|^2\right)^{1/2}, \qquad |||(\boldsymbol{v}, q)||| := \left(|||\boldsymbol{v}|||^2 + \alpha \|q\|^2\right)^{1/2} \tag{3.5}$$

on V and $V \times Q$ with $\alpha = \alpha(\varepsilon, c_0, c_\infty, b_\infty)$ given by

$$\alpha = \frac{1}{\max\left\{\varepsilon + \dfrac{c_\infty^2}{c_0} C_{\mathrm{F}}^2, b_\infty^2\right\}},$$

where C_{F} denotes the constant in the Friedrichs' inequality for the domain Ω, $c_\infty := \|c\|_\infty$ and $b_\infty := \|\boldsymbol{b}\|_\infty$. Note that the term c_∞^2 / c_0 is bounded since we assumed $0 < c_0 \leq c$ in Ω and $c \in L^\infty(\Omega)$ in the introduction.

LEMMA 3.2 Let $(\boldsymbol{u}, p)$ and $(\boldsymbol{u}_N, p_N)$ denote the solutions of (2.1) and (3.3), respectively. Then, the Galerkin orthogonality

$$A\big((\boldsymbol{u}-\boldsymbol{u}_N, p-p_N), (\boldsymbol{v}_N, q_N)\big) = 0 \tag{3.6}$$

holds true for all $(\boldsymbol{v}_N, q_N) \in V_N \times Q_N$.

Proof. Subtract (3.3) from (2.1), use $V_N \subset V$, $Q_N \subset Q$ and $\operatorname{div} \boldsymbol{u} = 0$. □

Our analysis relies on an upper bound for the parameter γ of the grad-div stabilization.

ASSUMPTION 3.3 Let γ be chosen such that

$$\gamma_\infty^2 \le \frac{C_\gamma}{2\alpha}$$

holds true with some positive constant C_γ where $\gamma_\infty := \|\gamma\|_\infty$.

Let us define for any $q_N \in Q_N$ the positive number

$$B_0(N,\varepsilon,q_N) := \begin{cases} \dfrac{1}{\|q_N\|} \sup\limits_{\boldsymbol{v}_N \in V_N} \dfrac{(q_N, \operatorname{div} \boldsymbol{v}_N)}{\|\nabla \boldsymbol{v}_N\|}, & q_N \neq 0, \\ 1, & q_N = 0. \end{cases} \tag{3.7}$$

Note that by definition and Assumption 3.1 it follows that

$$B_0(N,\varepsilon,q_N) \ge \beta_0(N,\varepsilon).$$

The stability of the bilinear form A can be shown similarly to Matthies & Tobiska (2015, Lemma 3.1). Since we have to take into consideration that here the constant α is defined differently, c and γ are no longer constant and we will use $B_0(N,\varepsilon,q_N)$ instead of $\beta_0(N,\varepsilon)$; the proof will be given here.

LEMMA 3.4 Under Assumptions 3.1 and 3.3, there exists for each $N > 0$, $\varepsilon > 0$ and each pair $(\boldsymbol{v}_N, q_N) \in V_N \times Q_N$ a positive number $B(N,\varepsilon,q_N)$ such that

$$|||(\boldsymbol{v}_N, q_N)||| \le \frac{1}{B(N,\varepsilon,q_N)} \sup_{(\boldsymbol{w}_N, r_N)} \frac{A\big((\boldsymbol{v}_N, q_N), (\boldsymbol{w}_N, r_N)\big)}{|||(\boldsymbol{w}_N, r_N)|||} \tag{3.8}$$

holds true.

Proof. Let $(\boldsymbol{v}_N, q_N)$ be an arbitrary element of $V_N \times Q_N$ and $B_0 = B_0(N,\varepsilon,q_N)$. We obtain immediately

$$A((\boldsymbol{v}_N, q_N), (\boldsymbol{v}_N, q_N)) \ge \varepsilon\|\nabla \boldsymbol{v}_N\|^2 + c_0\|\boldsymbol{v}_N\|^2 + \|\gamma \operatorname{div} \boldsymbol{v}_N\|^2 = |||\boldsymbol{v}_N|||^2\,. \tag{3.9}$$

It follows from the definition (3.7) of $B_0(N,\varepsilon,q_N)$ that for $q_n \in Q_n$ there exists a $z_N = z_N(q_N) \in V_N$ such that

$$(\operatorname{div} z_N, q_N) = -\alpha \|q_N\|^2, \qquad \|\nabla z_N\| = \frac{\alpha}{B_0}\|q_N\|, \tag{3.10}$$

where we already have used the stronger estimate with B_0 instead of $\beta_0(N,\varepsilon)$. Hence, we have

$$A((\boldsymbol{v}_N,q_N),(z_N,0)) = \varepsilon(\nabla \boldsymbol{v}_N,\nabla z_N) - ((\boldsymbol{b}\cdot\nabla)\boldsymbol{v}_N, z_N) + (c\boldsymbol{v}_N, z_N) + (\gamma \operatorname{div} \boldsymbol{v}_N, \gamma \operatorname{div} z_N) + \alpha\|q_N\|^2.$$

Now, we estimate the first four terms of that equation. We obtain directly using Friedrichs' inequality and (3.10),

$$\begin{aligned}
|\varepsilon(\nabla \boldsymbol{v}_N,\nabla z_N) + (c\boldsymbol{v}_N, z_N)| &\le \left(\varepsilon\|\nabla \boldsymbol{v}_N\|^2 + c_0\|\boldsymbol{v}_N\|^2\right)^{1/2}\left(\varepsilon\|\nabla z_N\|^2 + \frac{c_\infty^2}{c_0}\|z_N\|^2\right)^{1/2}\\
&\le \left(\varepsilon + \frac{c_\infty^2}{c_0}C_{\mathrm{F}}^2\right)^{1/2}\frac{\alpha}{B_0}\|q_N\|\,|||\boldsymbol{v}_N|||\\
&\le \frac{\alpha}{4}\|q_N\|^2 + \frac{1}{B_0^2}|||\boldsymbol{v}_N|||^2
\end{aligned}$$

and

$$\begin{aligned}
|-((\boldsymbol{b}\cdot\nabla)\boldsymbol{v}_N, z_N)| = |((\boldsymbol{b}\cdot\nabla)z_N, \boldsymbol{v}_N)| &\le b_\infty\|\nabla z_N\|\|\boldsymbol{v}_N\| \le \frac{\alpha}{4}\|q_N\|^2 + \frac{\alpha b_\infty^2}{B_0^2}\|\boldsymbol{v}_N\|^2\\
&\le \frac{\alpha}{4}\|q_N\|^2 + \frac{1}{B_0^2}|||\boldsymbol{v}_N|||^2,
\end{aligned}$$

where the definition of α was used. For the stabilization term, we are using (3.10) and obtain

$$\begin{aligned}
(\gamma \operatorname{div} \boldsymbol{v}_N, \gamma \operatorname{div} z_N) &\le \|\gamma \operatorname{div} \boldsymbol{v}_N\|\|\gamma \operatorname{div} z_N\|\\
&\le |||\boldsymbol{v}_N|||\sqrt{2}\,\gamma_\infty\|\nabla z_N\| \le |||\boldsymbol{v}_N|||\frac{C_\gamma^{1/2}}{\alpha^{1/2}}\frac{\alpha}{B_0}\|q_N\|\\
&\le \frac{\alpha}{4}\|q_N\|^2 + \frac{C_\gamma}{B_0^2}|||\boldsymbol{v}_N|||^2
\end{aligned}$$

with the constant C_γ from Assumption 3.3. Combining above estimates we get

$$A((\boldsymbol{v}_N,q_N),(z_N,0)) \ge \frac{\alpha}{4}\|q_N\|^2 - \frac{\widetilde{C}}{4}|||\boldsymbol{v}_N|||^2, \tag{3.11}$$

where $\tilde{C} := 4(2+C_\gamma)/B_0^2$.

Thus, we define for $(\boldsymbol{v}_N,q_N) \in V_N \times Q_N$ the pair $(\boldsymbol{w}_N, r_N) \in V_N \times Q_N$ by

$$(\boldsymbol{w}_N, r_N) := (\boldsymbol{v}_N,q_N) + \frac{4}{1+\widetilde{C}}(z_N,0)$$

and obtain with (3.9) and (3.11),

$$A((\boldsymbol{v}_N,q_N),(\boldsymbol{w}_N,r_N)) \geq \frac{\alpha}{1+\widetilde{C}}\|q_N\|^2 + \left(1-\frac{\widetilde{C}}{1+\widetilde{C}}\right)|||\boldsymbol{v}_N|||^2 = \frac{1}{1+\widetilde{C}}\,|||(\boldsymbol{v}_N,q_N)|||^2\,.$$

It remains to show that $|||(\boldsymbol{w}_N,r_N)||| \leq C\,|||(\boldsymbol{v}_N,q_N)|||$. To this end, we estimate

$$|||(\boldsymbol{w}_N,r_N)||| \leq |||(\boldsymbol{v}_N,q_N)||| + \frac{4}{1+\widetilde{C}}\,|||\boldsymbol{z}_N|||$$

and with Friedrichs' inequality and (3.10),

$$|||\boldsymbol{z}_N|||^2 = \varepsilon\|\nabla \boldsymbol{z}_N\|^2 + c_0\|\boldsymbol{z}_N\|^2 + \|\gamma\,\mathrm{div}\,\boldsymbol{z}_N\|^2 \leq \left(\varepsilon+\frac{c_\infty^2}{c_0}C_{\mathrm{F}}^2\right)\|\nabla \boldsymbol{z}_N\|^2 + \frac{C_\gamma\alpha}{B_0^2}\|q_N\|^2 \leq \frac{\widetilde{C}}{4}\alpha\|q_N\|^2.$$

Taking into account that $\sup_{0<t<\infty}\frac{4}{1+t}\sqrt{\frac{t}{4}} = 1$ we get

$$|||(\boldsymbol{w}_N,r_N)||| \leq 2\,|||(\boldsymbol{v}_N,q_N)|||\,.$$

Hence, with

$$B(N,\varepsilon,q_N) := \frac{1}{2\left(1+\widetilde{C}\right)} = \frac{B_0(N,\varepsilon,q_N)^2}{2B_0(N,\varepsilon,q_N)^2+16+8C_\gamma},$$

the stated condition holds. □

3.1 *Interpolation error*

ASSUMPTION 3.5 There exists an interpolation operator $I=(I_1,I_2):(C(\Omega))^2\to V_N$ such that the anisotropic interpolation error bounds

$$\|w_m - I_m w_m\|_{L^p(\tau_{ij})} \leq C\left(h_i^{k+1}\left\|\partial_x^{k+1}w_m\right\|_{L^p(\tau_{ij})} + \ell_j^{k+1}\left\|\partial_y^{k+1}w_m\right\|_{L^p(\tau_{ij})}\right),\quad m=1,2, \tag{3.12}$$

$$\|(w_m - I_m w_m)_x\|_{L^p(\tau_{ij})} \leq C\left(h_i^{k}\left\|\partial_x^{k+1}w_m\right\|_{L^p(\tau_{ij})} + \ell_j^{k}\left\|\partial_x^{k}\partial_y w_m\right\|_{L^p(\tau_{ij})}\right),\quad m=1,2, \tag{3.13}$$

and similarly for the y-derivative, hold true for $p\in[1,\infty]$ and arbitrary $\boldsymbol{w}=(w_1,w_2)\in(W^{k+1,p}(\Omega))^2$. Furthermore, we have the stability estimates

$$\|I_m w_m\|_{\infty,\tau} \leq C\|w_m\|_{\infty,\tau},\quad m=1,2, \tag{3.14}$$

and

$$\|\partial_x I_m w_m\|_{\infty,\tau} \leq C\|\partial_x w_m\|_{\infty,\tau} \quad\text{and}\quad \|\partial_y I_m w_m\|_{\infty,\tau} \leq C\|\partial_y w_m\|_{\infty,\tau},\quad m=1,2. \tag{3.15}$$

ASSUMPTION 3.6 There exists an interpolation operator $J : C(\Omega) \to Q_N$ such that

$$\|q - Jq\|_\infty \le CE(N,\varepsilon)^k$$

hold true for arbitrary $q \in H^k(\Omega)$ where $E(N,\varepsilon)$ is given by (2.8).

