# Weierstraß-Institut
## für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint ISSN 0946 – 8633


# Error analysis of the SUPG finite element discretization of evolutionary convection-diffusion-reaction equations


Volker John[1], Julia Novo[2]



submitted: March 23, 2010



[1] Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstr. 39, 10117 Berlin, Germany and Department of Mathematics and Computer Science, Arnimallee 6, 14195 Berlin, Germany
email: Volker.John@wias-berlin.de

[2] Departamento de Matemáticas, Universidad Autónoma de Madrid, Instituto de Ciencias Matemáticas CSIC-UAM-UC3M-UCM, Spain
email: julia.novo@uam.es



No. 1494
Berlin 2010

2010 *Mathematics Subject Classification.* 65M12, 65M60.

*Key words and phrases.* Evolutionary convection-diffusion-reaction equation, Streamline-Upwind Petrov-Galerkin (SUPG) finite element method, backward Euler scheme, error analysis, time-continuous problem .

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
10117 Berlin
Germany

Fax: + 49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

**Abstract**

Conditions on the stabilization parameters are explored for different approaches in deriving error estimates for the SUPG finite element stabilization of time-dependent convection-diffusion-reaction equations that is combined with the backward Euler method. Standard energy arguments lead to estimates for stabilization parameters that depend on the length of the time step. The stabilization vanishes in the time-continuous limit. However, based on numerical experiences, this seems not to be the correct behavior. For this reason, the time-continuous case is analyzed under certain conditions on the coefficients of the equation and the finite element method. An error estimate with the standard order of convergence is derived for stabilization parameters of the same form that is optimal for the steady-state problem. Numerical studies support the analytical results.

# 1 Introduction

Evolutionary convection-diffusion-reaction equations model the transport and reaction of species. In applications, typically the size of the diffusion is much smaller than the size of the convective term and solutions develop sharp layers. In this case, it is well known that standard finite element methods perform poorly and exhibit non-physical oscillations. Stabilization techniques are required in order to get physically sound numerical approximations. This paper studies one of the currently most popular finite element stabilizations, the Streamline-Upwind Petrov–Galerkin (SUPG) method that was introduced for steady-state equations in [8, 2]. Meanwhile, some results on the numerical analysis of the SUPG method for time-dependent convection-diffusion-reaction equations and a number of numerical studies can be found in the literature.

Concerning the numerical analysis, the case of the transient convection equation without diffusive and reactive term is considered in [3]. It is shown that a finite element discretization in space coupled with the backward Euler, the Crank–Nicolson or the second order backward differentiation formula in time leads to the classical error bound for the SUPG method in the $L^2$ norm (suboptimal by a factor of one half) and also to an optimal error bound in the norm of the material derivative. The results are obtained under certain regularity conditions on the data and with stability parameters that depend only on the mesh size in the space variable. However, an optimal bound for the error in the streamline derivative is not proven. If the data are not sufficiently smooth or if the velocity field is non-solenoidal, then the bound for the backward Euler method is valid under the condition $\delta^2 = \mathcal{O}(k)$ and the bound for the Crank–Nicolson scheme is valid under the condition $\delta = \mathcal{O}(k)$, where $\delta$ is the SUPG stabilization parameter and $k$ the length of the time step. An analogous condition for

$\delta$ was found in [13] where a Galerkin least-squares method in space coupled with a $\theta$-scheme in time is analyzed. The analysis of [13] excludes the case $\theta = 1/2$. Finally, the stability of the SUPG finite element method for transient convection-diffusion equations is studied in [1]. However, as it is shown in [3], the coercivity result of [1] leads to suboptimal global estimates in time.

Numerical studies of the SUPG method, together with a discussion on similarities and differences to other stabilized finite element methods, can be found in [6]. In [10, 11], the SUPG method was compared in comprehensive studies with other stabilized finite element methods. The approach in these studies was as follows: 1) discretize the equation in time, 2) consider the equation in each discrete time as a steady-state convection-diffusion-reaction equation, 3) discretize this equation in space with a stabilized method and apply a parameter choice that is appropriate for this type of steady-state equation. This methodology leads to parameters that are (in the notation of formula (3.1) below) proportional to the length of the time step, see formulae (8) and (11) in [10]. The numerical results with this approach show large spurious oscillations compared with other methods. Such oscillations can be observed also if the SUPG method, derived in this way, is used in coupled systems coming from applications, as in [9].

Altogether, the numerical results obtained so far are not at all satisfactory. We think that the reason for this is the choice of the stabilization parameters that depend on the length of the time step. This opinion is also stated in [7]. There, another approach for deriving the fully discrete equation is considered: 1) discretize the equation in space with a stabilized method, 2) choose standard stabilization parameters for this equation, 3) discretize the equation in time. Because the temporal discretization is performed after the choice of the stabilization parameters, these parameters cannot depend on the time step. Numerical studies in [7] show that this approach leads to much more stable results for small time steps compared with the approach from [10, 11]. In addition, another parameter choice is proposed in [7] that, e.g., does not depend on the length of the time step if a steady-state solution is approached, the so-called element-vector-based parameter choice.

The goal of the present paper consists in exploring the conditions on the stabilization parameters for different approaches in the numerical analysis for deriving error estimates. In particular, error estimates that do not lead to a dependency of the stabilization parameter on the length of the time step are of interest. To our best knowledge, error estimates of this kind for the SUPG method applied to evolutionary convection-diffusion-reaction equations are not yet available. The main difficulty in the analysis of the method comes from the fact that the time derivative has to enter the stabilization term in order to ensure consistency and this adds a non-symmetric term that cannot be easily bounded using standard energy arguments.

This paper concentrates on the backward Euler scheme as temporal discretization. In Sections 3 and 4, stability bounds and error estimates are derived based on standard energy arguments. Two different ways to argue lead to error estimates under the conditions $\delta = \mathcal{O}(k)$ and $\delta = \mathcal{O}(k^{1/2}h)$, respectively. These conditions arise in the stability bounds from the stabilization term with the discretization of the time derivative. In both choices, the stabilization parameters tend to zero on a fixed spatial grid as the length of the time step approaches zero. As discussed above, this seems not to be the correct choice. This is also seen in numerical studies, e.g., in Example 6.2 below. Altogether, the limit of the time-continuous case could not be treated so far satisfactorily by standard energy arguments.

To obtain some insight in the time-continuous case, Section 5 studies a special problem, where the convection field and the reaction do not depend on time, and the convection field is solenoidal. The SUPG method is applied to $P_1$ finite elements on a uniform grid with width $h$. The stabilization parameters are chosen to be the same on all mesh cells, depending only on the coefficients of the equation and on $h$: $\delta = \mathcal{O}(h)$. Under certain regularity assumptions on the solution and extending the analysis of [3], an error estimate for the $L^2$ norm and the norm of the material derivative is derived with the standard order of convergence 3/2. In the next step, based on this result, an estimate for the error in the norm of the streamline derivative is proven with the same order of convergence. To our best knowledge, this is the first result that proves standard order of convergence for the SUPG method applied to evolutionary convection-diffusion-reaction equations with a parameter choice that is essentially the same as in the steady-state case.

The next part of the paper, Section 6, presents some numerical studies. First, an example with a smooth, given solution is considered. The simulations performed at this example support the error estimates from the previous sections. Second, a rotating body problem is studied for the $P_1$ finite element, on a given spatial grid, and for a very small length of the time step. The results show clearly that in this situation a choice of the stabilization parameter independently of the length of the time step has to be preferred.

The paper concludes in Section 7 with a summary of the results and an outlook to open questions.

**2. The SUPG Method and Preliminaries of the Analysis.** Throughout this paper, standard notations are used for Lebesgue and Sobolev spaces. Generic constant that do not depend on the mesh width or the length of the time step are denoted by $C$.

A linear time-dependent convection-diffusion-reaction equation is given by

$$\begin{aligned} u_t - \varepsilon\Delta u + \mathbf{b}\cdot\nabla u + cu &= f && \text{in } (0,T]\times\Omega, \\ u &= 0 && \text{on } [0,T]\times\partial\Omega, \\ u(0,\mathbf{x}) &= u_0(\mathbf{x}) && \text{in } \Omega, \end{aligned} \tag{2.1}$$

where $\Omega$ is a bounded open domain in $\mathbb{R}^d$, $d \in \{1,2,3\}$, with boundary $\partial\Omega$, $\mathbf{b}(t,\mathbf{x})$ and $c(t,\mathbf{x})$ are given functions, $\varepsilon > 0$ is a constant diffusion coefficient, $u_0(\mathbf{x})$ are given initial data and $T$ is a given final time. For simplicity, the case that $\Omega$ is a convex polygonal or polyhedral domain is considered. In the following, it is assumed that there is a constant $\mu_0 > 0$ such that

$$0 < \mu_0 \le \mu(\mathbf{x}) = \left(c - \frac{1}{2}\nabla\cdot\mathbf{b}\right)(\mathbf{x}), \quad \forall\, \mathbf{x}\in\Omega. \tag{2.2}$$

This is a standard assumption in the analysis of equations of type (2.1), [14].

Let $V = H_0^1(\Omega)$. A variational form of (2.1) reads as follows: Find $u : (0,T] \to V$ such that

$$(u_t, v) + (\varepsilon\nabla u, \nabla v) + (\mathbf{b}\cdot\nabla u + cu, v) = (f, v) \quad \forall\, v \in V, \tag{2.3}$$

and $u(0,\mathbf{x}) = u_0(\mathbf{x})$. Here, $(\cdot,\cdot)$ denotes the inner product in $L^2(\Omega)^d$, $d \in \{1,2,3\}$. In numerical simulations, $V$ is replaced by a finite dimensional (finite element) space $V_{h,r}$, where $h$ indicates the fineness of the underlying triangulation $\mathcal{T}_h$ and $r \in \mathbb{N}$ the degree of the local finite element polynomials. This paper considers the case of a

conforming finite element method, i.e. $V_{h,r} \subset V$. The time-continuous finite element problem aims to find a function $u_h \in V_{h,r}$ that fulfills a problem of form (2.3) for all test functions from $V_{h,r}$ with an appropriate approximation of $u_0(\mathbf{x})$ at the initial time.

Using now some temporal discretization, one obtains a finite element Galerkin method for solving (2.3). It is well known that in the case of small diffusion, in particular compared with the convection, the Galerkin method is instable and leads to solutions that are globally polluted with huge spurious oscillations. A stabilization of the Galerkin method becomes necessary. The probably most popular stabilized finite element method is the SUPG method. This residual-based method adds artificial diffusion along the streamlines of the solution. It has the form (time-continuous case): Find $u_h \ : \ (0,T] \to V_{h,r}$ such that

$$\begin{aligned}(u_{h,t}, v_h) + a_{\mathrm{SUPG}}(u_h, v_h) + \sum_{K\in\mathcal{T}_h} \delta_K (u_{h,t}, \mathbf{b}\cdot\nabla v_h)_K \\ = (f, v_h) + \sum_{K\in\mathcal{T}_h} \delta_K (f, \mathbf{b}\cdot\nabla v_h)_K \quad \forall\ v_h \in V_{h,r},\end{aligned}$$

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

$$\begin{aligned}a_{\mathrm{SUPG}}(u_h, v_h) = \varepsilon(\nabla u_h, \nabla v_h) + (\mathbf{b}\cdot\nabla u_h, v_h) + (c u_h, v_h) \\ + \sum_{K\in\mathcal{T}_h} \delta_K (-\varepsilon\Delta u_h + \mathbf{b}\cdot\nabla u_h + c u_h, \mathbf{b}\cdot\nabla v_h)_K. \qquad (2.4)\end{aligned}$$

Here, $K \in \mathcal{T}_h$ denotes the mesh cells of the triangulation, $(\cdot,\cdot)_K$ the inner product in $L^2(K)$ and $\{\delta_K\}$ are local parameters which has to be chosen appropriately.