Using these assumptions on the existence of interpolation operators, we estimate the interpolation errors and their derivatives in the next two lemmas.

LEMMA 3.7 Let $\sigma > k+1$, $r = v_1 + w_1 + z_1$ and $s = v_2 + w_2 + z_2$ from Assumption 2.1, $I = (I_1, I_2)$ the interpolation operator from Assumption 3.5, J the interpolation operator from Assumption 3.6 and $E(N,\varepsilon)$ be given by (2.8). Then, the interpolation errors in the L^2-norm can be bounded by

$$\|r - I_1 r\|_{\Omega_{11}} + \|s - I_2 s\|_{\Omega_{11}} \le CN^{-\sigma}\left(\varepsilon^{1/2} + N^{-1/2}\right), \tag{3.16}$$

$$\|r - I_1 r\|_{\Omega\setminus\Omega_{11}} + \|s - I_2 s\|_{\Omega\setminus\Omega_{11}} \le C\varepsilon^{1/2} E(N,\varepsilon)^{k+1}. \tag{3.17}$$

Moreover, the estimates

$$\|g_1 - I_1 g_1\|_{\Omega_{11}} + \|g_2 - I_2 g_2\|_{\Omega_{11}} \le CN^{-(k+1)},$$
$$\|g_1 - I_1 g_1\|_{\Omega\setminus\Omega_{11}} + \|g_2 - I_2 g_2\|_{\Omega\setminus\Omega_{11}} \le CE(N,\varepsilon)^{k+1}$$

hold true for the smooth parts g_1 and g_2.

Proof. Since the proof of the error bounds for s follows the same line as the proof for r, we restrict ourselves to the latter one. We start with Ω_{11}. Clearly, we have

$$\|r - I_1 r\|_{\Omega_{11}} \le \|r\|_{\Omega_{11}} + \|I_1 r\|_{\Omega_{11}}. \tag{3.18}$$

With Assumption 2.1, we get for $\|r\|_{\Omega_{11}}$ directly,

$$\|r\|_{\Omega_{11}} \le \|v_1\|_{\Omega_{11}} + \|w_1\|_{\Omega_{11}} + \|z_1\|_{\Omega_{11}} \le C\varepsilon^{1/2} N^{-\sigma}, \tag{3.19}$$

where we used integration and the definition of λ_x and λ_y. Let $\overline{h}$ and $\overline{\ell}$ denote the mesh width and height in Ω_{11}. In order to estimate $I_1 r$, we use an idea from Stynes & Tobiska (2003) and divide the domain Ω_{11} into $S = [\lambda_x + \overline{h}, 1] \times [\lambda_y + \overline{\ell}, 1]$ and $\Omega_{11} \setminus S$ with the mesh transition points λ_x and λ_y. Note that $\Omega_{11} \setminus S$ consists of only a single ply of $N-1$ mesh elements and $\overline{h}, \overline{\ell} \le CN^{-1}$. Thus, we obtain

$$\|I_1 r\|^2_{\Omega_{11}\setminus S} \le \sum_{\tau\subset\Omega_{11}\setminus S} \overline{h}\,\overline{\ell}\|I_1 r\|^2_{\infty,\tau} \le CN^{-1}\|r\|^2_{\infty,\Omega_{11}} \le CN^{-2\sigma-1} \tag{3.20}$$

using the stability (3.14) and in S,

$$\|I_1 r\|_S^2 \le C\left\{\int_{\Omega_{11}}\left(\varepsilon^2 e^{-2\beta_1 x/\varepsilon} + e^{-2\beta_2 y/\varepsilon}\left(1+\varepsilon^2 e^{-2\beta_1 x/\varepsilon}\right)\right)\right\} \le C\varepsilon N^{-2\sigma}. \tag{3.21}$$

Combining (3.18)–(3.21) completes the proof for (3.16).

On Ω_{12}, we get by using (3.12), (2.6) and (2.7),

$$\|v_1 - I_1 v_1\|_{\Omega_{12}}^2 \le C\left(\left\|h_i^{k+1}\partial_x^{k+1} v_1\right\|_{\Omega_{12}}^2 + \left\|\ell_j^{k+1}\partial_y^{k+1} v_1\right\|_{\Omega_{12}}^2\right) \le C\varepsilon^3 (N^{-1}\max|\psi'|)^{2(k+1)}.$$

Using the stability (3.14), we obtain for w_1 and z_1,

$$\|(w_1+z_1) - I_1(w_1+z_1)\|_{\Omega_{12}} \le C(\operatorname{meas}\Omega_{12})^{1/2}\|w_1+z_1\|_{\infty,\Omega_{12}} \le C\varepsilon^{1/2}N^{-\sigma}(\ln N)^{1/2}.$$

Similarly we have on Ω_{21},

$$\|w_1 - I_1 w_1\|_{\Omega_{21}}^2 \le C\varepsilon(N^{-1}\max|\psi'|)^{2(k+1)}$$

and

$$\|(v_1+z_1) - I_1(v_1+z_1)\|_{\Omega_{21}} \le C\varepsilon^{3/2}N^{-\sigma}(\ln N)^{1/2}.$$

Using the anisotropic interpolation error bounds (3.12), we get on Ω_{22} analogously,

$$\begin{aligned}
\|v_1 - I_1 v_1\|_{\Omega_{22}}^2 &\le C\varepsilon^4 \ln N(N^{-1}\max|\psi'| + \ell)^{2(k+1)},\\
\|w_1 - I_1 w_1\|_{\Omega_{22}}^2 &\le C\varepsilon^2 \ln N(h + N^{-1}\max|\psi'|)^{2(k+1)},\\
\|z_1 - I_1 z_1\|_{\Omega_{22}}^2 &\le C\varepsilon^4 (N^{-1}\max|\psi'|)^{2(k+1)}.
\end{aligned}$$

Together with $\sigma > k+1$ and the definition of $E(N,\varepsilon)$ we get (3.17).

For the smooth part, we present only the proofs for g_1 since g_2 can be estimated similarly. We have

$$\begin{aligned}
\|g_1 - I_1 g_1\|_{\Omega_{11}} &\le C(\operatorname{meas}\Omega_{11})^{1/2}\|g_1 - I_1 g_1\|_{\infty,\Omega_{11}} \le C(\overline{h}^{k+1} + \overline{\ell}^{k+1}) \le CN^{-(k+1)},\\
\|g_1 - I_1 g_1\|_{\Omega\setminus\Omega_{11}} &\le C(\operatorname{meas}\Omega\setminus\Omega_{11})^{1/2}\|g_1 - I_1 g_1\|_{\infty,\Omega\setminus\Omega_{11}} \le C(h + \ell + N^{-1})^{k+1},
\end{aligned}$$

where (3.12) in the L^∞-norm was used. □

LEMMA 3.8 Let $\sigma > k+1$ and $E(N,\varepsilon)$ be given by (2.8). Then the interpolation errors in the H^1-seminorm for $\boldsymbol{u} = (u_1,u_2)$ and $I = (I_1,I_2)$ from Assumption 3.5 can be bounded by

$$\begin{aligned}
\|\nabla(\boldsymbol{u} - I\boldsymbol{u})\|_{\Omega_{11}} &\le C\varepsilon^{-1/2}N^{-k},\\
\|\nabla(\boldsymbol{u} - I\boldsymbol{u})\|_{\Omega\setminus\Omega_{11}} &\le C\varepsilon^{-1/2}E(N,\varepsilon)^k
\end{aligned}$$

while the estimates

$$\|\operatorname{div}(\boldsymbol{u} - \boldsymbol{I}\boldsymbol{u})\|_{\Omega_{11}} \le CN^{-k},$$
$$\|\operatorname{div}(\boldsymbol{u} - \boldsymbol{I}\boldsymbol{u})\|_{\Omega\setminus\Omega_{11}} \le CE(N,\varepsilon)^k$$

for the divergence of the interpolation error hold true.

Proof. Since the proofs of the estimates in the H^1-seminorm are similar for u_1 and u_2, we show only the proofs for u_1. With the anisotropic interpolation error bounds (3.13) and its y-derivative counterpart we get for the smooth part,

$$\|\nabla(g_1 - I_1 g_1)\|_{\Omega_{11}} \le C\left(\|(g_1 - I_1 g_1)_x\|_{\infty,\Omega_{11}} + \|(g_1 - I_1 g_1)_y\|_{\infty,\Omega_{11}}\right) \le CN^{-k}$$

and similarly

$$\|\nabla(g_1 - I_1 g_1)\|_{\Omega\setminus\Omega_{11}} \le C(h + \ell + N^{-1})^k.$$

Now, we estimate the layer parts. Consider first the term v_1. Here we have by the $W^{1,\infty}$-stability (3.15),

$$\begin{aligned}
\|(v_1 - I_1 v_1)_x\|_{\Omega_{11}\cup\Omega_{21}} &\le C(\|v_{1,x}\|_{\Omega_{11}\cup\Omega_{21}} + (\operatorname{meas}\Omega_{11}\cup\Omega_{21})^{1/2}\|(I_1 v_1)_x\|_{\infty,\Omega_{11}\cup\Omega_{21}})\\
&\le C(\|v_{1,x}\|_{\Omega_{11}\cup\Omega_{21}} + \|v_{1,x}\|_{\infty,\Omega_{11}\cup\Omega_{21}})\\
&\le CN^{-\sigma}.
\end{aligned}$$

In the remaining regions, we use (3.13) of I_1 to deduce

$$\begin{aligned}
\|(v_1 - I_1 v_1)_x\|_{\Omega_{12}\cup\Omega_{22}} &\le C\left\{\left\|h_i^k \partial_x^{k+1} v_1\right\|_{\Omega_{12}\cup\Omega_{22}} + \left\|\ell_j^k \partial_x \partial_y^k v_1\right\|_{\Omega_{12}\cup\Omega_{22}}\right\}\\
&\le C\varepsilon^{1/2}\left(\ell + N^{-1}\max|\psi'|\right)^k.
\end{aligned}$$

The techniques for the y-derivatives are similar. Hence, we obtain

$$\begin{aligned}
\|(v_1 - I_1 v_1)_y\|_{\Omega_{11}\cup\Omega_{21}} &\le C(\|v_{1,y}\|_{\Omega_{11}\cup\Omega_{21}} + \|I_1 v_{1,y}\|_{\infty,\Omega_{11}\cup\Omega_{21}})\\
&\le C(\|v_{1,y}\|_{\Omega_{11}\cup\Omega_{21}} + \|v_{1,y}\|_{\infty,\Omega_{11}\cup\Omega_{21}})\\
&\le C\varepsilon N^{-\sigma}
\end{aligned}$$

and

$$\begin{aligned}
\|(v_1 - I_1 v_1)_y\|_{\Omega_{12}\cup\Omega_{22}} &\le C\left\{\left\|h_i^k \partial_x^k \partial_y v_1\right\|_{\Omega_{12}\cup\Omega_{22}} + \left\|\ell_j^k \partial_y^{k+1} v_1\right\|_{\Omega_{12}\cup\Omega_{22}}\right\}\\
&\le C\varepsilon^{3/2}(\ell + N^{-1}\max|\psi'|)^k.
\end{aligned}$$

For the second layer part w_1, we get analogously

$$\begin{aligned}
\|(w_1 - I_1 w_1)_x\|_{\Omega_{11}\cup\Omega_{12}} &\le CN^{-\sigma},\\
\|(w_1 - I_1 w_1)_x\|_{\Omega_{21}\Omega_{22}} &\le C\varepsilon^{1/2}(h + N^{-1}\max|\psi'|)^k,\\
\|(w_1 - I_1 w_1)_y\|_{\Omega_{11}\cup\Omega_{12}} &\le C(\varepsilon^{-1/2} + N)N^{-\sigma},\\
\|(w_1 - I_1 w_1)_y\|_{\Omega_{21}\cup\Omega_{22}} &\le C\varepsilon^{-1/2}(h + N^{-1}\max|\psi'|)^k.
\end{aligned}$$

Now, we estimate the error of the function z_1 using the $W^{1,\infty}$-stability (3.15) of I_1,

$$\|(z_1 - I_1 z_1)_x\|_{\Omega\setminus\Omega_{22}} \le C\left(\|z_{1,x}\|_{\Omega\setminus\Omega_{22}} + \|z_{1,x}\|_{\infty,\Omega\setminus\Omega_{22}}\right) \le CN^{-\sigma},$$

and in Ω_{22} with (3.13),

$$\|(z_1 - I_1 z_1)_x\|_{\Omega_{22}} \le C\left\{\left\|h_i^k\partial_x^{k+1}z_1\right\|_{\Omega_{22}} + \left\|\ell_j^k\partial_x\partial_y^k z_1\right\|_{\Omega_{22}}\right\} \le C\varepsilon(N^{-1}\max|\psi'|)^k.$$

The bounds for the y-derivatives follow analogously:

$$\begin{aligned}
\|(z_1 - I_1 z_1)_y\|_{\Omega\setminus\Omega_{22}} &\le CN^{-\sigma},\\
\|(z_1 - I_1 z_1)_y\|_{\Omega_{22}} &\le C\varepsilon(N^{-1}\max|\psi'|)^k.
\end{aligned}$$

Putting all these estimates together, we obtain with the definition of $E(N,\varepsilon)$ the bounds for the H^1-seminorms. We remark that the bounds of the x-derivatives of u_1 are better than those of the y-derivatives. For u_2 and I_2, we obtain similar results with the difference that the x-derivatives are worse than the y-derivatives. Now combining the x-derivatives of u_1 and the y-derivatives of u_2, we get

$$\begin{aligned}
\|\operatorname{div}(\boldsymbol{u} - \boldsymbol{I}\boldsymbol{u})\|_{\Omega_{11}} &\le \|(u_1 - I_1 u_1)_x\|_{\Omega_{11}} + \|(u_2 - I_2 u_2)_y\|_{\Omega_{11}} \le CN^{-k},\\
\|\operatorname{div}(\boldsymbol{u} - \boldsymbol{I}\boldsymbol{u})\|_{\Omega\setminus\Omega_{11}} &\le C(\varepsilon\ln N)^{1/2}(h + \ell + N^{-1}\max|\psi'|)^k \le CE(N,\varepsilon)^k,
\end{aligned}$$

which are the last two statements of this lemma. □

LEMMA 3.9 Let $\sigma > k+1$, $I = (I_1, I_2)$, the interpolation operator from Assumption 3.5 and J the interpolation operator from Assumption 3.6. Then the estimate

$$|||(\boldsymbol{u} - \boldsymbol{I}\boldsymbol{u}, p - Jp)||| \le CE(N,\varepsilon)^k \tag{3.22}$$

holds true where $E(N,\varepsilon)$ is given by (2.8).

Proof. For this proof, we use the previous interpolation error estimates. The definition of the energy norm gives

$$|||(\boldsymbol{u} - \boldsymbol{I}\boldsymbol{u}, p - Jp)|||^2 = \varepsilon\|\nabla(\boldsymbol{u} - \boldsymbol{I}\boldsymbol{u})|_0^2 + c_0\|\boldsymbol{u} - \boldsymbol{I}\boldsymbol{u}\|^2 + \alpha\|p - Jp\|^2 + \|\gamma\operatorname{div}(\boldsymbol{u} - \boldsymbol{I}\boldsymbol{u})\|^2. \tag{3.23}$$

For the first term on the right-hand side of (3.23), we get with Lemma 3.8,

$$\varepsilon^{1/2}\|\nabla(\boldsymbol{u}-I\boldsymbol{u})\| \le \varepsilon^{1/2}\big(\|\nabla(u_1-I_1u_1)\| + \|\nabla(u_2-I_2u_2)\|\big) \le CE(N,\varepsilon)^k.$$

The second term of (3.23) yields with Lemma 3.7,

$$\|\boldsymbol{u}-I\boldsymbol{u}\| \le C\big(\|u_1-I_1u_1\| + \|u_2-I_2u_2\|\big) \le CE(N,\varepsilon)^{k+1}.$$

The bound

$$\|p-Jp\| \le CE(N,\varepsilon)^k$$

for the pressure term follows from Assumption 3.6. The stabilization term can be bounded by the divergence bounds of Lemma 3.8,

$$\|\gamma\,\mathrm{div}(\boldsymbol{u}-I\boldsymbol{u})\| \le \gamma_\infty\|\,\mathrm{div}(\boldsymbol{u}-I\boldsymbol{u})\| \le CE(N,\varepsilon)^k.$$

Summarizing all these estimates finishes the proof. □

3.2 *A priori error estimates*

LEMMA 3.10 Let $\gamma \ge \gamma_0 > 0$ be independent of N and ε, and $\sigma > k+1$. The solutions of (2.1) and (3.3) are denoted by $(\boldsymbol{u},p)$ and $(\boldsymbol{u}_N,p_N)$, respectively. Then, the estimate

$$|||(I\boldsymbol{u}-\boldsymbol{u}_N, Jp-p_N)||| \le C\frac{1}{B(N,\varepsilon,Jp-p_N)}E(N,\varepsilon)^k$$

holds true where $C\to\infty$ for $\gamma_0\to 0$.

Proof. Abbreviating $B = B(N,\varepsilon,Jp-p_N)$ in this proof, the estimate (3.8) gives

$$\begin{aligned}|||(I\boldsymbol{u}-\boldsymbol{u}_N, Jp-p_N)||| &\le \frac{1}{B}\sup_{(\boldsymbol{w}_N,r_N)\in V_N\times Q_N}\frac{A((I\boldsymbol{u}-\boldsymbol{u}_N,Jp-p_N),(\boldsymbol{w}_N,r_N))}{|||(\boldsymbol{w}_N,r_N)|||}\\ &= \frac{1}{B}\sup_{(\boldsymbol{w}_N,r_N)\in V_N\times Q_N}\frac{A((I\boldsymbol{u}-\boldsymbol{u},Jp-p),(\boldsymbol{w}_N,r_N))}{|||(\boldsymbol{w}_N,r_N)|||}\end{aligned}$$

due to the Galerkin orthogonality of Lemma 3.2. The expression

$$\begin{aligned}&A((I\boldsymbol{u}-\boldsymbol{u},Jp-p),(\boldsymbol{w}_N,r_N))\\ &\quad=\varepsilon(\nabla(I\boldsymbol{u}-\boldsymbol{u}),\nabla\boldsymbol{w}_N)-(b\cdot\nabla(I\boldsymbol{u}-\boldsymbol{u}),\boldsymbol{w}_N)+(c(I\boldsymbol{u}-\boldsymbol{u}),\boldsymbol{w}_N)\\ &\qquad-(Jp-p,\mathrm{div}\,\boldsymbol{w}_N)+(r_N,\mathrm{div}(I\boldsymbol{u}-\boldsymbol{u}))+(\gamma\,\mathrm{div}(I\boldsymbol{u}-\boldsymbol{u}),\gamma\,\mathrm{div}(\boldsymbol{w}_N))\end{aligned}$$

will be estimated term by term. We get with the interpolation error bounds of Lemmas 3.7 and 3.8

$$
\begin{aligned}
&|\varepsilon(\nabla(I\boldsymbol{u}-\boldsymbol{u}),\nabla \boldsymbol{w}_N)+(c(I\boldsymbol{u}-\boldsymbol{u}),\boldsymbol{w}_N)|\\
&\quad\le \left(\varepsilon\|\nabla(I\boldsymbol{u}-\boldsymbol{u})\|^2+\frac{c_\infty^2}{c_0}\|I\boldsymbol{u}-\boldsymbol{u}\|^2\right)^{1/2}(\varepsilon\|\nabla \boldsymbol{w}_N\|^2+c_0\|\boldsymbol{w}_N\|^2)^{1/2}\\
&\quad\le CE(N,\varepsilon)^k\,|||(\boldsymbol{w}_N,r_N)|||\,,
\end{aligned}
$$

by Lemma 3.7 for the pressure,

$$
\begin{aligned}
|(Jp-p,\operatorname{div}\boldsymbol{w}_N)| &= |(Jp-p,\operatorname{div}\boldsymbol{w}_N)|\\
&\le \|Jp-p\|\gamma_0^{-1}\,|||(\boldsymbol{w}_N,r_N)|||\\
&\le C\gamma_0^{-1}E(N,\varepsilon)^k\,|||(\boldsymbol{w}_N,r_N)|||
\end{aligned}
$$

and with the divergence bounds of Lemma 3.8,

$$
\begin{aligned}
|(r_N,\operatorname{div}(I\boldsymbol{u}-\boldsymbol{u}))+(\gamma\operatorname{div}(I\boldsymbol{u}-\boldsymbol{u}),\gamma\operatorname{div}\boldsymbol{w}_N))| &\le \|\operatorname{div}(I\boldsymbol{u}-\boldsymbol{u})\|\big(\|r_N\|+\gamma_\infty\|\gamma\operatorname{div}\boldsymbol{w}_N\|\big)\\
&\le (\alpha^{-1/2}+\gamma_\infty)\|\operatorname{div}(I\boldsymbol{u}-\boldsymbol{u})\|\,|||(\boldsymbol{w}_N,r_N)|||\\
&\le C(1+\gamma_\infty)E(N,\varepsilon)^k\,|||(\boldsymbol{w}_N,r_N)|||\,.
\end{aligned}
$$

The last two estimates show that both γ_0^{-1} and γ_∞ should be bounded independent of N and ε. Hence, γ should be contained in the interval $[\gamma_0,\gamma_\infty]$ with end points independent of N and ε.