Next, preliminaries for the analysis are introduced. The elliptic projection $\pi_h \ : \ V \to V_{h,r}$ is defined by

$$(\nabla(u - \pi_h u), \nabla v_h) = 0 \quad \forall\ v_h \in V_{h,r}.$$

Note that the functions of $V_{h,r}$ do not depend on time. Hence, for all $v_h \in V_{h,r}$ hold

$$\begin{aligned}0 &= (\nabla(u_t - \pi_h(u_t)), \nabla v_h) = ((\nabla u)_t - \nabla\pi_h(u_t), \nabla v_h), \\ 0 &= \frac{d}{dt}(\nabla(u - \pi_h u), \nabla v_h) = ((\nabla u)_t - (\nabla\pi_h u)_t, \nabla v_h),\end{aligned}$$

and this inner product defines a norm in $V_{h,r}$, it follows

$$(\pi_h u)_t = \pi_h(u_t) = \pi_h u_t. \qquad (2.5)$$

Assuming that the meshes are quasi-uniform, the following inverse inequality holds for each $v_h \in V_{h,r}$, see, e.g., [4, Theorem 3.2.6],

$$\|v_h\|_{W^{m,q}(K)} \le c_{\mathrm{inv}} h_K^{l-m-d\left(\frac{1}{q'}-\frac{1}{q}\right)} \|v_h\|_{W^{l,q'}(K)}, \qquad (2.6)$$

where $0 \le l \le m \le 1$, $1 \le q' \le q \le \infty$, $h_K$ is the size (diameter) of the mesh cell $K \in \mathcal{T}_h$ and $\|\cdot\|_{W^{m,q}(K)}$ is the norm in $W^{m,q}(K)$. The following interpolation error estimate for $u \in V \cap H^{r+1}(\Omega)$ is well known, [5, 16]

$$\|u - \pi_h u\|_0 + h\|u - \pi_h u\|_1 \le C h^{r+1} \|u\|_{r+1}, \qquad (2.7)$$

where $\|\cdot\|_r$ denotes the norm in $H^r(\Omega)$ with $H^0(\Omega) = L^2(\Omega)$. In particular, stability estimates for $u \in H_0^1(\Omega)$ of the form

$$\|\pi_h u\|_0 \leq \|u - \pi_h u\|_0 + \|u\|_0 \leq Ch\|u\|_1 + \|u\|_0 \leq C\|u\|_1 \tag{2.8}$$

can be derived.

It is assumed that the space $V_{h,r}$ satisfies the following local approximation property: for each $u \in V \cap H^{r+1}(\Omega)$ there exists $\hat{u}_h \in V_{h,r}$ such that

$$\|u - \hat{u}_h\|_{0,K} + h_K\|\nabla(u - \hat{u}_h)\|_{0,K} + h_K^2\|\Delta(u - \hat{u}_h)\|_{0,K} \leq Ch_K^{r+1}\|u\|_{r+1,K} \tag{2.9}$$

for all $K \in \mathcal{T}_h$. For example, this property is given for Lagrange finite elements on mesh cells which allow an affine transform to a reference mesh cell.

LEMMA 2.1. *With the local approximation property (2.9) follows for all* $u \in V \cap H^{r+1}(\Omega)$

$$\sum_{K\in\tau_h} \|\Delta(u - \pi_h u)\|_{0,K}^2 \leq Ch^{2r-2}\|u\|_{r+1}^2. \tag{2.10}$$

*Proof.* The triangle inequality, the local approximation property (2.9) and the inverse inequality (2.6) give

$$\begin{aligned}\|\Delta(u - \pi_h u)\|_{0,K} &\leq \|\Delta(u - \hat{u}_h)\|_{0,K} + \|\Delta(\hat{u}_h - \pi_h u)\|_{0,K} \\ &\leq ch_K^{r-1}\|u\|_{r+1,K} + c_{\text{inv}}h_K^{-1}\|\hat{u}_h - \pi_h u\|_{1,K}.\end{aligned}$$

Squaring this inequality, taking the sum over all mesh cells and using the quasi-uniformity of the mesh lead to

$$\sum_{K\in\mathcal{T}_h} \|\Delta(u - \pi_h u)\|_{0,K}^2 \leq ch^{2r-2}\sum_{K\in\mathcal{T}_h} \|u\|_{r+1,K}^2 + ch^{-2}\sum_{K\in\mathcal{T}_h} \|\hat{u}_h - \pi_h u\|_{1,K}^2. \tag{2.11}$$

The last term can be estimated using the interpolation error estimate (2.7) and the local approximation property (2.9)

$$\sum_{K\in\mathcal{T}_h} \|\hat{u}_h - \pi_h u\|_{1,K}^2 \leq 2\|u - \pi_h u\|_1^2 + 2\sum_{K\in\mathcal{T}_h} \|u - \hat{u}_h\|_{1,K}^2 \leq ch^{2r}\|u\|_{r+1}^2.$$

Substituting this estimate into (2.11) gives the statement of the lemma. ∎

The coercivity of the bilinear form $a_{\text{SUPG}}(\cdot,\cdot)$ under the condition that the parameters $\{\delta_K\}$ are appropriately bounded from above is a well-known result.

LEMMA 2.2. COERCIVITY OF $a_{\text{SUPG}}(\cdot,\cdot)$. *Let (2.2) be satisfied. If the SUPG parameters are chosen such that*

$$\delta_K \leq \frac{\mu_0}{2\|c\|_{K,\infty}^2}, \quad \delta_K \leq \frac{h_K^2}{2\varepsilon c_{\text{inv}}^2}, \tag{2.12}$$

*then the bilinear form* $a_{\text{SUPG}}(\cdot,\cdot)$ *associated with the SUPG method satisfies*

$$a_{\text{SUPG}}(u_h, u_h) \geq \frac{1}{2}\|u_h\|_{\text{SUPG}}^2 \tag{2.13}$$

*with*

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

*Proof.* See, e.g., [15, Lemma 10.3]. ∎

For linear finite elements, the condition $\delta_K \leq h_K^2/(2\varepsilon c_{\text{inv}}^2)$ can be omitted.

The analysis of a time-continuous problem requires a Gronwall-type estimate.

LEMMA 2.3. GRONWALL-TYPE ESTIMATE. *Let $t > 0$, $a, b, c \in L^1(0,t)$ nonnegative functions and $d, \gamma \in \mathbb{R} \geq 0$. From the inequality*

$$a(t) + \int_0^t b(\tau)\ d\tau \leq \gamma \int_0^t a(\tau)\ d\tau + \int_0^t c(\tau)\ d\tau + d$$

*follows*

$$a(t) + \int_0^t b(\tau)\ d\tau \leq \exp(\gamma t) \left( \int_0^t c(\tau)\ d\tau + d \right)$$

*Proof.* Set

$$\alpha(t) = \gamma \int_0^t a(\tau)\ d\tau + \int_0^t c(\tau)\ d\tau + d - a(t) - \int_0^t b(\tau)\ d\tau \geq 0,$$
$$s(t) = a(t) + \int_0^t b(\tau) + \alpha(t).$$

Note, the last two terms in $s(t)$ are nonnegative. Differentiating $s(t)$ gives

$$s_t(t) = \gamma a(t) + c(t) \leq \gamma s(t) + c(t).$$

Multiplying this inequality with the integrating factor $\exp(-\gamma t)$, integrating in $(0,t)$, and using $s(0) = a(0) + \alpha(0) = d$ give the statement of the lemma. ∎

**3. Stability for stabilization parameters depending on the length of the time step.** This section studies a fully discrete method for solving (2.3). Besides the finite element SUPG discretization (2.4), the temporal derivative is approximated with the backward or implicit Euler scheme.

The approaches used in this section for deriving stability bounds apply standard energy arguments. It turns out that this analysis proposes parameter choices in the SUPG method that depend on the length of the time step.

Consider the case of a fixed time step $k = \Delta t$. The fully discrete solution at time $t_n = nk$ will be denoted by $U_h^n$. The backward Euler/SUPG method reads as follows: For $n = 1, 2, \ldots$ find $U_h^n \in V_{h,r}$ such that

$$\left( \frac{U_h^n - U_h^{n-1}}{k}, \varphi \right) + \varepsilon(\nabla U_h^n, \nabla\varphi) + (\mathbf{b}\cdot\nabla U_h^n, \varphi) + (cU_h^n, \varphi) = (f^n, \varphi)$$
$$+ \sum_{K\in\mathcal{T}_h} \delta_K \left( f^n - \left( \frac{U_h^n - U_h^{n-1}}{k} \right) + \varepsilon\Delta U_h^n - \mathbf{b}\cdot\nabla U_h^n - cU_h^n, \mathbf{b}\cdot\nabla\varphi \right)_K \qquad (3.1)$$

for all $\varphi \in V_{h,r}$ and $U_h^0(\mathbf{x}) = u_h(0,\mathbf{x})$. Method (3.1) can be written equivalently in the form

$$(U_h^n - U_h^{n-1}, \varphi) + k a_{\text{SUPG}}(U_h^n, \varphi) = k(f^n, \varphi) + k \sum_{K\in\mathcal{T}_h} \delta_K (f^n, \mathbf{b}\cdot\nabla\varphi)_K$$
$$- \sum_{K\in\mathcal{T}_h} \delta_K (U_h^n - U_h^{n-1}, \mathbf{b}\cdot\nabla\varphi)_K. \qquad (3.2)$$

THEOREM 3.1. STABILITY, STABILIZATION PARAMETERS PROPORTIONAL TO THE LENGTH OF THE TIME STEP. *Let (2.2) and (2.12) be fulfilled. With the additional condition*

$$\delta_K \leq \frac{k}{4} \quad \forall\, K \in \mathcal{T}_h, \tag{3.3}$$

*the solution of (3.1) satisfies at* $t_n = nk$

$$\|U_h^n\|_0^2 + \frac{k}{2}\sum_{j=1}^{n} \|U_h^j\|_{\mathrm{SUPG}}^2 \leq \|U_h^0\|_0^2 + k\left(\frac{2}{\mu_0} + k\right)\sum_{j=1}^{n}\|f^j\|_0^2.$$

*Proof.* The proof starts in the usual way by setting $\varphi = U_h^n$. This gives with (3.2)

$$\begin{aligned}(U_h^n - U_h^{n-1}, U_h^n) + k a_{\mathrm{SUPG}}(U_h^n, U_h^n) = k(f^n, U_h^n) + k\sum_{K\in\mathcal{T}_h}\delta_K(f^n, \mathbf{b}\cdot\nabla U_h^n)_K \\ - \sum_{K\in\mathcal{T}_h}\delta_K(U_h^n - U_h^{n-1}, \mathbf{b}\cdot\nabla U_h^n)_K.\end{aligned}$$

A straightforward calculation shows

$$(U_h^n - U_h^{n-1}, U_h^n) = \frac{1}{2}\left(\|U_h^n\|_0^2 - \|U_h^{n-1}\|_0^2 + \|U_h^n - U_h^{n-1}\|_0^2\right),$$

such that, with (2.13),

$$\begin{aligned}&\frac{1}{2}\left(\|U_h^n\|_0^2 - \|U_h^{n-1}\|_0^2 + \|U_h^n - U_h^{n-1}\|_0^2\right) + \frac{k}{2}\|U_h^n\|_{\mathrm{SUPG}} \\ &\leq |k(f^n, U_h^n)| + \left|k\sum_{K\in\mathcal{T}_h}\delta_K(f^n, \mathbf{b}\cdot\nabla U_h^n)_K\right| + \left|\sum_{K\in\mathcal{T}_h}\delta_K(U_h^n - U_h^{n-1}, \mathbf{b}\cdot\nabla U_h^n)_K\right|.\end{aligned} \tag{3.4}$$

The first two terms on the right hand side are estimated using the Cauchy–Schwarz inequality and Young's inequality

$$\begin{aligned}|k(f^n, U_h^n)| &= k\left(\frac{f^n}{\mu^{1/2}}, \mu^{1/2}U_h^n\right) \leq k\left\|\frac{f^n}{\mu^{1/2}}\right\|_0^2 + \frac{k}{4}\|\mu^{1/2}U_h^n\|_0^2 \\ &\leq \frac{k}{\mu_0}\|f^n\|_0^2 + \frac{k}{4}\|\mu^{1/2}U_h^n\|_0^2,\end{aligned}$$

and

$$\left|k\sum_{K\in\mathcal{T}_h}\delta_K(f^n, \mathbf{b}\cdot\nabla\varphi)\right| \leq 2k\sum_{K\in\mathcal{T}_h}\delta_K\|f^n\|_{0,K}^2 + \frac{k}{8}\sum_{K\in\mathcal{T}_h}\delta_K\|\mathbf{b}\cdot\nabla U_h^n\|_{0,K}^2.$$

The estimate of the last term on the right hand side of (3.4) uses condition (3.3) on the stabilization parameters

$$\begin{aligned}&\left|\sum_{K\in\mathcal{T}_h}\delta_K(U_h^n - U_h^{n-1}, \mathbf{b}\cdot\nabla U_h^n)_K\right| \\ &\leq \frac{2}{k}\sum_{K\in\mathcal{T}_h}\delta_K\|U_h^n - U_h^{n-1}\|_{0,K}^2 + \frac{k}{8}\sum_{K\in\mathcal{T}_h}\delta_K\|\mathbf{b}\cdot\nabla U_h^n\|_{0,K}^2 \\ &\leq \frac{1}{2}\|U_h^n - U_h^{n-1}\|_0^2 + \frac{k}{8}\sum_{K\in\mathcal{T}_h}\delta_K\|\mathbf{b}\cdot\nabla U_h^n\|_{0,K}^2.\end{aligned}$$

Inserting all estimates leads to

$$\|U_h^n\|_0^2 + \frac{k}{2}\|U_h^n\|_{\mathrm{SUPG}}^2 \leq \|U_h^{n-1}\|_0^2 + \frac{2k}{\mu_0}\|f^n\|_0^2 + 4k\sum_{K\in\mathcal{T}_h}\delta_K\|f^n\|_{0,K}^2. \qquad (3.5)$$

Summation of the time steps $j = 1, \ldots, n$, and using once more condition (3.3) gives the statement of the theorem. □

Note that $k\sum_{j=1}^n \|U_h^j\|_{\mathrm{SUPG}}^2$ is an approximation of $\|U_h^j\|_{L^2(0,T;\mathrm{SUPG})}^2$ by a Riemann sum using as node in the quadrature rule always the right end of the time intervals.