For the convective term, we estimate the terms of

$$
\begin{aligned}
|(\boldsymbol{b}\cdot\nabla(I\boldsymbol{u}-\boldsymbol{u}),\boldsymbol{w}_N)| \le\;& |(b_1(I_1u_1-u_1)_x,w_{N,1})|+|(b_2(I_1u_1-u_1)_y,w_{N,1})|\\
&+|(b_1(I_2u_2-u_2)_x,w_{N,2})|+|(b_2(I_2u_2-u_2)_y,w_{N,2})|
\end{aligned}
$$

individually. Here we have directly,

$$
\begin{aligned}
|(b_1(I_1u_1-u_1)_x,w_{N,1})| &\le C\|(I_1u_1-u_1)_x\|\|w_{N,1}\|\le CE(N,\varepsilon)^k\,|||(\boldsymbol{w}_N,r_N)|||\,,\\
|(b_2(I_2u_2-u_2)_y,w_{N,2})| &\le C\|(I_2u_2-u_2)_y\|\|w_{N,2}\|\le CE(N,\varepsilon)^k\,|||(\boldsymbol{w}_N,r_N)|||\,.
\end{aligned}
$$

For the other terms, we use integration by parts to obtain

$$
\begin{aligned}
|(b_2(I_1u_1-u_1)_y,w_{N,1})| &\le C\left(\|I_1u_1-u_1\|\|w_{N,1}\|+|(b_2(I_1u_1-u_1),w_{N,1,y})|\right)\\
&\le C\Big(\|I_1u_1-u_1\|\|w_{N,1}\|+\|I_1u_1-u_1\|_{\Omega_{11}}\|w_{N,1,y}\|_{\Omega_{11}}\\
&\qquad+\|I_1u_1-u_1\|_{\Omega\setminus\Omega_{11}}\|w_{N,1,y}\|_{\Omega\setminus\Omega_{11}}\Big)
\end{aligned}
$$

$$\leq C\left(\|u_1 - u_1\| + N\|I_1 u_1 - u_1\|_{\Omega_{11}} + \varepsilon^{-1/2}\|I_1 u_1 - u_1\|_{\Omega\setminus\Omega_{11}}\right) |||(\boldsymbol{w}_N, r_N)|||$$
$$\leq CE(N,\varepsilon)^k\, |||(\boldsymbol{w}_N, r_N)||| \,.$$

The remaining term can be estimated analogously. Upon summarizing the proof is done. □

THEOREM 3.11 Let $\gamma \geq \gamma_0 > 0$ be independent of N and ε, and $\sigma > k+1$. The error between the solution $(\boldsymbol{u}, p)$ of (2.1) and the discrete solution $(\boldsymbol{u}_N, p_N)$ of (3.3) can be bounded by

$$|||(\boldsymbol{u} - \boldsymbol{u}_N, p - p_N)||| \leq C\left(1 + \frac{1}{B(N,\varepsilon, Jp - p_N)}\right) E(N,\varepsilon)^k,$$

where $E(N,\varepsilon)$ is given by (2.8) and J is an interpolation operator fulfilling Assumption 3.6.

Proof. It follows directly from Lemmas 3.9 and 3.10, and the triangle inequality. □

COROLLARY 3.12 If $\gamma = 0$ is chosen on the whole domain, then we have the standard Galerkin method. For the solution of this method we have the estimate

$$|||(\boldsymbol{u} - \boldsymbol{u}_N, p - p_N)||| \leq C\left(1 + \frac{1}{B(N,\varepsilon, Jp - p_N)}\right) (N + (\ln N)^{1/2}) E(N,\varepsilon)^k,$$

i.e., the convergence order reduces by 1 compared to the stabilized case.

Proof. The only difference to the estimation before is the bounding of the term $|(Jp - p, \operatorname{div} \boldsymbol{w}_N)|$. This term cannot be estimated against $\|\gamma \operatorname{div} \boldsymbol{w}_N\|$ due to $\gamma = 0$. Alternatively, we get

$$\begin{aligned}
|(Jp - p, \operatorname{div} \boldsymbol{w}_N)| &= |(Jp - p, \operatorname{div} \boldsymbol{w}_N)| \\
&\leq C(\|Jp - p\|_{\Omega_{11}} \|\nabla \boldsymbol{w}_N\|_{\Omega_{11}} + \|Jp - p\|_{\Omega\setminus\Omega_{11}} \|\nabla \boldsymbol{w}_N\|_{\Omega\setminus\Omega_{11}}) \\
&\leq C(N\|Jp - p\|_{\Omega_{11}} \|\boldsymbol{w}_N\|_{\Omega_{11}} + (\varepsilon \ln N)^{1/2} \|Jp - p\|_{\infty,\Omega\setminus\Omega_{11}} \|\nabla \boldsymbol{w}_N\|_{\Omega\setminus\Omega_{11}}) \\
&\leq C(N\|Jp - p\| + (\ln N)^{1/2}\|Jp - p\|_\infty)\, |||(\boldsymbol{w}_N, r_N)||| \\
&\leq C(N + (\ln N)^{1/2}) E(N,\varepsilon)^k\, |||(\boldsymbol{w}_N, r_N)|||
\end{aligned}$$

and the reduced order can be seen. □

COROLLARY 3.13 If we choose $\gamma = 0$ only in the layer region, we can recover the convergence rate k compared with $\gamma = 0$ on the whole domain.

Proof. See the proof before. In the layer region, we obtain an additional factor of only $(\ln N)^{1/2}$ and not N. □

4. Examples for suitable spaces

We will consider two families of finite element pairs on the introduced meshes. The first one is the Taylor–Hood family $Q_k \times Q_{k-1}$, $k \geq 2$, and the second one is $Q_k \times P_{k-1}^{\mathrm{disc}}$, $k \geq 2$.

On isotropic meshes, the pairs $Q_k \times Q_{k-1}$, $k \geq 2$, and $Q_k \times P_{k-1}^{\mathrm{disc}}$, $k \geq 2$, are known to fulfil Assumption 3.1 even with a constant $\beta_0 > 0$ that is independent of N; see Girault & Raviart (1986, Ch. II), and Assumption 3.1 becomes just the well-known inf–sup or Babuška–Brezzi condition. Since the ε-dependence of $\beta_0(N, \varepsilon)$ comes from the underlying mesh only, there is also no ε-dependence on general isotropic meshes.

On anisotropic meshes, the situation is much more complicated. The existence of an inf–sup constant that is independent of the maximal aspect ratio of the underlying mesh depends strongly on the precise mesh structure and the chosen spaces for approximating velocity and pressure. Proofs for the existence of inf–sup constants are mainly available for situations where the difference between the polynomial orders of velocity and pressure is 2; see Schötzau & Schwab (1998), Schötzau *et al.* (1999) for the cases $Q_k \times Q_{k-2}$ on two-dimensional domains. A review of results on inf–sup constants for the Stokes problem on several types of anisotropic meshes can be found in Apel & Maharavo Randrianarivony (2003). Nonconforming discretizations of the Stokes problem on anisotropic rectangular meshes were investigated in Apel & Matthies (2008). Although the four nonconforming velocity approximation spaces in Apel & Matthies (2008) differ only slightly, just two of them provide an inf–sup constant independent of the aspect ratio. Another approach for generating inf–sup-stable pairs by eliminating certain pressure modes is presented in Ainsworth *et al.* (2015).

In Section 5.1.1, we investigate the values of $\beta_0 = \beta_0(N, \varepsilon)$ and $B_0 = B_0(N, \varepsilon, q_N)$ for both element pairs. It can be observed that for the continuous pressure space, β_0 is bounded independent of ε but not independent of N, while for the discontinuous pressure space it goes to zero for $\varepsilon \to 0$. In contrast to this negative observation, the values of B_0 stay bounded away from zero, independent of N and ε, for both cases if q_N is the difference between the interpolation of p and the discrete solution p_N. This shows that B_0 might be more appropriate to estimate the errors although we do not have a theoretical insight into the boundedness of B_0.

In order to define the interpolation operators, let I_k be the scalar nodal interpolator of order k. Then we set $I = (I_k, I_k)$. In Apel (1999, Theorem 2.7) the anisotropic interpolation error bounds for Assumption 3.5 are given. The stability estimates are standard and can be shown directly.

For the interpolation operator into the discrete pressure space Q_N, we have to distinguish two cases. In the case $Q_N = P_{k-1}^{\mathrm{disc}}$, we can use the L^2-projection into Q_N as J. The global L^2-projection furthermore localizes to L^2-projections on each mesh cell. Hence, the proof of Assumption 3.6 can exploit the anisotropic error estimates for the L^2-projection that follow the same reasoning as in Apel (1999).

In the case $Q_N = Q_{k-1}$, the construction of J is two stage. Let I_{k-1} denote the scalar nodal interpolation into the space of continuous, piecewise Q_{k-1} functions. Then we set

$$Jq = I_{k-1}q - \bar{c}, \quad \text{where} \quad \bar{c} = \frac{1}{|\Omega|}\int_\Omega I_{k-1}q$$

to ensure that J maps into $L_0^2(\Omega)$. Using again the anisotropic interpolation error estimates by Apel (1999), we have

$$\|p - I_{k-1}p\|_{\Omega_{11}} \leq CN^{-k} \quad \text{and} \quad \|p - I_{k-1}p\|_{\Omega\setminus\Omega_{11}} \leq CE(N, \varepsilon)^k.$$

From the definition of J and the fact that constants are $L^2(\Omega)$-orthogonal to functions from $L_0^2(\Omega)$, we get

$$\|p - I_{k-1}p\|^2 = \|p - Jp - \bar{c}\|^2 = \|p - Jp\|^2 + \|\bar{c}\|^2$$

and therefore

$$\|p - Jp\| \le \|p - I_{k-1}p\| \le CE(N,\varepsilon)^k.$$

Furthermore, we obtain in L^∞,

$$\|p - Jp\|_\infty \le \|p - I_{k-1}p\|_\infty + |\bar{c}| \le \|p - I_{k-1}p\|_\infty + \left|\int_\Omega I_{k-1}p - p\right| \le CE(N,\varepsilon)^k$$

which provides Assumption 3.6.

COROLLARY 4.1 For both choices of the discrete pressure space, Corollary 3.12 can be improved slightly.

If $Q_N = P_{k-1}^{\mathrm{disc}}$ then $\|Jp - p\|_{\Omega_{11}} \le CN^{-k}$ and the final lines in the proof of Corollary 3.12 become

$$\begin{aligned}|(Jp - p, \operatorname{div} \boldsymbol{w}_N)| &\le C(N\|Jp - p\|_{\Omega_{11}} + (\ln N)^{1/2}\|Jp - p\|_{\infty,\Omega\setminus\Omega_{11}})\, |||(\boldsymbol{w}_N, r_N)||| \\ &\le C(N^{-k+1} + (\ln N)^{1/2}E(N,\varepsilon)^k)\, |||(\boldsymbol{w}_N, r_N)||| \\ &\le CE(N,\varepsilon)^{k-1}\, |||(\boldsymbol{w}_N, r_N)||| \,.\end{aligned}$$

If $Q_N = Q_{k-1}$ then $(Jp - p, \operatorname{div} \boldsymbol{w}_N) = (I_{k-1}p - p, \operatorname{div} \boldsymbol{w}_N)$ and the proof of Corollary 3.12 can be done with I_{k-1} instead of J. Using $\|I_{k-1}p - p\|_{\Omega_{11}} \le CN^{-k}$ leads to

$$\begin{aligned}|(Jp - p, \operatorname{div} \boldsymbol{w}_N)| &= |(I_{k-1}p - p, \operatorname{div} \boldsymbol{w}_N)| \\ &\le C(N^{-k+1} + (\ln N)^{1/2}E(N,\varepsilon)^k)\, |||(\boldsymbol{w}_N, r_N)||| \\ &\le CE(N,\varepsilon)^{k-1}\, |||(\boldsymbol{w}_N, r_N)|||.\end{aligned}$$

5. Numerical results

In the numerical section, we show the results of computations supporting our theoretical findings. We start with a problem where the exact solution is known and end with some considerations on problems with unknown exact solutions where our theory does not hold to its full extent.

5.1 *Problem with known solution*

We consider the singularly perturbed Oseen equations

$$\begin{aligned} -\varepsilon\Delta\boldsymbol{u} - \left(\begin{pmatrix}2+3x\\3+2y^2\end{pmatrix}\cdot\nabla\right)\boldsymbol{u} + (1+2xy)\boldsymbol{u} + \nabla p = \boldsymbol{f} \text{ in } \Omega = (0,1)^2, \\ \operatorname{div}\boldsymbol{u} = 0 \text{ in } \Omega, \\ \boldsymbol{u} = 0 \text{ on } \partial\Omega, \end{aligned} \tag{5.1}$$

where the right-hand side $\boldsymbol{f}$ of (5.1) was chosen such that

$$u_1(x,y) = \frac{\partial \Psi(x,y)}{\partial y}, \quad u_2(x,y) = -\frac{\partial \Psi(x,y)}{\partial x}, \quad p(x,y) = 2\cos(x)\sin(y) - 2\sin(1)(1-\cos(1))$$

with $\Psi(x,y) := F(x)G(y)$ and

$$\begin{aligned} F(x) &:= c_3(1-x)^3 + c_2(1-x)^2 + c_1(1-x) - \varepsilon\left(e^{-2x/\varepsilon} - e^{-2/\varepsilon}\right) - \sin(1-x),\\ G(y) &:= d_3(1-y)^3 + d_2(1-y)^2 + d_1(1-y) - 1 + \varepsilon\left(e^{-3y/\varepsilon} - e^{-3/\varepsilon}\right) + \cos(1-y) \end{aligned}$$

and constants

$$\begin{aligned} &c_1 := \left(1 + 2\,e^{-2/\varepsilon}\right), \quad c_2 := -\left((4+3\,\varepsilon)\,e^{-2/\varepsilon} + 4 + \cos(1) - 3\,\sin(1) - 3\,\varepsilon\right),\\ &c_3 := \left((2+2\,\varepsilon)\,e^{-2/\varepsilon} + 3 + \cos(1) - 2\,\varepsilon - 2\,\sin(1)\right),\\ &d_1 := -3\,e^{-3/\varepsilon}, \quad d_2 := \left((6+3\,\varepsilon)\,e^{-3/\varepsilon} + 6 - 3\,\cos(1) - \sin(1) - 3\,\varepsilon\right),\\ &d_3 := -\left((3+2\,\varepsilon)\,e^{-3/\varepsilon} + 5 - 2\,\cos(1) - \sin(1) - 2\,\varepsilon\right) \end{aligned}$$

is the solution of (5.1). The function u_1 shows a strong exponential boundary layer at $y=0$ and a weak exponential boundary layer at $x=0$, and vice versa for u_2. The pressure shows no layer behaviour. Moreover, Assumption 2.1 is satisfied. For a visualization of the solution, see Fig. 1.

All calculations were done in MATLAB using the finite element suite $\mathbb{SOFE}$ developed by Lars Ludwig; see Ludwig (2016). In the following, ‘order’ will always denote the exponent r in a convergence order of form $\mathcal{O}(N^{-r})$ while ‘ln-order’ corresponds to the exponent r in a convergence order of form $\mathcal{O}((N^{-1}\ln N)^r)$. If not stated otherwise then the results on the two finest meshes were used to calculate the convergence order.

We have chosen $\gamma = 1$ on the whole domain, unless otherwise stated. The parameter σ was chosen to be $\sigma = k+2$, where k is the order of the elements for the velocity $\boldsymbol{u}$. The errors will be measured in the norm

$$|||(\boldsymbol{u},p)|||_\varepsilon^2 := \varepsilon\|\nabla\boldsymbol{u}\|^2 + \|\boldsymbol{u}\|^2 + \|p\|^2 + \|\gamma\,\operatorname{div}\boldsymbol{u}\|^2; \tag{5.2}$$

this means that α is set to 1.

5.1.1 *Dependence of β_0 and B_0 on N and ε.* First we will consider the dependence of $\beta_0 = \beta_0(N,\varepsilon)$ and $B_0 = B_0(N,\varepsilon,Jp-p_N)$ on N and ε. The pictures in Fig. 3 show the dependence of β_0 on N and ε for the discretizations $Q_2\times Q_1$ and $Q_3\times Q_2$ on Shishkin meshes. Clearly, there is an ε-uniform bound for any fixed N while the value of β_0 will decrease with increasing N if ε is fixed. If N and ε are fixed, the constant β_0 decreases with increasing approximation order k. This effect appears also on isotropic meshes and spectral methods; see Bernardi & Maday (1997). The behaviour of β_0 on Bakhvalov–Shishkin meshes

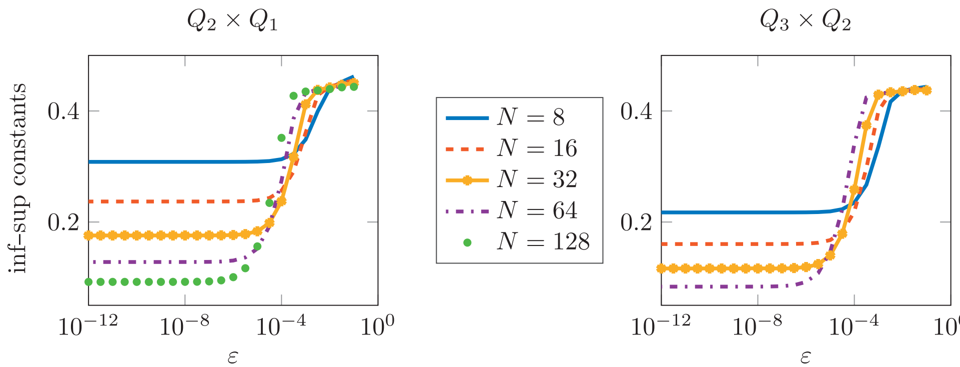

FIG. 3. Problem (5.1): The inf–sup constants for different values of N as a function of ε on Shishkin meshes: $Q_2 \times Q_1$ (left) and $Q_3 \times Q_2$ (right).

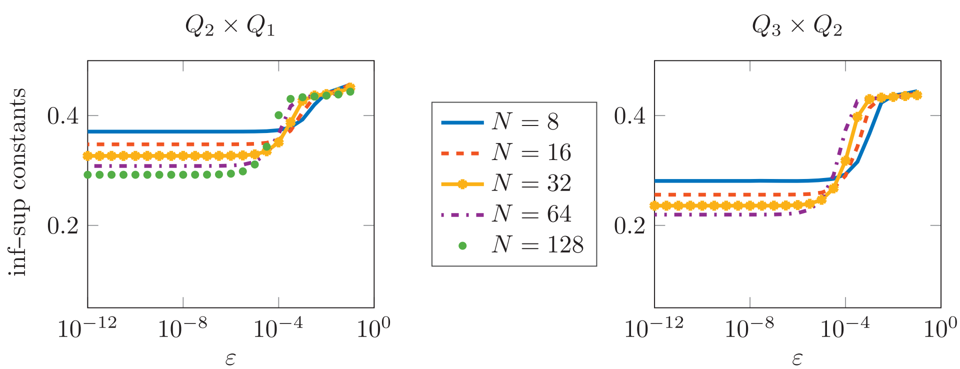

FIG. 4. Problem (5.1): The inf–sup constants for different values of N as a function of ε on Bakhvalov–Shishkin meshes: $Q_2 \times Q_1$ (left) and $Q_3 \times Q_2$ (right).

is presented in Fig. 4. We observe that β_0 takes much larger values and the dependence on N is much weaker compared to Shishkin meshes. Numerical experiments for discretizations $Q_k \times P_{k-1}^{\text{disc}}$ show that β_0 approaches zero if ε tends to zero.

Figure 5 shows the dependence of β_0 and B_0 on N for $\varepsilon = 10^{-8}$. We observe that the values of B_0 are much larger than the corresponding values of β_0. Moreover, the value of B_0 is bounded away from zero for both continuous and discontinuous pressure spaces and is nondecreasing in N. Additional studies showed that there is also an ε-uniform lower bound for B_0. Hence, Theorem 3.11 provides an ε-uniform error bound.