Theorem 3.1 covers the case that the stabilization parameter is proportional to the length of the time step. On a fixed spatial grid, the stabilization becomes small for small time steps and it vanishes in the time-continuous limit. This behavior does not seem to be correct, see the discussion in the introduction. The desired situation in the convection-dominated regime, $\delta_K \sim h_K$, is obtained if spatial and temporal mesh width are proportional $h \sim k$. Note that for the mesh width and the time step being of the same order, the parameter choice of [10, 11] leads also to $\delta \sim k \sim h$.

THEOREM 3.2. STABILITY, STABILIZATION PARAMETERS PROPORTIONAL TO SOME FUNCTION OF THE LENGTH OF THE TIME STEP. *Let (2.2) and (2.12) be fulfilled. With the choice*

$$\delta_K = \frac{\sigma(k)h_K}{\|\mathbf{b}\|_{\infty,K}c_{\mathrm{inv}}} \quad \textit{with} \quad 0 < \sigma(k) \leq \frac{1}{4} \quad \forall\, K \in \mathcal{T}_h, \qquad (3.6)$$

*where $\sigma(k)$ is a function to be specified later, the solution of (3.1) satisfies at $t_n = nk$*

$$\begin{aligned}
&\|U_h^n\|_0^2 + \frac{k}{2}\sum_{j=1}^n \|U_h^j\|_{\mathrm{SUPG}}^2 \\
&\leq (1+2\sigma^2(k))^n \left[\|U_h^0\|_0^2 + 2k\sum_{j=1}^n\left(\frac{1}{\mu_0}\|f^j\|_0^2 + \sum_{K\in\mathcal{T}_h}\delta_K\|f^j\|_{0,K}^2\right)\right]. \qquad (3.7)
\end{aligned}$$

*Proof.* The proof starts exactly as the proof of Theorem 3.1 until estimate (3.4) is reached. The first two terms on the right hand side of (3.4) are estimated also in the same way as in the proof of Theorem 3.1

$$\begin{aligned}
|k(f^n, U_h^n)| &\leq \frac{k}{\mu_0}\|f^n\|_0^2 + \frac{k}{4}\|\mu^{1/2}U_h^n\|_0^2, \\
\left|k\sum_{K\in\mathcal{T}_h}\delta_K(f^n, \mathbf{b}\cdot\nabla\varphi)\right| &\leq k\sum_{K\in\mathcal{T}_h}\delta_K\|f^n\|_{0,K}^2 + \frac{k}{4}\sum_{K\in\mathcal{T}_h}\delta_K\|\mathbf{b}\cdot\nabla U_h^n\|_{0,K}^2.
\end{aligned}$$

The last term on the right hand side of (3.4) will now not be absorbed into $\frac{k}{2}\|U_h^n\|_{\mathrm{SUPG}}$.

It is estimated by using the inverse inequality (2.6) and Young's inequality

$$
\begin{aligned}
&\left|\sum_{K\in\mathcal{T}_h}\delta_K(U_h^n-U_h^{n-1},\mathbf{b}\cdot\nabla U_h^n)_K\right| \\
&=\left|\sum_{K\in\mathcal{T}_h}\delta_K(U_h^n-U_h^{n-1},\mathbf{b}\cdot\nabla(U_h^n-U_h^{n-1}))_K+\sum_{K\in\mathcal{T}_h}\delta_K(U_h^n-U_h^{n-1},\mathbf{b}\cdot\nabla U_h^{n-1})_K\right| \\
&\le\sum_{K\in\mathcal{T}_h}\delta_K\frac{\|\mathbf{b}\|_{\infty,K}c_{\mathrm{inv}}}{h_K}\|U_h^n-U_h^{n-1}\|_{0,K}^2 \\
&\quad+\sum_{K\in\mathcal{T}_h}\delta_K\|\mathbf{b}\|_{\infty,K}\|U_h^n-U_h^{n-1}\|_{0,K}\|\nabla U_h^{n-1}\|_{0,K} \\
&\le\sum_{K\in\mathcal{T}_h}\delta_K\frac{\|\mathbf{b}\|_{\infty,K}c_{\mathrm{inv}}}{h_K}\|U_h^n-U_h^{n-1}\|_{0,K}^2+\sum_{K\in\mathcal{T}_h}\frac{1}{4}\|U_h^n-U_h^{n-1}\|_{0,K}^2 \\
&\quad+\sum_{K\in\mathcal{T}_h}\delta_K^2\|\mathbf{b}\|_{\infty,K}^2\|\nabla U_h^{n-1}\|_{0,K}^2 \\
&\le\sum_{K\in\mathcal{T}_h}\left(\delta_K\frac{\|\mathbf{b}\|_{\infty,K}c_{\mathrm{inv}}}{h_K}+\frac{1}{4}\right)\|U_h^n-U_h^{n-1}\|_{0,K}^2+\sum_{K\in\mathcal{T}_h}\delta_K^2\frac{\|\mathbf{b}\|_{\infty,K}^2c_{\mathrm{inv}}^2}{h_K^2}\|U_h^{n-1}\|_{0,K}^2.
\end{aligned}
$$

The first term can be absorbed into the left hand side of (3.4) if

$$
\delta_K\frac{\|\mathbf{b}\|_{\infty,K}c_{\mathrm{inv}}}{h_K}+\frac{1}{4}\le\frac{1}{2}\implies\delta_K\le\frac{h_K}{4\|\mathbf{b}\|_{\infty,K}c_{\mathrm{inv}}}.
$$

Set the stabilization parameter as in (3.6), then it follows

$$
\left|\sum_{K\in\mathcal{T}_h}\delta_K(U_h^n-U_h^{n-1},\mathbf{b}\cdot\nabla U_h^n)_K\right|\le\frac{1}{2}\|U_h^n-U_h^{n-1}\|_0^2+\sigma^2(k)\|U_h^{n-1}\|_0^2.
$$

Collecting all estimates leads to the recursion

$$
\|U_h^n\|_0^2+\frac{k}{2}\|U_h^n\|_{\mathrm{SUPG}}^2\le(1+2\sigma^2(k))\|U_h^{n-1}\|_0^2+\frac{2k}{\mu_0}\|f^n\|_0^2+2k\sum_{K\in\mathcal{T}_h}\delta_K\|f^n\|_{0,K}^2. \quad (3.8)
$$

Now, one obtains by induction

$$
\begin{aligned}
&\|U_h^n\|_0^2+\frac{k}{2}\|U_h^n\|_{\mathrm{SUPG}}^2 \\
&\le(1+2\sigma^2(k))^n\|U_h^0\|_0^2+2k\sum_{j=1}^n\left(1+2\sigma^2(k)\right)^{n-j}\left(\frac{\|f^j\|_0^2}{\mu_0}+\sum_{K\in\mathcal{T}_h}\delta_K\|f^j\|_{0,K}^2\right) \\
&\le(1+2\sigma^2(k))^n\left[\|U_h^0\|_0^2+2k\sum_{j=1}^n\left(\frac{\|f^j\|_0^2}{\mu_0}+\sum_{K\in\mathcal{T}_h}\delta_K\|f^j\|_{0,K}^2\right)\right]. \qquad (3.9)
\end{aligned}
$$

Summation of (3.8) gives

$$
\begin{aligned}
\|U_h^n\|_0^2+\frac{k}{2}\sum_{j=1}^n\|U_h^j\|_{\mathrm{SUPG}}^2\le{}&2\sigma^2(k)\sum_{j=1}^{n-1}\|U_h^j\|_0^2+(1+2\sigma^2(k))\|U_h^0\|_0^2 \\
&+2k\sum_{j=1}^n\left(\frac{\|f^j\|_0^2}{\mu_0}+\sum_{K\in\mathcal{T}_h}\delta_K\|f^j\|_{0,K}^2\right).
\end{aligned}
$$

Inserting (3.9) and applying some estimates for the sake of simplifying the representation lead to

$$
\begin{aligned}
&\|U_h^n\|_0^2 + \frac{k}{2}\sum_{j=1}^{n}\|U_h^j\|_{\mathrm{SUPG}}^2 \\
&\leq 2\sigma^2(k)\sum_{j=1}^{n-1}(1+2\sigma^2(k))^j\|U_h^0\|_0^2 + (1+2\sigma^2(k))\|U_h^0\|_0^2 \\
&\quad +2\sigma^2(k)\sum_{j=1}^{n-1}(1+2\sigma^2(k))^j 2k\sum_{l=1}^{j}\left(\frac{\|f^l\|_0^2}{\mu_0} + \sum_{K\in\mathcal{T}_h}\delta_K\|f^l\|_{0,K}^2\right) \\
&\quad +2k\sum_{j=1}^{n}\left(\frac{\|f^j\|_0^2}{\mu_0} + \sum_{K\in\mathcal{T}_h}\delta_K\|f^j\|_{0,K}^2\right) \\
&\leq \left(2\sigma^2(k)\frac{(1+2\sigma^2(k))^n - (1+2\sigma^2(k))}{1+2\sigma^2(k)-1} + 1 + 2\sigma^2(k)\right)\|U_h^0\|_0^2 \\
&\quad +2k\left[\left(1+2\sigma^2(k)\sum_{j=1}^{n-1}(1+2\sigma^2(k))^j\right)\sum_{j=1}^{n}\left(\frac{\|f^j\|_0^2}{\mu_0} + \sum_{K\in\mathcal{T}_h}\delta_K\|f^j\|_{0,K}^2\right)\right] \\
&\leq (1+2\sigma^2(k))^n\left[\|U_h^0\|_0^2 + 2k\sum_{j=1}^{n}\left(\frac{\|f^j\|_0^2}{\mu_0} + \sum_{K\in\mathcal{T}_h}\delta_K\|f^j\|_{0,K}^2\right)\right].
\end{aligned}
$$

□

Consider a finite time interval $[0,T]$ and a fixed length of the time step. Then, Theorem 3.2 gives stability with the desired stability parameter (in the convection-dominated regime) $\delta_K = \mathcal{O}(h_K)$ without a coupling of the mesh width to the time step by choosing $\sigma(k) = const \leq 1/4$. However, the stability bound blows up for $\sigma(k) = const$ in the time-continuous limit $k \to 0$. Given a length of the time step $k$, the number of time steps to solve the equation in $[0,T]$ is $n = T/k$. The stability estimate will not blow up for $k \to 0$ if $(1+\sigma^2(k))^{1/k}$ is bounded uniformly. A possible choice is $\sigma(k) = \delta_0\sqrt{k}$ leading to the stabilization parameter

$$
\delta_K = \delta_0\frac{\sqrt{k}h_K}{\|\mathbf{b}\|_{\infty,K}c_{\mathrm{inv}}}, \tag{3.10}
$$

where $\delta_0$ has to be chosen such that $\delta_0\sqrt{k} \leq 1/4$. For fixed $h$ and sufficiently small $k$, the parameter from (3.10) is larger than the parameter from (3.3).