5.1.2 *Uniformity in ε.* For this study, we have chosen $\gamma = 1$ on the whole domain in the computations. Table 2 presents the numerical results on Shishkin meshes (S-mesh) and Bakhvalov–Shishkin meshes (B–S-mesh) with $N = 32$ and ε varying from 10^{-1} to 10^{-10}. We show second-order approximations by $Q_2 \times Q_1$ and $Q_2 \times P_1^{\text{disc}}$ as well as third-order approximations by $Q_3 \times Q_2$ and $Q_3 \times P_2^{\text{disc}}$. Note that for large values of ε the mesh is equidistant while the characteristic meshes are obtained for smaller values of ε. We clearly see from Table 2 that the error for small ε is independent of ε, as predicted by our theory

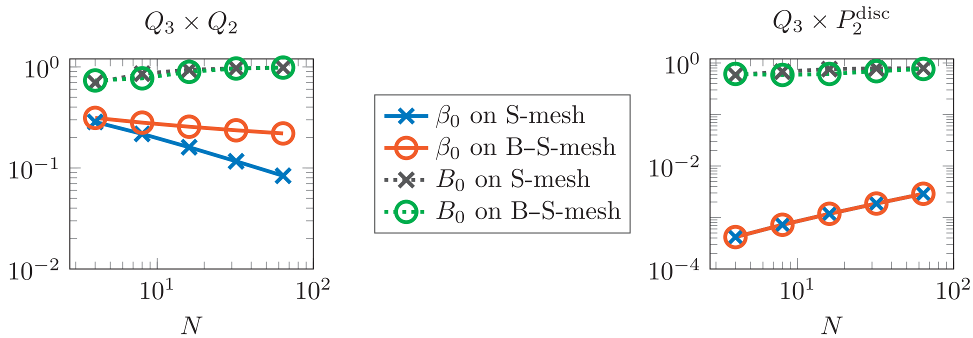

FIG. 5. Problem (5.1): The behaviour of β_0 and B_0 with $\varepsilon = 10^{-8}$ as a function of N for $Q_3 \times Q_2$ (left) and $Q_3 \times P_2^{\mathrm{disc}}$ (right), respectively.

TABLE 2 *Problem* (5.1)*: Error in energy norm* $|||(\cdot,\cdot)|||_\varepsilon$ *for* $N = 32$

	$Q_2 \times Q_1$		$Q_2 \times P_1^{\mathrm{disc}}$		$Q_3 \times Q_2$		$Q_3 \times P_2^{\mathrm{disc}}$	
ε	S-mesh	B–S-mesh	S-mesh	B–S-mesh	S-mesh	B–S-mesh	S-mesh	B–S-mesh
10^{-1}	1.7439-02	8.6659-03	1.7439-02	8.6661-03	1.4779-03	1.4779-03	1.4779-03	1.4779-03
10^{-2}	2.4899-02	3.5717-03	2.4901-02	3.5810-03	3.4299-03	1.2697-04	3.4299-03	1.2702-04
10^{-3}	2.5197-02	3.7450-03	2.5198-02	3.7565-03	3.4685-03	1.2840-04	3.4685-03	1.2850-04
10^{-4}	2.5230-02	3.8047-03	2.5233-02	3.8166-03	3.4723-03	1.2854-04	3.4723-03	1.2865-04
10^{-5}	2.5234-02	3.8146-03	2.5237-02	3.8273-03	3.4727-03	1.2856-04	3.4727-03	1.2866-04
10^{-6}	2.5235-02	3.8156-03	2.5237-02	3.8292-03	3.4727-03	1.2856-04	3.4727-03	1.2866-04
10^{-7}	2.5235-02	3.8157-03	2.5237-02	3.8282-03	3.4727-03	1.2856-04	3.4727-03	1.2866-04
10^{-8}	2.5235-02	3.8157-03	2.5237-02	3.8272-03	3.4727-03	1.2856-04	3.4727-03	1.2866-04
10^{-9}	2.5235-02	3.8157-03	2.5237-02	3.8271-03	3.4727-03	1.2856-04	3.4727-03	1.2866-04
10^{-10}	2.5235-02	3.8157-03	2.5237-02	3.8271-03	3.4727-03	1.2856-04	3.4727-03	1.2866-04

and the ε-uniform bound for B_0. Furthermore, the errors for discretizations of the same order and on the same mesh family differ only slightly. Moreover, the errors on Bakhvalov–Shishkin meshes are smaller than those on Shishkin meshes.

5.1.3 *Convergence studies*. Table 3 shows for Shishkin meshes and $\varepsilon = 10^{-8}$, which is sufficiently small to have a solution with layers, the errors in the energy norm for pairs of second-, third-, and fourth-order with continuous and discontinuous pressure approximations. It can be seen that the theoretically predicted convergence orders are achieved. The differences in the energy norm between the two different pressure approximation spaces are again very small. Here the different pressure errors are similar and in both cases dominated by the velocity errors.

The results on Bakhvalov–Shishkin meshes for $\varepsilon = 10^{-8}$ are shown in Table 4. The same approximation spaces as on Shishkin meshes have been used. The convergence orders predicted by our theory are obtained. Moreover, one sees that the errors on Bakhvalov–Shishkin meshes are smaller than those on the corresponding Shishkin meshes. This is due to the logarithmic factor which is present only for Shishkin meshes. The difference between Shishkin and Bakhvalov–Shishkin meshes becomes larger with

TABLE 3 *Problem (5.1): Error in energy norm* $|||(\cdot,\cdot)|||_\varepsilon$ *on Shishkin meshes with* $\varepsilon = 10^{-8}$

N	$Q_2 \times Q_1$	$Q_2 \times P_1^{\text{disc}}$	$Q_3 \times Q_2$	$Q_3 \times P_2^{\text{disc}}$	$Q_4 \times Q_3$	$Q_4 \times P_3^{\text{disc}}$
4	2.7153-01	2.5880-01	7.2169-02	7.2220-02	2.5802-02	2.5803-02
8	1.3134-01	1.3131-01	3.6858-02	3.6859-02	1.0660-02	1.0660-02
16	6.1849-02	6.1859-02	1.2918-02	1.2918-02	2.6828-03	2.6828-03
32	2.5235-02	2.5237-02	3.4727-03	3.4727-03	4.7297-04	4.7297-04
64	9.2647-03	9.2650-03	7.8324-04	7.8324-04	6.5527-05	6.5527-05
ln-order	1.96	1.96	2.92	2.92	3.87	3.87
Theory	2	2	3	3	4	4

TABLE 4 *Problem (5.1): Error in energy norm* $|||(\cdot,\cdot)|||_\varepsilon$ *on Bakhvalov–Shishkin meshes with* $\varepsilon = 10^{-8}$

N	$Q_2 \times Q_1$	$Q_2 \times P_1^{\text{disc}}$	$Q_3 \times Q_2$	$Q_3 \times P_2^{\text{disc}}$	$Q_4 \times Q_3$	$Q_4 \times P_3^{\text{disc}}$
4	2.2054-01	2.0379-01	3.1063-02	3.1183-02	8.3515-03	8.3526-03
8	5.6584-02	5.6509-02	6.1054-03	6.1147-03	9.5490-04	9.5494-04
16	1.4800-02	1.4845-02	9.3378-04	9.3470-04	7.8247-05	7.8249-05
32	3.8157-03	3.8272-03	1.2856-04	1.2866-04	5.5671-06	5.5672-06
64	9.7042-04	9.7364-04	1.6851-05	1.6864-05	3.7078-07	3.7079-07
Order	1.98	1.97	2.93	2.93	3.91	3.91
Theory	2	2	3	3	4	4

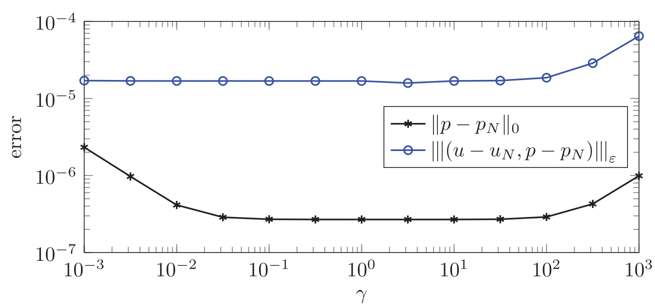

FIG. 6. Problem (5.1): Dependency of errors on the stabilization parameter γ.

increasing approximation order and already reaches two orders of magnitude for the fourth-order pairs $Q_4 \times Q_3$ and $Q_4 \times P_3^{\text{disc}}$.

5.1.4 *Necessity of stabilization.* The size of the stabilization parameter γ is arbitrary and our analysis gives only the restriction that it should be constant with respect to N and ε. Figure 6 shows for the energy norm of the solution and the L^2-norm of the pressure alone the dependency on γ for a Bakhvalov–Shishkin

TABLE 5 *Problem (5.1): Error in energy norm* $|||(\cdot,\cdot)|||_\varepsilon$ *for* $Q_3 \times Q_2$ *and several choices of* γ *on Bakhvalov–Shishkin meshes with* $\varepsilon = 10^{-8}$

	$\gamma = 1$ in Ω		$\gamma = 1$ in Ω_{11}		$\gamma = 0$ in Ω	
N	Error	Order	Error	Order	Error	Order
4	3.1063-02		3.1075-02		3.1302-02	
8	6.1054-03	2.35	6.1061-02	2.35	6.1727-03	2.34
16	9.3378-04	2.71	9.3381-04	2.71	9.6048-04	2.68
32	1.2856-04	2.86	1.2856-04	2.86	1.4026-04	2.78
64	1.6851-05	2.93	1.6851-05	2.93	2.1901-05	2.68
128	2.1567-06	2.97	2.1567-06	2.97	4.1048-06	2.42
196	6.0560-07	2.98	6.0560-07	2.98	1.6071-06	2.20

mesh with fixed $N = 64$, $\varepsilon = 10^{-8}$, $Q_3 \times Q_2$ and stabilization in the full domain. It can clearly be seen that a moderate value of γ gives best results. Furthermore, we observe almost constant error values on a wide range of values for γ that include $\gamma = 1$. This is the reason for choosing this particular value in the numerical studies.

Table 5 shows the influence of the grad-div stabilization for $\varepsilon = 10^{-8}$ and $Q_3 \times Q_2$. We consider three choices for γ: first $\gamma = 1$ in Ω, second $\gamma = 1$ in Ω_{11} only, and third $\gamma = 0$ in Ω. The last choice gives the standard Galerkin method, thus no stabilization. The errors are measured in the corresponding energy norm $|||(\cdot,\cdot)|||_\varepsilon$. We observe that the errors for discrete problems without grad-div stabilization are larger than those for both stabilized discrete problems. Moreover, the convergence order for problems without stabilization is less than for problems with stabilization. This is in agreement with Corollary 3.12. Furthermore, we observe that the errors of both stabilized methods are almost indistinguishable, confirming Corollary 3.13.

5.2 *Problems with unknown solution*

In this part of our numerical study, we do not prescribe a solution but give only the data for the problem. Therefore it is not clear whether Assumption 2.1 holds. In order to calculate the errors and the value of B_0, we replace the exact solution with a reference solution computed on a Bakhvalov–Shishkin mesh with $N = 128$ and $Q_4 \times Q_3$. This approach is a variant of the double-mesh principle; see Farrell *et al.* (2000).

Let us start with

$$\begin{aligned} -\varepsilon\Delta \boldsymbol{u} - \left(\begin{pmatrix}1\\1\end{pmatrix}\cdot\nabla\right)\boldsymbol{u} + \boldsymbol{u} + \nabla p &= \boldsymbol{f} \text{ in } \Omega = (0,1)^2, \\ \operatorname{div} \boldsymbol{u} &= 0 \text{ in } \Omega, \\ \boldsymbol{u} &= 0 \text{ on } \partial\Omega, \end{aligned} \tag{5.3}$$

where the right-hand side $\boldsymbol{f} = (1, x)^{\mathrm{T}}$ was chosen to be simple but not a gradient. First we look at the convergence of our method with full stabilization for $\varepsilon = 10^{-8}$ on a Bakhvalov–Shishkin mesh. We present the results for $Q_2 \times Q_1$ and $Q_3 \times Q_2$ only as the counterparts with discontinuous pressure

TABLE 6 *Problem (5.3): Error in energy norm* $|||(\cdot,\cdot)|||_\varepsilon$ *on Bakhvalov–Shishkin meshes with* $\varepsilon = 10^{-8}$

	$\boldsymbol{f} = (1, x)^{\mathrm{T}}$	
N	$Q_2 \times Q_1$	$Q_3 \times Q_2$
8	3.178-02	2.244-02
16	1.787-02	1.258-02
32	1.121-02	8.035-03
64	6.780-03	4.891-03
Order	0.73	0.72

	$\boldsymbol{f} = (1-x-y)(x-y)(1,-1)^{\mathrm{T}}$	
N	$Q_2 \times Q_1$	$Q_3 \times Q_2$
8	9.403-03	1.300-03
16	2.488-03	2.495-04
32	6.322-04	4.794-05
64	1.571-04	9.191-06
Order	2.00	2.38

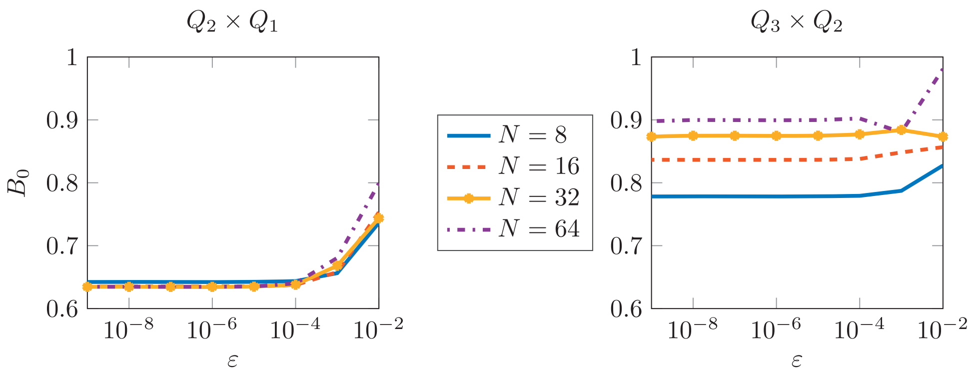

FIG. 7. Problem (5.3): The value of B_0 for different values of N as a function of ε on Bakhvalov–Shishkin meshes: $Q_2 \times Q_1$ (left) and $Q_3 \times Q_2$ (right).

approximations behave similarly. The left part of Table 6 shows the error behaviour measured in the energy norm. Clearly the convergence order is far from the theoretically predicted values and does not improve upon increasing the polynomial order. A possible explanation is that Assumption 2.1 is not fulfilled and some corner singularities arise. This is similar to convection–diffusion problems where additional compatibility conditions of the right-hand side have to be fulfilled; see, e.g., Linß & Stynes (2001), Kellogg & Stynes (2005, 2007). Thus, we modify our problem by changing the right-hand side to

$$\boldsymbol{f}_2 := \begin{pmatrix} 1 \\ -1 \end{pmatrix} (1-x-y)(x-y).$$

This function $\boldsymbol{f}_2$ is zero in the corners of Ω which is a minimal compatibility condition in convection–diffusion problems. The convergence results are presented in the right part of Table 6. We do observe the optimal convergence order for $Q_2 \times Q_1$ but still a reduced order for the higher-order pair. Thus again, not enough compatibility conditions are fulfilled to ensure that Assumption 2.1 holds for the higher derivatives. It remains an open question which conditions are needed.

Using $\boldsymbol{f}_2$ we can also look into the constant B_0. The obtained results are shown in Fig. 7. Again we observe for all values of N and ε that B_0 is bounded away from zero. In order to get more insight into the

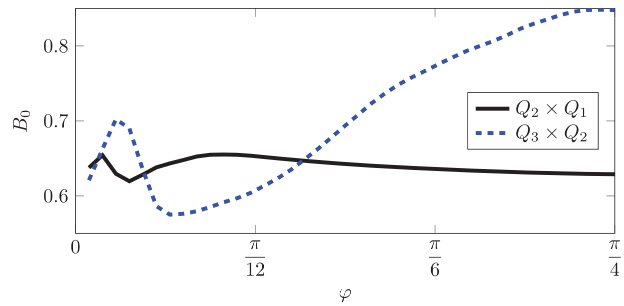

FIG. 8. Problem (5.3): The value of B_0 as function of φ on Bakhvalov–Shishkin meshes.

behaviour of B_0 , we now change the vector $\boldsymbol{b}$ in (5.3) to

$$\boldsymbol{b} = \begin{pmatrix} \cos(\varphi) \\ \sin(\varphi) \end{pmatrix}$$

for $\varphi \in (0, \pi/4]$. For fixed values of $N = 32$ and $\varepsilon = 10^{-8}$, we vary in Fig. 8 the angle φ. Note that $\varphi = 0$ would give characteristic layers that we will investigate in a moment. Although the value of B_0 depends on φ, it is bounded away from zero over the full range of φ.

Let us now consider the case of characteristic layers, e.g.,

$$\begin{aligned} -\varepsilon \Delta \boldsymbol{u} - \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot \nabla \right) \boldsymbol{u} + \boldsymbol{u} + \nabla p = \begin{pmatrix} 1 \\ -1 \end{pmatrix} (1-x-y)(x-y) & \text{ in } \Omega = (0,1)^2, \\ \operatorname{div} \boldsymbol{u} = 0 \qquad & \text{ in } \Omega, \\ \boldsymbol{u} = 0 \qquad & \text{ on } \partial\Omega. \end{aligned} \tag{5.4}$$

Here our theory about the decomposition of $\boldsymbol{u}$ does not hold at all. Nevertheless, we can look at the corresponding convection–diffusion case where a decomposition is known; see Kellogg & Stynes (2005, 2007), and use the analogy of the layer structure. Thus we expect u_1 to have characteristic layers at $y = 0$ and $y = 1$ and u_2 to have an exponential layer at $x = 0$. An approximation to the solution is shown in Fig. 9.

With this assumption on the layer structure, a suitable layer-adapted mesh can be constructed in the usual way and the analysis presented could also be done. We present only the convergence results in Table 7 that are similar to those of only exponential layers with reduced regularity. Also the constant B_0 is bounded away from zero for all ε and N.

6. Outlook

This article is a starting point for the analysis of more complex domains and problems. The analysis relies on a solution decomposition and therefore a precise knowledge of the layer structure of $\boldsymbol{u}$. Having such a structure, it is relatively easy to extend the techniques of this article to the three-dimensional case

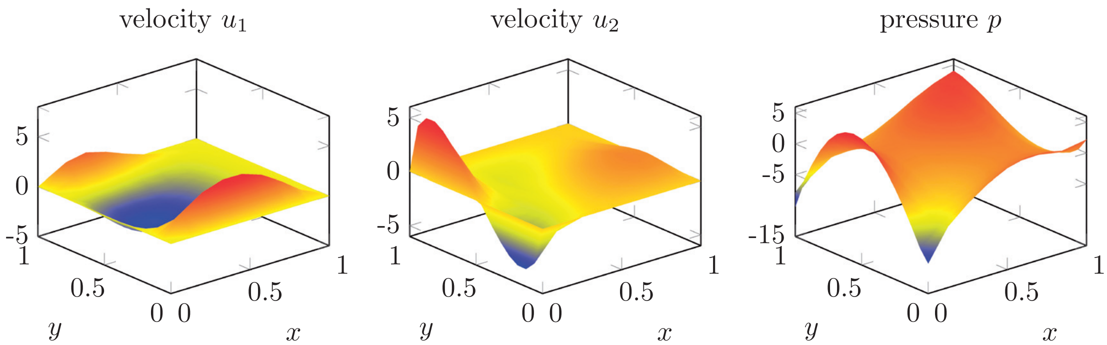

FIG. 9. Problem (5.4): Solution of (5.4) with exponential and characteristic layers in the velocity.

TABLE 7 *Problem (5.4): Error in energy norm* $|||(\cdot,\cdot)|||_\varepsilon$ *on Bakhvalov–Shishkin meshes with* $\varepsilon = 10^{-8}$

N	$Q_2 \times Q_1$	$Q_3 \times Q_2$
8	9.403-03	1.300-03
16	2.488-03	2.495-04
32	6.322-04	4.794-05
64	1.571-04	9.191-06
Order	2.00	2.38

or more general domains. Numerical experiments do actually indicate that the solution structure in three dimensions is similar to the two-dimensional situation.

In the case of second-order reaction–diffusion problems, more general but smooth domains were investigated in, e.g., Xenophontos & Fulton (2003), Kopteva (2009). Their techniques could be incorporated into the system case.

For more general domains and problems, the type of layers can change from exponential to characteristic (see Roos *et al.*, 2008 for these types of layers), but the exact transition from one type to another is not obvious. A technique to handle characteristic layers is the use of problem-specific enriched function spaces where the enrichment could be based on the law of wall; see Krank & Wall (2016) for a recent development.

For flows in channels and around obstacles in unbounded domains, one tries to prevent strong boundary layers at the outflow. The idea is to apply suitable boundary conditions that allow us to restrict the domains without significant changes of the flow solutions. Such methods can be found, e.g., in Griffiths (1997), Braack & Mucha (2014).

Another topic for further research is to show that B_0 is actually bounded away from zero, independent of N and ε. Having this would provide a tool for proving the uniform error estimation.