**4. Error estimates for stabilization parameters depending on the length of the time step.** For the following error analysis, it is assumed that all functions are sufficiently regular. Summaries of these assumptions are given below in theorems. The error analysis for (3.1) starts by decomposing the error into an interpolation error and the difference of the interpolation and the solution

$$
U_h^n - u(t_n) = (U_h^n - \pi_h u(t_n)) + (\pi_h u(t_n) - u(t_n)).
$$

The interpolation error can be estimated with (2.7). For brevity, denote

$$
\pi_h^n u := \pi_h u(t_n), \quad e_h^n = U_h^n - \pi_h u(t_n).
$$

Straightforward calculations yield the following error equation

$$\begin{aligned}
&(e_h^n - e_h^{n-1}, \varphi) + k a_{\mathrm{SUPG}}(e_h^n, \varphi) \\
&= k(\tilde{T}_{\mathrm{zero}}^n, \varphi) + k(T_{\mathrm{conv}}^n, \varphi) + k \sum_{K \in \mathcal{T}_h} \delta_K (\tilde{T}_{\mathrm{stab},K}^n, \mathbf{b} \cdot \nabla \varphi)_K \\
&\quad - \sum_{K \in \mathcal{T}_h} \delta_K (e_h^n - e_h^{n-1}, \mathbf{b} \cdot \nabla \varphi)_K,
\end{aligned}$$

with

$$\begin{aligned}
\tilde{T}_{\mathrm{zero}}^n &= (u_t(t_n) - \pi_h^n u_t) + c\,(u(t_n) - \pi_h^n u) + \left( \pi_h u_t(t_n) - \frac{\pi_h^n u - \pi_h^{n-1} u}{k} \right), \\
T_{\mathrm{conv}}^n &= \mathbf{b} \cdot \nabla (u(t_n) - \pi_h^n u), \\
\tilde{T}_{\mathrm{stab},K}^n &= \left( \tilde{T}_{\mathrm{zero}}^n + T_{\mathrm{conv}}^n + \varepsilon \Delta (\pi_h^n u - u(t_n)) \right)\big|_K.
\end{aligned}$$

Using integration by parts and assuming $\delta_K > 0$, the convective term can be distributed to the term with the zeroth order derivatives (with respect to space) and the stabilization term

$$\begin{aligned}
(T_{\mathrm{conv}}^n, \varphi) &= -((\nabla \cdot \mathbf{b})(\pi_h^n u - u(t_n)), \varphi) - (\pi_h^n u - u(t_n), \mathbf{b} \cdot \nabla \varphi) \\
&= -((\nabla \cdot \mathbf{b})(\pi_h^n u - u(t_n)), \varphi) - \sum_{K \in \mathcal{T}_h} \delta_K \left( \frac{\pi_h^n u - u(t_n)}{\delta_K}, \mathbf{b} \cdot \nabla \varphi \right)_K.
\end{aligned}$$

Redefining the zeroth order and the stabilization term

$$T_{\mathrm{zero}}^n = \tilde{T}_{\mathrm{zero}}^n - (\nabla \cdot \mathbf{b})(\pi_h^n u - u(t_n)), \quad T_{\mathrm{stab},K}^n = \tilde{T}_{\mathrm{stab},K}^n - \frac{\pi_h^n u - u(t_n)}{\delta_K},$$

leads to the error equation

$$\begin{aligned}
(e_h^n - e_h^{n-1}, \varphi) + k a_{\mathrm{SUPG}}(e_h^n, \varphi) &= k(T_{\mathrm{zero}}^n, \varphi) + k \sum_{K \in \mathcal{T}_h} \delta_K (T_{\mathrm{stab},K}^n, \mathbf{b} \cdot \nabla \varphi)_K \\
&\quad - \sum_{K \in \mathcal{T}_h} \delta_K (e_h^n - e_h^{n-1}, \mathbf{b} \cdot \nabla \varphi)_K. \qquad (4.1)
\end{aligned}$$

This error equation is similar to equation (3.1), only the arguments on the first two terms on the right hand side are not the same.

Deriving error estimates from (4.1) starts essentially in the same way as the derivation of the stability bounds. After this, the arising terms have to be bounded by norms of the solution of the continuous equation (2.3). Since the stability bounds derived in Theorems 3.1 and 3.2 are similar, the detailed analysis for the error estimates is presented here only for the case that was considered in Theorem 3.1.

For proving stability of (4.1), only the last two terms cannot be combined in the summation of the analog to (3.5). One gets

$$\|e_h^n\|_0^2 + \frac{k}{2} \sum_{j=1}^n \|e_h^j\|_{\mathrm{SUPG}}^2 \leq \|e_h^0\|_0^2 + \frac{2k}{\mu_0} \sum_{j=1}^n \|T_{\mathrm{zero}}^j\|_0^2 + 4k \sum_{j=1}^n \sum_{K \in \mathcal{T}_h} \delta_K \|T_{\mathrm{stab},K}^j\|_{0,K}^2. \qquad (4.2)$$

Using the triangle inequality and (2.7), one obtains

$$\|T^j_{\text{zero}}\|_0^2 \le Ch^{2r+2}\Big(\|u_t(t_j)\|_{r+1}^2 + \|c\|^2_{L^\infty(0,T;L^\infty)}\|u(t_j)\|_{r+1}^2 + \|\nabla\cdot\mathbf{b}\|^2_{L^\infty(0,T;L^\infty)}\|u(t_j)\|_{r+1}^2\Big) + C\left\|\pi_h u_t(t_j) - \frac{\pi_h^j u - \pi_h^{j-1}u}{k}\right\|_0^2.$$

The last term is in essence the approximation error of $u_t(t_j)$ by a backward finite difference, hence an estimate of $\mathcal{O}(k)$ can be expected. The derivation of this estimate uses Taylor's formula with remainder in integral form, the application of (2.5), the Cauchy–Schwarz inequality, and the stability estimate (2.8)

$$\begin{aligned}\left\|\pi_h u_t(t_j) - \frac{\pi_h^j u - \pi_h^{j-1}u}{k}\right\|_0^2 &= \frac{1}{k^2}\left\|\int_{t_{j-1}}^{t_j}(t-t_{j-1})\pi_h u_{tt}\,dt\right\|_0^2\\ &\le \frac{1}{k^2}\left(\left(\int_{t_{j-1}}^{t_j}(t-t_{j-1})^2\,dt\right)^{1/2}\left(\int_{t_{j-1}}^{t_j}\|\pi_h u_{tt}\|_0^2\,dt\right)^{1/2}\right)^2\\ &\le Ck\int_{t_{j-1}}^{t_j}\|u_{tt}\|_1^2\,dt = Ck\|u_{tt}\|^2_{L^2(t_{j-1},t_j;H^1)}.\end{aligned}$$

Summation over the time steps, taking into account that the number of time steps $n$ is inverse proportional to the length of the time step, and assuming that all norms are uniformly (in time) bounded gives

$$k\sum_{j=1}^n\|T^j_{\text{zero}}\|_0^2 \le Cknh^{2r+2} + Ck^2\|u_{tt}\|^2_{L^2(0,t_n;H^1)} \le C\left(h^{2r+2}+k^2\right).$$

The estimate of the first term can be applied, in combination with (2.10), to obtain an estimate for the second term on the right hand side of (4.2)

$$\begin{aligned}&\sum_{K\in\mathcal{T}_h}\delta_K\|T^j_{\text{stab},K}\|^2_{0,K}\\ &\le C\left(\max_{K\in\mathcal{T}_h}\delta_K\right)\Big(h^{2r+2}\left(\|u_t(t_j)\|_{r+1}^2 + \|u(t_j)\|_{r+1}^2\right)\\ &\quad + k\|u_{tt}\|^2_{L^2(t_{j-1},t_j;H^1)} + \|\mathbf{b}\|^2_{L^\infty(0,T;L^\infty)}h^{2r}\|u(t_j)\|_{r+1}^2\\ &\quad + \varepsilon^2h^{2r-2}\|u(t_j)\|_{r+1}^2\Big) + C\left(\min_{K\in\mathcal{T}_h}\delta_K\right)^{-1}h^{2r+2}\|u(t_j)\|_{r+1}^2.\end{aligned}$$

Hence,

$$k\sum_{j=1}^n\sum_{K\in\mathcal{T}_h}\delta_K\|T^j_{\text{stab},K}\|^2_{0,K} \le C\left(\left(\max_{K\in\mathcal{T}_h}\delta_K\right)\left(h^{2r+2}+k^2+h^{2r}+\varepsilon^2h^{2r-2}\right) + \left(\min_{K\in\mathcal{T}_h}\delta_K\right)^{-1}h^{2r+2}\right).$$

Inserting all estimates into (4.2) and applying the triangle inequality leads to the following error estimates.

THEOREM 4.1. ERROR ESTIMATES FOR THE STABILIZATION PARAMETER OBEYING (3.3). *Suppose* $\mathbf{b} \in L^\infty(0,T;(L^\infty)^d)$, $\nabla \cdot \mathbf{b}, c \in L^\infty(0,T;L^\infty)$ *for the coefficients in (2.3) and* $u, u_t \in L^\infty(0,T;H^{r+1})$, $u_{tt} \in L^2(0,T;H^1)$ *for the solution of (2.3). Let the stabilization parameters* $\{\delta_K\}$ *fulfill (2.12), (3.3) and* $\delta_K > 0$ *for all* $K \in \mathcal{T}_h$. *Denote* $\delta = \max_{\in \mathcal{T}_h} \delta_K$. *Then, the error* $U_h^n - u(t_n)$ *satisfies*

$$\|U_h^n - u(t_n)\|_0 \leq C\Bigg[h^{r+1} + k + h^{r-1}\delta^{1/2}\Big(h^2 + h + \varepsilon\Big) + \frac{h^{r+1}}{(\min_{K\in\mathcal{T}_h}\delta_K)^{1/2}} + \|\pi_h u_0 - U_h^0\|_0\Bigg], \tag{4.3}$$

*and*

$$\left(k\sum_{j=1}^{n}\|U_h^j - u(t_j)\|_{\mathrm{SUPG}}^2\right)^{1/2} \leq C\Bigg[h^r(\varepsilon^{1/2} + \delta^{1/2} + h) + k + h^{r-1}\delta^{1/2}\Big(h^2 + h + \varepsilon\Big) + \frac{h^{r+1}}{(\min_{K\in\mathcal{T}_h}\delta_K)^{1/2}} + \|\pi_h u_0 - U_h^0\|_0\Bigg], \tag{4.4}$$

*where the constants* $C$ *depend on* $u, u_t, u_{tt}, \mathbf{b}, \nabla\cdot\mathbf{b}$ *and* $c$.

Applying the analysis of Theorem 3.2 to estimate (4.1) and using (3.7) leads essentially to (4.2), only with an additional factor of $(1+2\sigma^2(k))^n$ on the right hand side. The same analysis as in the proof of Theorem 4.1 gives the following error estimates.

THEOREM 4.2. ERROR ESTIMATES FOR THE STABILIZATION PARAMETER PROPORTIONAL TO SOME FUNCTION OF THE LENGTH OF THE TIME STEP. *Let the assumptions on the coefficients and solution of (2.3) be the same as in Theorem 4.1. Let the stabilization parameters* $\{\delta_K\}$ *defined in (3.6) such that (2.12) is fulfilled, too, and* $\delta_K > 0$ *for all* $K \in \mathcal{T}_h$. *Denote* $\delta = \max_{\in\mathcal{T}_h}\delta_K$. *Then, the error* $U_h^n - u(t_n)$ *satisfies*

$$\|U_h^n - u(t_n)\|_0 \leq C(1+2\sigma^2(k))^n\Bigg[h^{r+1} + k + h^{r-1}\delta^{1/2}\Big(h^2 + h + \varepsilon\Big) + \frac{h^{r+1}}{(\min_{K\in\mathcal{T}_h}\delta_K)^{1/2}} + \|\pi_h u_0 - U_h^0\|_0\Bigg], \tag{4.5}$$

*and*

$$\left(k\sum_{j=1}^{n}\|U_h^j - u(t_j)\|_{\mathrm{SUPG}}^2\right)^{1/2} \leq C(1+2\sigma^2(k))^n\Bigg[h^r(\varepsilon^{1/2} + \delta^{1/2} + h) + k + h^{r-1}\delta^{1/2}\Big(h^2 + h + \varepsilon\Big) + \frac{h^{r+1}}{(\min_{K\in\mathcal{T}_h}\delta_K)^{1/2}} + \|\pi_h u_0 - U_h^0\|_0\Bigg], \tag{4.6}$$

*where the constants* $C$ *depend on* $u, u_t, u_{tt}, \mathbf{b}, \nabla\cdot\mathbf{b}$ *and* $c$.

**5. Error analysis of a special time-continuous problem with stabilization parameters depending not on the length of the time step.** The numerical analysis presented so far is only valid if, for a constant mesh and a small time step, the stabilization parameters are sufficiently small. In the time-continuous limit, the SUPG stabilization even vanishes. As discussed in the introduction and as demonstrated in the numerical studies, Example 6.2, we think that this is not the correct asymptotic of the stabilization parameters. This section shows that error estimates with stabilization parameters proportional to the mesh width can be derived for a special time-continuous problem.

In the first step, an error estimate for the material derivative is derived, Theorem 5.1. The analysis of this step uses some ideas from [3], like the application of a special test function to obtain (5.9). Extensions of the analysis from [3] were necessary to include diffusion and reaction. Based on the estimate for the material derivative, an error estimate for the streamline derivative is proven in a second step.

Lets consider problem (2.1) with $\mathbf{b}_t(t,\mathbf{x}) = \mathbf{0}$, $c_t(t,\mathbf{x}) = 0$, i.e., $\mathbf{b} = \mathbf{b}(\mathbf{x}), c = c(\mathbf{x})$ and $\nabla \cdot \mathbf{b} = 0$ for all $\mathbf{x} \in \Omega$. Condition (2.2) reads in this case

$$0 < \mu_0 = \inf_{\mathbf{x}\in\Omega} \mu(\mathbf{x}) = \inf_{\mathbf{x}\in\Omega} c(\mathbf{x}).$$

From the divergence-free condition on $\mathbf{b}$ follows

$$(v, \mathbf{b}\cdot\nabla v) = 0 \quad \forall\ v \in H_0^1(\Omega). \tag{5.1}$$

It is assumed that all functions are sufficiently smooth such that all norms appearing below are well defined. Further, it is assumed that a uniform mesh with width $h$ and $P_1$ finite elements are used. It follows that the stabilization term with the Laplacian does not appear. In addition, only the convection-dominated regime is considered, i.e. it is assumed that $\varepsilon \leq h$. Then, the stabilization parameters are set to be

$$\delta_K = \delta = \min\left\{ \frac{h}{4c_{\rm inv}\|\mathbf{b}\|_\infty} \min\left\{1, \mu_0^{1/2}, \frac{1}{\|c\|_\infty^{1/2}}, \frac{\mu_0^{1/2}}{\|c\|_\infty^{1/2}}, \frac{\mu_0^{1/2}}{\|c\|_\infty}\right\}, \frac{\mu_0^{1/2}}{4\|\mathbf{b}\|_\infty\|\nabla c\|_\infty}, 2\right\}. \tag{5.2}$$

Hence, the stabilization parameters are proportional to the mesh width and they are bounded from above by data of the problem.

Consider a finite time interval $[0,T]$ and let $t \in [0,T]$. In the analysis of this section, a formally steady-state problem derived from (2.1) is used. Let $\Pi_h u(t) \in V_h = V_{h,1}$ be the solution of

$$a_{\rm SUPG}(\Pi_h u(t), v_h) = (f(t) - u_t(t), v_h) + \delta(f(t) - u_t(t), \mathbf{b}\cdot\nabla v_h) \quad \forall\ v_h \in V_h. \tag{5.3}$$

The corresponding continuous equation is solved by $u(t)$. Hence, firstly the Galerkin orthogonality of the SUPG method gives

$$a_{\rm SUPG}(\Pi_h u(t), v_h) = a_{\rm SUPG}(u(t), v_h) \quad \forall\ v_h \in V_h.$$

Secondly, error estimates of the form

$$\|u(t) - \Pi_h u(t)\|_{\rm SUPG} \leq Ch^{3/2}\|u(t)\|_2 \quad t \in [0,T], \tag{5.4}$$

can be proven, see [14]. A straightforward calculation, using the linearity of the equation and the time-independency of convection, reaction and the test functions, shows

$$(\Pi_h u(t))_t = \Pi_h(u_t(t)) = \Pi_h u_t. \tag{5.5}$$

For brevity, the dependency on time will be omitted from now in the notations.

Let $u_h : (0,T] \to V_h$ be the finite element solution of the continuous-in-time SUPG method

$$(u_{h,t}, v_h) + a_{\mathrm{SUPG}}(u_h, v_h) = (f, v_h) + \delta(f - u_{h,t}, \mathbf{b}\cdot\nabla v_h) \quad \forall\, v_h \in V_h \tag{5.6}$$

with $u_h(0)$ given.

For the error analysis, the following norms in $V_h$ are introduced

$$\|v_h\|_{\mathbf{b}} := \left(\|v_h\|_0^2 + \delta^2\|\mathbf{b}\cdot\nabla v_h\|_0^2\right)^{1/2}, \quad \|v_h\|_{\mathrm{mat}} := \delta^{1/2}\|v_{h,t} + \mathbf{b}\cdot\nabla v_h\|_0.$$

The expression in the second norm is the material derivative. Note, $\|\cdot\|_{\mathbf{b}}$ is equivalent to the $L^2$ norm, since by using the inverse inequality and the definition (5.2) of the stabilization parameter, one obtains

$$\|v_h\|_0 \le \|v_h\|_{\mathbf{b}} \le \left(\|v_h\|_0^2 + \delta^2\|\mathbf{b}\|_\infty^2 c_{\mathrm{inv}}^2 h^{-2}\|v_h\|_0^2\right)^{1/2} \le \frac{\sqrt{17}}{4}\|v_h\|_0.$$

Denote the error between the continuous-in-time finite element solution and the solution of the steady-state problem by $e_h = u_h - \Pi_h u$ and let $T_{\mathrm{trunc}} = u_t - \Pi_h u_t$. An error equation is obtain by subtracting (5.3) from (5.6)

$$(e_{h,t}, v_h) + a_{\mathrm{SUPG}}(e_h, v_h) = (T_{\mathrm{trunc}}, v_h) + \delta(T_{\mathrm{trunc}}, \mathbf{b}\cdot\nabla v_h) - \delta(e_{h,t}, \mathbf{b}\cdot\nabla v_h) \;\forall\, v_h \in V_h. \tag{5.7}$$

Setting in (5.7) $v_h = e_h$ and using (5.1) give

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|e_h\|_0^2 + \varepsilon\|\nabla e_h\|_0^2 + \delta\|\mathbf{b}\cdot\nabla e_h\|_0^2 + \|c^{1/2}e_h\|_0^2 + \delta(e_{h,t}, \mathbf{b}\cdot\nabla e_h)\\ = (T_{\mathrm{trunc}}, e_h + \delta\mathbf{b}\cdot\nabla e_h) - \delta(ce_h, \mathbf{b}\cdot\nabla e_h).\end{aligned} \tag{5.8}$$

Analogously, one obtains for $v_h = e_{h,t}$ in (5.7)

$$\begin{aligned}\|e_{h,t}\|_0^2 + \frac{\varepsilon}{2}\frac{d}{dt}\|\nabla e_h\|_0^2 + (\mathbf{b}\cdot\nabla e_h, e_{h,t}) + \frac{\delta}{2}\frac{d}{dt}\|\mathbf{b}\cdot\nabla e_h\|_0^2 + \frac{1}{2}\frac{d}{dt}\|c^{1/2}e_h\|_0^2\\ = (T_{\mathrm{trunc}}, (e_h + \delta\mathbf{b}\cdot\nabla e_h)_t) - \delta(ce_h, \mathbf{b}\cdot\nabla e_{h,t}).\end{aligned} \tag{5.9}$$

The addition of $\delta$ times (5.9) to (5.8) leads to

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|e_h\|_{\mathbf{b}}^2 + \varepsilon\|\nabla e_h\|_0^2 + \|c^{1/2}e_h\|_0^2 + \|e_h\|_{\mathrm{mat}}^2 + \frac{\varepsilon\delta}{2}\frac{d}{dt}\|\nabla e_h\|_0^2 + \frac{\delta}{2}\frac{d}{dt}\|c^{1/2}e_h\|_0^2\\ = (T_{\mathrm{trunc}}, e_h + \delta\mathbf{b}\cdot\nabla e_h) + \delta(T_{\mathrm{trunc}}, (e_h + \delta\mathbf{b}\cdot\nabla e_h)_t)\\ -\delta(ce_h, \mathbf{b}\cdot\nabla e_h) - \delta^2(ce_h, \mathbf{b}\cdot\nabla e_{h,t}),\end{aligned} \tag{5.10}$$

where the definition of $\|\cdot\|_{\mathbf{b}}$ and

$$\delta\left(\|e_{h,t}\|_0^2 + 2(e_{h,t}, \mathbf{b}\cdot\nabla e_h) + \|\mathbf{b}\cdot\nabla e_h\|_0^2\right) = \|e_h\|_{\mathrm{mat}}^2$$

have been used. Using the inverse inequality and the definition (5.2) of the stabilization parameter yields

$$\begin{aligned}\delta(ce_h, \mathbf{b}\cdot\nabla e_h) &= \delta(c^{1/2}e_h, c^{1/2}\mathbf{b}\cdot\nabla e_h) \le \delta\|c^{1/2}e_h\|_0\|c^{1/2}\|_\infty\|\mathbf{b}\|_\infty c_{\mathrm{inv}}h^{-1}\|e_h\|_0\\ &\le \delta\frac{\|c\|_\infty^{1/2}\|\mathbf{b}\|_\infty c_{\mathrm{inv}}}{h\mu_0^{1/2}}\|c^{1/2}e_h\|_0^2 \le \frac{1}{4}\|c^{1/2}e_h\|_0^2.\end{aligned}$$

Considering the last term of (5.10), $e_{h,t}$ has to be absorbed by the material derivative on the left hand side of (5.10). To this end, integration by parts and $\nabla \cdot \mathbf{b}(\mathbf{x}) = 0$ for all $\mathbf{x} \in \Omega$ give

$$\delta^2(ce_h, \mathbf{b}\cdot\nabla e_{h,t}) = -\delta^2(\mathbf{b}\cdot\nabla(ce_h), e_{h,t} + \mathbf{b}\cdot\nabla e_h) + \delta^2(\mathbf{b}\cdot\nabla(ce_h), \mathbf{b}\cdot\nabla e_h).$$

The estimate of the left hand term on the right hand side is obtained with the Cauchy–Schwarz inequality, Young's inequality, the product rule, and the definition of $\delta$ from (5.2)

$$\begin{aligned}
&\delta^2(\mathbf{b}\cdot\nabla(ce_h), e_{h,t} + \mathbf{b}\cdot\nabla e_h)\\
&\leq \frac{\delta^3}{2}\|\mathbf{b}\cdot\nabla(ce_h)\|_0^2 + \frac{1}{2}\|e_h\|_{\mathrm{mat}}^2\\
&\leq \delta^3\frac{\|\mathbf{b}\|_\infty^2}{\mu_0}\left(\|\nabla c\|_\infty^2 + \frac{\|c\|_\infty^2 c_{\mathrm{inv}}^2}{h^2}\right)\|c^{1/2}e_h\|_0^2 + \frac{1}{2}\|e_h\|_{\mathrm{mat}}^2\\
&\leq \left(\frac{\delta}{16} + \frac{\delta}{16}\right)\|c^{1/2}e_h\|_0^2 + \frac{1}{2}\|e_h\|_{\mathrm{mat}}^2 \leq \frac{1}{4}\|c^{1/2}e_h\|_0^2 + \frac{1}{2}\|e_h\|_{\mathrm{mat}}^2.
\end{aligned}$$

With the previous estimate, one obtains

$$\begin{aligned}
\delta^2(\mathbf{b}\cdot\nabla(ce_h), \mathbf{b}\cdot\nabla e_h) &\leq \frac{\delta^2}{2}\|\mathbf{b}\cdot\nabla(ce_h)\|_0^2 + \frac{\delta^2}{2}\|\mathbf{b}\cdot\nabla e_h\|_0^2\\
&\leq \frac{1}{8}\|c^{1/2}e_h\|_0^2 + \frac{\delta^2}{2}\frac{\|\mathbf{b}\|_\infty^2 c_{\mathrm{inv}}^2}{\mu_0 h^2}\|c^{1/2}e_h\|_0^2\\
&\leq \left(\frac{1}{8} + \frac{1}{32}\right)\|c^{1/2}e_h\|_0^2 \leq \frac{1}{4}\|c^{1/2}e_h\|_0^2.
\end{aligned}$$

For the special case of $c$ being a constant, an inspection of the estimates shows that some conditions in the definition of the stabilization parameter (5.2) can be omitted. Inserting all estimated into (5.10) gives