As indicated in Fig. 8, the value of B_0 stays away from zero even for the transition from exponential to characteristic layers. Assuming a suitable layer structure for the solution of the Oseen problem with characteristic layers, a convergence behaviour similar to the case with only exponential layers is obtained; compare Tables 6 (right) and 7.

Funding

K.H. was funded by the grant FR 3052/2-1 of the German Research Foundation (DFG).

References

Ainsworth, M., Barrenechea, G. R. & Wachtel, A. (2015) Stabilization of high aspect ratio mixed finite elements for incompressible flow. *SIAM J. Numer. Anal.*, **53**, 1107–1120.

Apel, T. (1999) *Anisotropic Finite Elements: Local Estimates and Applications*. Advances in Numerical Mathematics. Stuttgart: B. G. Teubner.

Apel, T., Knopp, T. & Lube, G. (2008) Stabilized finite element methods with anisotropic mesh refinement for the Oseen problem. *Appl. Numer. Math.*, **58**, 1830–1843.

Apel, T. & Maharavo Randrianarivony, H. (2003) Stability of discretizations of the Stokes problem on anisotropic meshes. *Math. Comput. Simul.*, **61**, 437–447.

Apel, T. & Matthies, G. (2008) Nonconforming, anisotropic, rectangular finite elements of arbitrary order for the Stokes problem. *SIAM J. Numer. Anal.*, **46**, 1867–1891.

Bakhvalov, N. S. (1969) The optimization of methods of solving boundary value problems with a boundary layer. *U.S.S.R. Comput. Math. Math. Phys.*, **9**, 139–166.

Bernardi, C. & Maday, Y. (1997) Spectral methods. *Handbook of Numerical Analysis*, vol. V, pp. 209–485. Amsterdam: North-Holland.

Braack, M. & Mucha, P. B. (2014) Directional do-nothing condition for the Navier-Stokes equations. *J. Comput. Math.*, **32**, 507–521.

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

Burman, E. & Linke, A. (2008) Stabilized finite element schemes for incompressible flow using Scott–Vogelius elements. *Appl. Numer. Math.*, **58**, 1704–1719.

Dallmann, H., Arndt, D. & Lube, G. (2016) Local projection stabilization for the Oseen problem. *IMA J. Numer. Anal.*, **36**, 796–823.

Farrell, P. A., Hegarty, A. F., Miller, J. J. H., O'Riordan, E. & Shishkin, G. I. (2000) *Robust Computational Techniques for Boundary Layers*. Applied Mathematics (Boca Raton), vol. 16. Boca Raton, FL: Chapman & Hall/CRC.

Franca, L. P. & Frey, S. L. (1992) Stabilized finite element methods: II. The incompressible Navier–Stokes equations. *Comput. Methods Appl. Mech. Eng.*, **99**, 209–233.

Franz, S. (2008) Continuous interior penalty method on a Shishkin mesh for convection-diffusion problems with characteristic boundary layers. *Comput. Meth. Appl. Mech. Eng.*, **197**, 3679–3686.

Franz, S., Höhne, K. & Waurick, M. (2016) A solution decomposition for a singularly perturbed 4th order problem (submitted).

Franz, S., Linß, T. & Roos, H.-G. (2008) Superconvergence analysis of the SDFEM for elliptic problems with characteristic layers. *Appl. Numer. Math.*, **58**, 1818–1829.

Franz, S. & Matthies, G. (2010) Local projection stabilisation on S-type meshes for convection-diffusion problems with characteristic layers. *Computing*, **87**, 135–167.

Gelhard, T., Lube, G., Olshanskii, M. A. & Starcke, J.-H. (2005) Stabilized finite element schemes with LBB-stable elements for incompressible flows. *J. Comput. Appl. Math.*, **177**, 243–267.

Girault, V. & Raviart, P.-A. (1986) *Finite Element Methods for Navier–Stokes Equations*. Berlin: Springer.

Griffiths, D. F. (1997) The "no boundary condition" outflow boundary condition. *Int. J. Numer. Methods Fluids*, **24**, 393–411.

Hansbo, P. & Szepessy, A. (1990) A velocity-pressure streamline diffusion method for the incompressible Navier–Stokes equations. *Comput. Methods Appl. Mech. Eng.*, **84**, 175–192.

HUGHES, T. J. R., FRANCA, L. P. & BALESTRA, M. (1986) A new finite element formulation for computational fluid dynamics. V: circumventing the Babuska–Brezzi condition: a stable Petrov–Galerkin formulation of the Stokes problem accommodating equal-order interpolations. *Comput. Methods Appl. Mech. Eng.*, **59**, 85–99.

JENKINS, E. W., JOHN, V., LINKE, A. & REBHOLZ, L. G. (2014) On the parameter choice in grad-div stabilization for the Stokes equations. *Adv. Comput. Math.*, **40**, 491–516.

JOHNSON, C. & SARANEN, J. (1986) Streamline diffusion methods for the incompressible Euler and Navier–Stokes equations. *Math. Comput.*, **47**, 1–18.

KELLOGG, R. B. & STYNES, M. (2005) Sharpened and corrected version of: corner singularities and boundary layers in a simple convection-diffusion problem. *J. Differ. Equ.*, **213**, 81–120.

KELLOGG, R. B. & STYNES, M. (2007) Sharpened bounds for corner singularities and boundary layers in a simple convection-diffusion problem. *Appl. Math. Lett.*, **20**, 539–544.

KOPTEVA, N. (2009) Numerical analysis of a 2d singularly perturbed semilinear reaction-diffusion problem. *Numerical Analysis and Its Applications* (S. Margenov, L. G. Vulkov & J. Waśniewski eds). Lecture Notes in Computer Science, vol. 5434. Berlin: Springer, pp. 80–91.

KRANK, B. & WALL, W. A. (2016) A new approach to wall modeling in LES of incompressible flow via function enrichment. *J. Comput. Phys.*, **316**, 94–116.

LINKE, A. (2014) On the role of the Helmholtz decomposition in mixed methods for incompressible flows and a new variational crime. *Comput. Methods Appl. Mech. Eng.*, **268**, 782–800.

LINKE, A. & REBHOLZ, L. G. (2013) On a reduced sparsity stabilization of grad-div type for incompressible flow problems. *Comput. Methods Appl. Mech. Eng.*, **261/262**, 142–153.

LINß, T. (1999) An upwind difference scheme on a novel Shishkin-type mesh for a linear convection-diffusion problem. *J. Comput. Appl. Math.*, **110**, 93–104.

LINß, T. (2000) Analysis of a Galerkin finite element method on a Bakhvalov-Shishkin mesh for a linear convection-diffusion problem. *IMA J. Numer. Anal.*, **20**, 621–632.

LINß, T. (2010) *Layer-Adapted Meshes for Reaction-Convection-Diffusion Problems*. Lecture Notes in Mathematics, vol. 1985. Berlin: Springer.

LINß, T. & STYNES, M. (2001) Asymptotic analysis and Shishkin-type decomposition for an elliptic convection-diffusion problem. *J. Math. Anal. Appl.*, **261**, 604–632.

LUDWIG, L. (2017) https://github.com/SOFE-Developers/SOFE, last accessed 17 March 2017.

MATTHIES, G., LUBE, G. & RÖHE, L. (2009) Some remarks on residual-based stabilisation of inf-sup stable discretisations of the generalised Oseen problem. *Comput. Methods Appl. Math.*, **9**, 368–390.

MATTHIES, G. & TOBISKA, L. (2015) Local projection type stabilization applied to inf–sup stable discretizations of the Oseen problem. *IMA J. Numer. Anal.*, **35**, 239–269.

MILLER, J. J. H., O'RIORDAN, E. & SHISHKIN, G. I. (1996) *Fitted Numerical Methods for Singular Perturbation Problems: Error Estimates in the Maximum Norm for Linear Problems in One and Two Dimensions*. River Edge, NJ: World Scientific Publishing Co.

OLSHANSKII, M. A. (2002) A low order Galerkin finite element method for the Navier-Stokes equations of steady incompressible flow: a stabilization issue and iterative methods. *Comput. Methods Appl. Mech. Eng.*, **191**, 5515–5536.

OLSHANSKII, M. A., LUBE, G., HEISTER, T. & LÖWE, J. (2009) Grad-div stabilization and subgrid pressure models for the incompressible Navier-Stokes equations. *Comput. Methods Appl. Mech. Eng.*, **198**, 3975–3988.

OLSHANSKII, M. A. & REUSKEN, A. (2004) Grad-div stabilization for the Stokes equations. *Math. Comput.*, **73**, 1699–1718.

ROOS, H.-G. (2012) Robust numerical methods for singularly perturbed differential equations: a survey covering 2008–2012. *ISRN Appl. Math.*, **2012**, Article ID 379547.

ROOS, H.-G. & LINß, T. (1999) Sufficient conditions for uniform convergence on layer-adapted grids. *Computing*, **63**, 27–45.

ROOS, H.-G., STYNES, M. & TOBISKA, L. (2008) *Robust Numerical Methods for Singularly Perturbed Differential Equations*. Springer Series in Computational Mathematics, vol. 24, 2nd edn. Berlin: Springer.

Schötzau, D. & Schwab, C. (1998) Mixed *hp*-FEM on anisotropic meshes. *Math. Models Methods Appl. Sci.*, **8**, 787–820.

Schötzau, D., Schwab, C. & Stenberg, R. (1999) Mixed *hp*-FEM on anisotropic meshes. II. Hanging nodes and tensor products of boundary layer meshes. *Numer. Math.*, **83**, 667–697.

Stynes, M. & O'Riordan, E. (1997) A uniformly convergent Galerkin method on a Shishkin mesh for a convection-diffusion problem. *J. Math. Anal. Appl.*, **214**, 36–54.

Stynes, M. & Tobiska, L. (2003) The SDFEM for a convection-diffusion problem with a boundary layer: optimal error analysis and enhancement of accuracy. *SIAM J. Numer. Anal.*, **41**, 1620–1642.

Stynes, M. & Tobiska, L. (2008) Using rectangular Q_p elements in the SDFEM for a convection-diffusion problem with a boundary layer. *Appl. Numer. Math.*, **58**, 1709–1802.

Tezduyar, T. E., Mittal, S., Ray, S. E. & Shih, R. (1992) Incompressible flow computations with stabilized bilinear and linear equal order interpolation velocity pressure elements. *Comput. Methods Appl. Mech. Eng.*, **95**, 221–242.

Tobiska, L. & Lube, G. (1991) A modified streamline diffusion method for solving the stationary Navier–Stokes equations. *Numer. Math.*, **59**, 13–29.

Xenophontos, C. & Fulton, S. R. (2003) Uniform approximation of singularly perturbed reaction-diffusion problems by the finite element method on a Shishkin mesh. *Numer. Methods Partial Differ. Equ.*, **19**, 89–111.