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\|e_h\|_{\mathbf{b}}^2 + \varepsilon\|\nabla e_h\|_0^2 + \frac{1}{4}\|c^{1/2}e_h\|_0^2 + \frac{1}{2}\|e_h\|_{\mathrm{mat}}^2 + \frac{\varepsilon\delta}{2}\frac{d}{dt}\|\nabla e_h\|_0^2 + \frac{\delta}{2}\frac{d}{dt}\|c^{1/2}e_h\|_0^2\\
&\qquad \leq (T_{\mathrm{trunc}}, e_h + \delta\mathbf{b}\cdot\nabla e_h) + \delta(T_{\mathrm{trunc}}, (e_h + \delta\mathbf{b}\cdot\nabla e_h)_t).
\end{aligned}$$

Integration in $(0,t)$ leads to

$$\begin{aligned}
&\frac{1}{2}\|e_h(t)\|_{\mathbf{b}}^2 + \varepsilon\|\nabla e_h\|_{L^2(0,t;L^2)}^2 + \frac{1}{4}\|c^{1/2}e_h\|_{L^2(0,t;L^2)}^2 + \frac{1}{2}\|e_h\|_{L^2(0,t;\mathrm{mat})}^2\\
&\quad + \frac{\varepsilon\delta}{2}\|\nabla e_h(t)\|_0^2 + \frac{\delta}{2}\|(c^{1/2}e_h)(t)\|_0^2\\
&\leq \frac{1}{2}\|e_h(0)\|_{\mathbf{b}}^2 + \frac{\varepsilon\delta}{2}\|\nabla e_h(0)\|_0^2 + \frac{\delta}{2}\|(c^{1/2}e_h)(0)\|_0^2 + \int_0^t (T_{\mathrm{trunc}}, e_h + \delta\mathbf{b}\cdot\nabla e_h)\, d\tau\\
&\quad + \delta\int_0^t (T_{\mathrm{trunc}}, (e_h + \delta\mathbf{b}\cdot\nabla e_h)_t)\, d\tau. \qquad (5.11)
\end{aligned}$$

Now, the terms on the right hand side of (5.11) have to be bounded. It is

$$\begin{aligned}
&\frac{1}{2}\|e_h(0)\|_{\mathbf{b}}^2 + \frac{\varepsilon\delta}{2}\|\nabla e_h(0)\|_0^2 + \frac{\delta}{2}\|(c^{1/2}e_h)(0)\|_0^2\\
&\leq \left(\frac{17}{32} + \frac{\varepsilon\delta c_{\mathrm{inv}}^2}{2h^2} + \frac{\delta\|c\|_\infty}{2}\right)\|e_h(0)\|_0^2 \leq C\|e_h(0)\|_0^2
\end{aligned}$$

since $\varepsilon \leq h$ is assumed. For the next term, one obtains with the Cauchy–Schwarz inequality, Young's inequality and the definition of $\|\cdot\|_{\mathbf{b}}$

$$\begin{aligned}\int_0^t (T_{\text{trunc}}, e_h + \delta \mathbf{b}\cdot\nabla e_h) &\leq 2T \int_0^t \|T_{\text{trunc}}\|_0^2 \, d\tau + \frac{1}{8T}\int_0^t \|e_h + \delta\mathbf{b}\cdot\nabla e_h\|_0^2 \, d\tau \\ &\leq 2T \int_0^t \|T_{\text{trunc}}\|_0^2 \, d\tau + \frac{1}{4T}\int_0^t \|e_h\|_{\mathbf{b}}^2 \, d\tau.\end{aligned}$$

The last term in (5.11) is integrated by parts in time and then a similar estimate is applied

$$\begin{aligned}&\delta \int_0^t (T_{\text{trunc}}, (e_h + \delta\mathbf{b}\cdot\nabla e_h)_t) \, d\tau \\ &= \delta\big((T_{\text{trunc}}(t), (e_h + \delta\mathbf{b}\cdot\nabla e_h)(t)) - (T_{\text{trunc}}(0), (e_h + \delta\mathbf{b}\cdot\nabla e_h)(0))\big) \\ &\quad -\delta \int_0^t (T_{\text{trunc},t}, e_h + \delta\mathbf{b}\cdot\nabla e_h) \, d\tau \\ &\leq 2\delta^2 \|T_{\text{trunc}}(t)\|_0^2 + \frac{1}{4}\|e_h(t)\|_{\mathbf{b}}^2 + 2\delta^2\|T_{\text{trunc}}(0)\|_0^2 + \frac{1}{4}\|e_h(0)\|_{\mathbf{b}}^2 \\ &\quad +2\delta^2 T \int_0^t \|T_{\text{trunc},t}\|_0^2 \, d\tau + \frac{1}{4T}\int_0^t \|e_h\|_{\mathbf{b}}^2 \, d\tau.\end{aligned}$$

Inserting these estimates into (5.11) and using $\|e_h(0)\|_{\mathbf{b}}^2 \leq C\|e_h(0)\|_0^2$ yield

$$\begin{aligned}&\|e_h(t)\|_{\mathbf{b}}^2 + 4\varepsilon\|\nabla e_h\|_{L^2(0,t;L^2)}^2 + \|c^{1/2}e_h\|_{L^2(0,t;L^2)}^2 + 2\|e_h\|_{L^2(0,t;\text{mat})}^2 \\ &\quad +2\varepsilon\delta\|\nabla e_h(t)\|_0^2 + 2\delta\|(c^{1/2}e_h)(t)\|_0^2 \\ &\leq C\|e_h(0)\|_0^2 + 8T\int_0^t \|T_{\text{trunc}}\|_0^2 \, d\tau + 8\delta^2 T\int_0^t \|T_{\text{trunc},t}\|_0^2 \, d\tau \\ &\quad +16\delta^2\|T_{\text{trunc}}\|_{L^\infty(0,T;L^2)}^2 + \frac{2}{T}\int_0^t \|e_h\|_{\mathbf{b}}^2 \, d\tau.\end{aligned}$$

The Gronwall inequality from Lemma 2.3 leads to

$$\begin{aligned}&\|e_h(t)\|_{\mathbf{b}}^2 + 4\varepsilon\|\nabla e_h\|_{L^2(0,t;L^2)}^2 + \|c^{1/2}e_h\|_{L^2(0,t;L^2)}^2 + 2\|e_h\|_{L^2(0,t;\text{mat})}^2 \\ &\quad +2\varepsilon\delta\|\nabla e_h(t)\|_0^2 + 2\delta\|(c^{1/2}e_h)(t)\|_0^2 \\ &\leq \exp\left(\frac{2t}{T}\right)\Big(C\|e_h(0)\|_0^2 + 8T\int_0^t \|T_{\text{trunc}}\|_0^2 \, d\tau + 8\delta^2 T\int_0^t \|T_{\text{trunc},t}\|_0^2 \, d\tau \\ &\quad +16\delta^2\|T_{\text{trunc}}\|_{L^\infty(0,T;L^2)}^2\Big).\end{aligned}$$

The next step of the error analysis uses that the convection and reaction do not depend on time. Hence (5.3) can be differentiated with respect to time. Using (5.5), one obtains steady-state SUPG problems for $\Pi_h u_t(t)$ and $\Pi_h u_{tt}(t)$ with corresponding error estimates of type (5.4)

$$\|T_{\text{trunc}}(t)\|_{\text{SUPG}} \leq Ch^{3/2}\|u_t(t)\|_2, \quad \|T_{\text{trunc},t}(t)\|_{\text{SUPG}} \leq Ch^{3/2}\|u_{tt}(t)\|_2.$$

It follows

$$\|T_{\text{trunc}}(t)\|_0 \leq C\frac{h^{3/2}}{\mu_0^{1/2}}\|u_t(t)\|_2, \quad \|T_{\text{trunc},t}(t)\|_0 \leq C\frac{h^{3/2}}{\mu_0^{1/2}}\|u_{tt}(t)\|_2. \tag{5.12}$$

Noting $t \leq T$ and summarizing all constant into a generic constant, the following theorem is proven.

THEOREM 5.1. ERROR ESTIMATE FOR NORM INVOLVING THE MATERIAL DERIVATIVE. *Let $t \leq T < \infty$ and let $u_t \in L^\infty(0;T;H^2(\Omega))$, $u_{tt} \in L^2(0,T;H^2(\Omega))$. Then, the error $e_h = u_h - \Pi_h u$ satisfies*

$$\begin{aligned}
&\|e_h(t)\|_{\mathbf{b}} + \Big(\varepsilon\|\nabla e_h\|^2_{L^2(0,t;L^2)} + \|e_h\|^2_{L^2(0,t;\mathrm{mat})} + \|c^{1/2}e_h\|^2_{L^2(0,t;L^2)}\Big)^{1/2} \\
&\quad + \delta^{1/2}\left(\varepsilon^{1/2}\|\nabla e_h(t)\|_0 + \|(c^{1/2}e_h)(t)\|_0\right) \\
&\leq C\Big[\|e_h(0)\|_0 + h^{3/2}\Big(T^{1/2}\|u_t\|_{L^2(0,t;H^2)} + \delta T^{1/2}\|u_{tt}\|_{L^2(0,t;H^2)} \\
&\quad + \delta\|u_t\|_{L^\infty(0,T;H^2)}\Big)\Big],
\end{aligned} \tag{5.13}$$

*where $C$ depends on $\|\mathbf{b}\|_\infty, \mu_0, \|c\|_\infty$ and $\|\nabla c\|_\infty$.*

An estimate for $u - u_h$ is now obtained by applying the triangle inequality and using (5.4) for estimating the terms with $u - \Pi_h u$.

In the second step, an estimate with the stronger SUPG norm $\delta\|\mathbf{b}\cdot\nabla e_h\|_{L^2(0,t;L^2)}$ instead of $\|e_h\|_{L^2(0,t;\mathrm{mat})}$ is derived. To this end, insert once more $v_h = e_h$ into the error equation (5.7) and apply a standard analysis by using the coercivity (2.13)

$$\begin{aligned}
\frac{1}{2}\frac{d}{dt}\|e_h\|_0^2 + \frac{1}{2}\|e_h\|^2_{\mathrm{SUPG}} \leq \frac{\|T_{\mathrm{trunc}}\|_0^2}{\mu_0} + \frac{\|c^{1/2}e_h\|_0^2}{4} + 2\delta\|T_{\mathrm{trunc}}\|_0^2 \\
+ 2\delta\|e_{h,t}\|_0^2 + \delta\frac{\|\mathbf{b}\cdot\nabla e_h\|_0^2}{4}.
\end{aligned} \tag{5.14}$$

The second and the last term can be absorbed into the left hand side. The first and the third are estimated by (5.12). The estimate for the fourth term uses once more that convection and reaction are functions independent of time. Hence, (5.3) and (5.6) can be differentiated with respect to time, leading to the same type of equations. Now, the error analysis for $e_h$ leading to (5.13) can be carried out in the same way for $e_{h,t}$. Using the equivalence of the $L^2$ norm and $\|\cdot\|_{\mathbf{b}}$ gives

$$\begin{aligned}
\|e_{h,t}(t)\|_{L^\infty(0,t;L^2)} \leq C\Big[\|e_{h,t}(0)\|_0 + h^{3/2}\Big(T^{1/2}\|u_{tt}\|_{L^2(0,t;H^2)} \\
+ \delta T^{1/2}\|u_{ttt}\|_{L^2(0,t;H^2)} + \delta\|u_{tt}\|_{L^\infty(0,T;H^2)}\Big)\Big],
\end{aligned} \tag{5.15}$$

since the norms are monotonically increasing. Now, $\|e_{h,t}(0)\|_0$ has to be bounded in terms of $e_h(0)$ and $T_{\mathrm{trunc}}(0)$ since it is not clear how to control $e_{h,t}(0)$ by an appropriate choice of $u_h(0)$. To this end, $e_{h,t}(t)$ is inserted into the error equation (5.7) leading to

$$\|e_{h,t}\|_0^2 = -a_{\mathrm{SUPG}}(e_h, e_{h,t}) + (T_{\mathrm{trunc}}, e_{h,t} + \delta\mathbf{b}\cdot\nabla e_{h,t}). \tag{5.16}$$

Applying the Cauchy–Schwarz inequality and the inverse inequality, using $\varepsilon \leq h$ and (5.2) yields

$$\begin{aligned}
a_{\mathrm{SUPG}}(u_h, v_h) &\leq \Big(\frac{\varepsilon c_{\mathrm{inv}}}{h}\|\nabla u_h\|_0 + \|\mathbf{b}\cdot\nabla u_h\|_0 + \|c\|_\infty^{1/2}\|c^{1/2}u_h\|_0 \\
&\quad + \frac{\delta c_{\mathrm{inv}}\|\mathbf{b}\|_\infty}{h}\|\mathbf{b}\cdot\nabla u_h\|_0 + \frac{\delta c_{\mathrm{inv}}\|\mathbf{b}\|_\infty\|c\|_\infty^{1/2}}{h}\|c^{1/2}u_h\|_0\Big)\|v_h\|_0 \\
&\leq C\left(\|\nabla u_h\|_0 + \|\mathbf{b}\cdot\nabla u_h\|_0 + \|c^{1/2}u_h\|_0\right)\|v_h\|_0,
\end{aligned}$$

where $C$ depends on $c_{\text{inv}}$ and $\|c\|_\infty$. Applying this estimate to (5.16), using (5.2) and (5.12) give

$$\begin{aligned}\|e_{h,t}\|_0^2 &\le C\left(\|\nabla e_h\|_0 + \|\mathbf{b}\cdot\nabla e_h\|_0 + \|c^{1/2}e_h\|_0 + \|T_{\text{trunc}}\|_0\right)\|e_{h,t}\|_0\\ &\le C\left(\|\nabla e_h\|_0 + \|\mathbf{b}\cdot\nabla e_h\|_0 + \|c^{1/2}e_h\|_0 + h^{3/2}\|u_t(t)\|_2\right)\|e_{h,t}\|_0.\end{aligned}$$

Using this estimate for $t = 0$ in (5.15), inserting then (5.15) into (5.14), integrating in $(0, t)$, estimating

$$\int_0^t \|e_{h,t}(\tau)\|_0^2\, d\tau \le T\|e_{h,t}(\tau)\|^2_{L^\infty(0,t;L^2)},$$

and applying the triangle inequality leads to the following error estimate.

THEOREM 5.2. ERROR ESTIMATE FOR NORM INVOLVING THE SUPG NORM. *Let $T < \infty$ be the final time and let $u_t(t) \in H^2(\Omega)$ for all $t \in [0, T]$, $u_{tt} \in L^\infty(0; T; H^2(\Omega))$, $u, u_t, u_{ttt} \in L^2(0, T; H^2(\Omega))$. Then, the error estimate*

$$\begin{aligned}&\|(u-u_h)(t)\|_0 + \|u-u_h\|_{L^2(0,t;\text{SUPG})}\\ &\le C\left[\|e_h(0)\|_0 + \delta^{1/2}T^{1/2}\left(\|\nabla e_h(0)\|_0 + \|(\mathbf{b}\cdot\nabla)e_h(0)\|_0 + \|(c^{1/2}e_h)(0)\|_0\right)\right]\\ &\quad + Ch^{3/2}\left(\|u(t)\|_2 + \delta^{1/2}T^{1/2}\|u_t(0)\|_2 + \|u\|_{L^2(0,t;H^2)} + \|u_t\|_{L^2(0,t;H^2)}\right.\\ &\quad \left. + \delta^{1/2}T\|u_{tt}\|_{L^2(0,t;H^2)} + \delta^{1/2}T\|u_{ttt}\|_{L^2(0,t;H^2)} + \delta^{1/2}T^{1/2}\|u_{tt}\|_{L^\infty(0,t;H^2)}\right)\end{aligned} \tag{5.17}$$

*holds. The constants depend on $\|\mathbf{b}\|_\infty, \mu_0, \|c\|_\infty$, $\|\nabla c\|_\infty$, and $c_{\text{inv}}$.*

Choosing the initial finite element solution $u_h(0)$ such that $u_h(0)$ solves

$$\begin{aligned}a_{\text{SUPG}}(u_h(0), v_h) &= (f(0) - u_t(0), v_h) + \delta(f(0) - u_t(0), \mathbf{b}\cdot\nabla v_h)\\ &= (-\varepsilon\Delta u_0 + \mathbf{b}\cdot\nabla u_0 + cu_0, v_h + \delta\mathbf{b}\cdot\nabla v_h) \quad \forall\ v_h \in V_h\end{aligned}$$

leads to $e_h(0) = 0$ such that all terms with $e_h(0)$ vanish in (5.17).

**6. Numerical studies.** Two examples will be presented in the numerical studies. The first one, possessing a given smooth solution, serves as support for the orders of convergence that are proven in the previous sections. The second example is the well-known rotating body problem from [12]. It demonstrates the superiority of the parameter choice from Section 5 compared with the choices from Sections 3 and 4 for small time steps on a fixed, rather coarse, spatial mesh.

EXAMPLE 6.1. SMOOTH SOLUTION. This example serves for supporting the error estimates (4.3) – (4.6) and (5.17). Consider (2.3) with $\Omega = (0, 1)^2$, $T = 1$, different values of $\varepsilon$, $\mathbf{b} = (1, -1)$, $c = 1$, and the right-hand side is chosen such that

$$u(t, x, y) = e^{\sin(2\pi t)}\sin(2\pi x)\sin(2\pi y)$$

is the solution of (2.3). The simulations were performed with $\varepsilon = 10^{-8}$ in the convection-dominated regime and with $\varepsilon = 1$ in the diffusion-dominated regime. Uniform triangular grids were used with the coarsest grid (level 0) obtained by dividing the unit square with a diagonal from $(0, 0)$ to $(1, 1)$. To prevent superconvergence, the convection field is chosen such that it is not parallel to any grid line.

Consider at the beginning the error estimates (4.3) – (4.6). First, optimal scalings of the mesh width $h$ and the length of the time step $k$ are derived from these estimates.

Then, the error estimates lead to only one asymptotic order of convergence that serves as criterion. The mesh width $h$ was defined by dividing the diameters of the mesh cells by $\sqrt{2}$.

The stabilization parameter for the estimates under the assumptions of Theorem 4.1 is set to be $\delta_K = \delta = k/4$, according to condition (3.3). In the convection-dominated regime, $\varepsilon \ll h$, the terms $\mathcal{O}(k)$ and $\mathcal{O}(h^{r+1}\delta^{-1/2}) = \mathcal{O}(h^{r+1}k^{-1/2})$ have to be balanced to obtain an optimal $L^2$-error estimate (4.3). This leads to the scaling $k = \mathcal{O}(h^{2(r+1)/3})$. The same reasoning applies for the SUPG error (4.4). If the final time $T = 1$ is not obtained exactly with the chosen time steps, the simulations were stopped at the first discrete time larger than $T$.

In the diffusion-dominated regime, $h \leq \varepsilon$, the terms $\mathcal{O}(k)$,$\mathcal{O}(k^{1/2}h^{r-1}\varepsilon)$, and $\mathcal{O}(h^{r+1}k^{-1/2})$ need to be balanced. This leads to $k = \mathcal{O}(h^{2(r+1)/3})$ or $k = (h^2/\varepsilon)$. If $h \ll \varepsilon$, the second scaling gives a better order of convergence for $r = 1$ (piecewise linear elements). Note, in this case, $\delta = k = \mathcal{O}(h^2/\varepsilon)$ is a standard choice of the stabilization parameter in the diffusion-dominated regime for steady-state problems. For $r = 2$, both scalings are essentially the same and for $r \geq 3$, the first scaling leads to a higher order of convergence. For the SUPG estimate, the same terms have to be balanced. In addition, the order of convergence is bounded by the term $\mathcal{O}(\varepsilon^{1/2}h^r)$, such that for $h \ll \varepsilon$ only first order convergence can be expected for $r = 1$.

Figure 6.1 presents the orders of convergence for the $P_1$, $P_2$, and $P_3$ finite element. It can be seen that all orders match the predictions from the analysis.

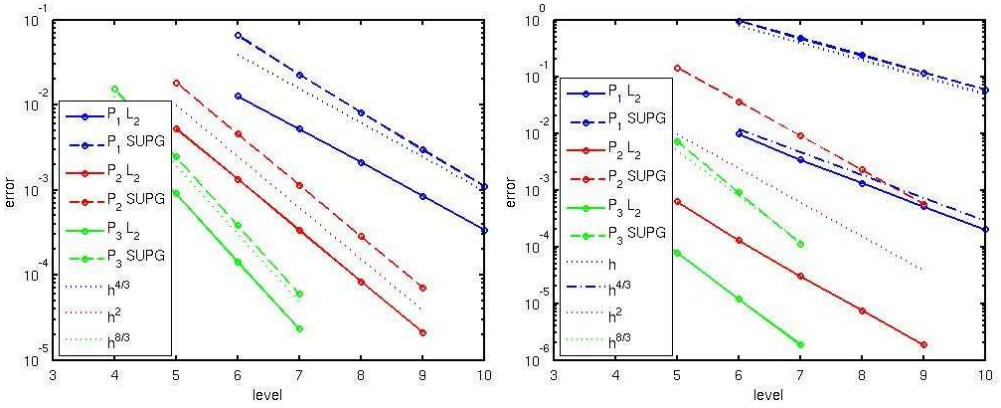


FIG. 6.1. *Example 6.1, orders of convergence for the estimates (4.3) and (4.4); left: convection-dominated regime; right: diffusion-dominated regime.*

Concerning the estimates of Theorem 4.2, the stabilization parameters were chosen to be $\delta_K = \sqrt{k}h_K/(4\|\mathbf{b}\|_2)$, with $\|\mathbf{b}\|_2$ being the (constant) Euclidean norm of the convection vector. In the convection-dominated regime, the terms $\mathcal{O}(k)$ and $\mathcal{O}(h^{r+1/2}k^{-1/4})$ have to be balanced. Thus, the optimal scaling is $k = \mathcal{O}(h^{4(r+1/2)/5})$. This turns out to be the optimal scaling also in the diffusion-dominated regime. For piecewise linear elements, $r = 1$, the stabilization parameter with this scaling is $\delta = \mathcal{O}(h^{8/5})$. Note that in this case, the condition $\delta_K \leq h_K^2/(2\varepsilon c_{\text{inv}}^2)$ does not apply, see the remark after Lemma 2.2. Again, for the SUPG error only first order convergence can be expected for $r = 1$ since the term $\mathcal{O}(\varepsilon^{1/2}h^r)$ occurs in (4.6). The numerical results for the estimates (4.5) and (4.6) are presented in Figure 6.2. They match well the predictions from the analysis.

In further numerical studies at this example, we could observe that also the

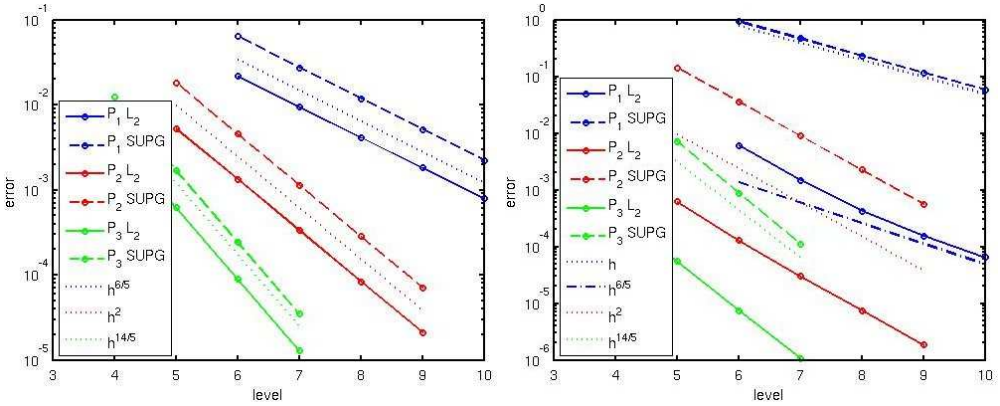


FIG. 6.2. *Example 6.1, orders of convergence for the estimates (4.5) and (4.6); left: convection-dominated regime; right: diffusion-dominated regime.*

Galerkin finite element method gives reasonable results. In particular, the simulations with this method do not blow up. Consequently, we could not observe a blow-up for the case $\delta_K \to 0$ and the term $\mathcal{O}(h^{r+1}\delta^{-1/2})$ is not visible in the computational results. We could not construct an example for that the Galerkin finite element method blows up and for that, consequently, a blow-up of the SUPG method for $\delta \to 0$ can be expected.

Next, estimate (5.17) for the time-continuous case is considered. From this estimate, one can expect convergence for the $L^2$ norm and the SUPG norm of order 3/2 for $P_1$ finite elements and sufficiently small time steps. The length of the time step was set to be $k = 10^{-6}$. As initial condition, the Lagrange interpolant of $u(0, x, y)$ was used. The results are presented in Figure 6.3. The observed order of convergence in the $L^2$ norm is even higher than the prediction by the analysis.

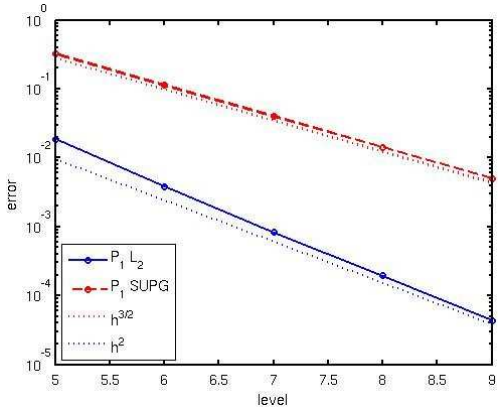


FIG. 6.3. *Example 6.1, orders of convergence for the estimate (5.17), convection-dominated regime.*

EXAMPLE 6.2. ROTATING BODY PROBLEM. This problem was studied numerically for finite element discretizations of convection-diffusion equations already in [10]. Here, exactly the same setting is used. The aim of this example is to illustrate that the choice of the stabilization parameter $\delta_K = \mathcal{O}(h_K)$ from Section 5 is much better

than the choices $\delta_K = \mathcal{O}(k)$, $\delta_K = \mathcal{O}(k^{1/2}h_K)$ from Sections 3 and 4 in the presence of very small time steps.

Let $\Omega = (0,1)^2$, $\varepsilon = 10^{-20}$, $\mathbf{b} = (0.5 - y, x - 0.5)^T$, and $c = f = 0$. The initial condition, consisting of three disjoint bodies, is presented in Figure 6.4. Each body lies within a circle with center $(x_0, y_0)$ and of radius $r_0 = 0.15$. The initial condition is zero outside the three bodies.

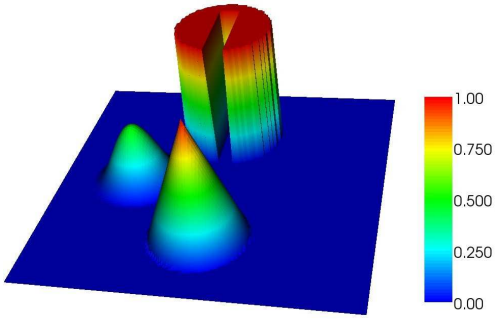


FIG. 6.4. *Example 6.2, initial condition and ideal solution after one rotation.*

Let $r(x,y) = \sqrt{(x-x_0)^2 + (y-y_0)^2}/r_0$. The center of the slotted cylinder is in $(x_0, y_0) = (0.5, 0.75)$ and its shape is given by

$$u(0;x,y) = \begin{cases} 1 & \text{if } r(x,y) \le 1,\ |x - x_0| \ge 0.0225 \text{ or } y \ge 0.85, \\ 0 & \text{else.} \end{cases}$$

The hump at the left hand side is defined by $(x_0, y_0) = (0.25, 0.5)$ and

$$u(0;x,y) = \frac{1}{4}\Big(1 + \cos(\pi \min\{r(x,y), 1\})\Big).$$

On the bottom, a conical body is given by $(x_0, y_0) = (0.5, 0.25)$ and

$$u(0;x,y) = 1 - r(x,y).$$

The rotation of the bodies occurs counter-clockwise. A full revolution takes $t = 6.28 \approx 2\pi$. With the extremely small diffusion, the solution after one revolution is essentially the same as the initial condition. Homogeneous Dirichlet boundary conditions were imposed.

In the simulations, a uniform grid consisting of $128 \times 128$ triangles was used. This leads to 16 641 degrees of freedom for the $P_1$ finite element method, including Dirichlet nodes. The length of the time step was chosen to be $k = 10^{-6}$. Computational studies were performed for the Galerkin finite element method ($\delta_K = 0$ for all mesh cells), the choice of the stabilization parameter from [10, formulae (8) and (11)], that results in $\delta_K = k$, the choice $\delta_K = \sqrt{k}h_K/4$ and $\delta_K = h_K/4$.

Analogously to [10], a measure for the spurious oscillations is given by

$$\mathrm{var}(t) := \max_{(x,y)\in\Omega} u_h(t;x,y) - \min_{(x,y)\in\Omega} u_h(t;x,y),$$

with the optimal value $\mathrm{var}(t) = 1$ for all $t$.

The spurious oscillations of the computed solutions are illustrated in Figure 6.5 and the solutions at the final time in Figure 6.6. It can be observed that by far the

best result was obtained with $\delta_K = h_K/4$. However, the computed solution with these parameters possesses still non-negligible spurious oscillations. Using the stabilization parameters from the analysis of Sections 3 and 4 leads for very small time steps on a fixed spatial grid to similar results as for the Galerkin finite element method. A slight damping of the spurious oscillations can be observed, see the ranges of the finite element solutions in Figure 6.6.

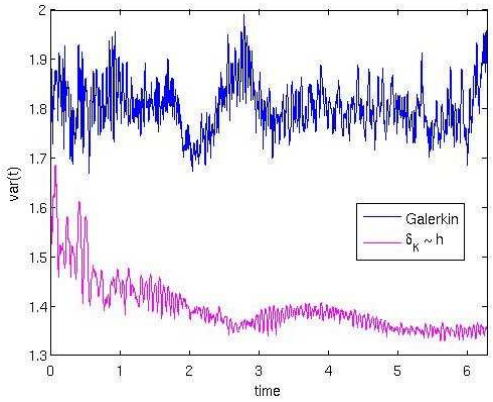


FIG. 6.5. *Example 6.2, spurious oscillations measured by* $var(t)$.

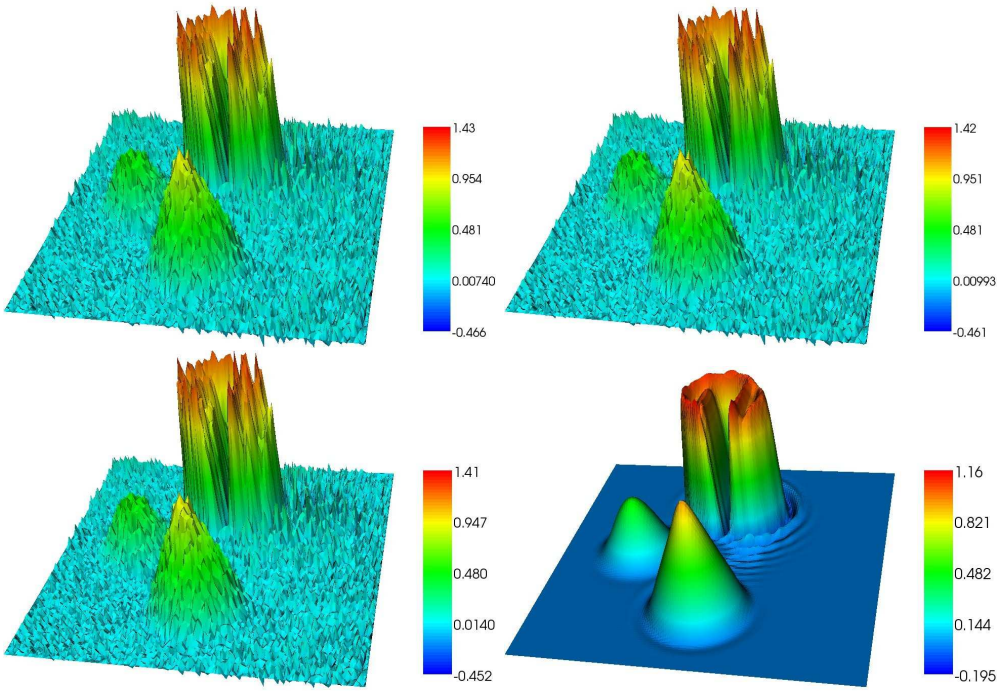


FIG. 6.6. *Example 6.2, computed solutions after one revolution: Galerkin finite element method, SUPG with* $\delta_K = \mathcal{O}(k)$, *SUPG with* $\delta_K = \mathcal{O}(\sqrt{k h_K})$, *SUPG with* $\delta_K = \mathcal{O}(h_K)$; *left to right, top to bottom.*

**7. Summary and Outlook.** This paper studied different ways to obtain error estimates for the SUPG finite element method applied to evolutionary convection-diffusion-reaction equations. For the definition of the fully discrete problem, the backward Euler temporal discretization was considered.

Standard energy arguments for the fully discrete problem yield error estimates under conditions that couple the choice of the stabilization parameters to the length of the time step. In particular, the SUPG stabilization vanishes in the time-continuous limit. Numerical evidence shows that this is not the correct behavior.

For this reason, the time-continuous case was considered for a problem with certain conditions on the coefficients and the $P_1$ finite element on a uniform grid. Error estimates with the expected order of convergence could be proven with the standard choice of the stabilization parameters in the convection-dominated regime $\delta = \mathcal{O}(h)$.

The analysis of the general time-continuous problem, with time-dependent coefficients, is open. An extension of the analysis from Section 5 seems to be hard, since this analysis uses several times that the original equation can be differentiated with respect to time yielding essentially the same equation. Also the cases of higher order finite elements and non-uniform grids in the time-continuous equation have still to be treated.

Concerning the fully discrete case, the deeper reasons for the coupling of the stabilization parameters with the length of the time step are not yet understood. Are these only technical difficulties which might be overcome? Or is there a worst case for that the stability or error analysis with stabilization parameters depending not on the time step is not valid?

With respect to the usage of the SUPG finite element method in time-dependent convection-diffusion-reaction equations, the results of Section 5, Example 6.2 and other numerical studies from the literature strongly suggest to define the stabilization parameters in the convection-dominated regime in the classical way by $\delta_K = \mathcal{O}(h_K)$.

REFERENCES


[1] P.B. Bochev, M.D. Gunzburger, and J.N. Shadid, *Stability of the supg finite element method for transient advection-diffusion problems*, Comput. Methods Appl. Mech. Engrg., 193 (2004), pp. 2301 – 2323.

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

[3] E. Burman, *Consistent SUPG-method for transient transport problems: Stability and convergence*, Comput. Methods Appl. Mech. Engrg., ?? (2010), p. ??

[4] P.G. Ciarlet, *The finite element method for elliptic problems*, North-Holland Publishing Company, Amsterdam – New York – Oxford, 1978.

[5] ———, *Basic error estimates for elliptic problems*, in Handbook of Numerical Analysis II, P.G. Ciarlet and J.L. Lions, eds., North-Holland Amsterdam, New York, Oxford, Tokyo, 1991, pp. 19 – 351.

[6] R. Codina, *Comparison of some finite element methods for solving the diffusion-convection-reaction equation*, Comput. Methods Appl. Mech. Engrg., 156 (1998), pp. 185 – 210.

[7] M. C. Hsu, Y. Bazilevs, V. M. Calo, T. E. Tezduyar, and T. J. R. Hughes, *Improving stability of stabilized and multiscale formulations in flow simulations at small time steps*, Comput. Methods Appl. Mech. Engrg., (2010). in press.

[8] T.J.R. Hughes and A.N. Brooks, *A multidimensional upwind scheme with no crosswind diffusion*, in Finite Element Methods for Convection Dominated Flows, AMD vol.34, T.J.R. Hughes, ed., ASME, New York, 1979, pp. 19 – 35.

[9] V. John, T. Mitkova, M. Roland, K. Sundmacher, L. Tobiska, and A. Voigt, *Simulations of population balance systems with one internal coordinate using finite element methods*, Chem. Engrg. Sci., 64 (2009), pp. 733 – 741.

[10] V. John and E. Schmeyer, *Stabilized finite element methods for time–dependent convection–diffusion–reaction equations*, Comput. Methods Appl. Mech. Engrg., 198 (2008), pp. 475 – 494.

[11] ———, *On finite element methods for 3d time–dependent convection–diffusion–reaction equations with small diffusion*, in BAIL 2008 – Boundary and Interior Layers, vol. 69 of Lecture Notes in Computational Science and Engineering, Springer, 2009, pp. 173 – 182.

[12] R.J. LeVeque, *High–resolution conservative algorithms for advection in incompressible flow*, SIAM J. Numer. Anal., 33 (1996), pp. 627 – 665.

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

[14] H.-G. Roos, M. Stynes, and L. Tobiska, *Robust Numerical Methods for Singularly Perturbed Differential Equations*, vol. 24 of Springer Series in Computational Mathematics, Springer, 2nd ed., 2008.

[15] M. Stynes, *Steady-state convection-diffusion problems*, in Acta Numerica, A. Iserles, ed., Cambridge University Press, 2005, pp. 445 – 508.

[16] L.B. Wahlbin, *Superconvergence in Galerkin Finite Element Methods*, no. 1605 in Lecture Notes in Math., Springer-Verlag, Berlin, 1975.