MATHEMATICS OF COMPUTATION
Volume 87, Number 313, September 2018, Pages 2113–2132
http://dx.doi.org/10.1090/mcom/3326
Article electronically published on January 23, 2018

A UNIFIED FRAMEWORK FOR TIME-DEPENDENT SINGULARLY PERTURBED PROBLEMS WITH DISCONTINUOUS GALERKIN METHODS IN TIME

SEBASTIAN FRANZ AND GUNAR MATTHIES

Abstract. In this paper we present a unified framework for the error analysis of time-dependent singularly perturbed problems with discontinuous Galerkin time discretisation. Its general analysis relies on spatial error estimates known from stationary problems and the properties of the discontinuous Galerkin time discretisation.

We present also applications of our framework to second- and fourth-order singularly perturbed problems in estimation and simulation.

1. Introduction

In the context of singularly perturbed problems many results deal with stationary problems. A widely used technique to obtain accurate solutions in layer regions is the application of layer-adapted meshes. Doing so yields for many finite element methods with or without stabilisation convergence results which are uniform in the perturbation parameter called ε.

When we look at time-dependent singularly perturbed problems and layer-adapted meshes, such uniform results in the literature are much rarer. In [16, 30] standard discretisations in time in combination with Shishkin meshes in space were considered for certain singularly perturbed problems. The analysis in those papers can be extended to more general S-type meshes [24].

In our paper we show, for the first time, a general numerical analysis of time-dependent singularly perturbed problems discretised in time by discontinuous Galerkin methods and applicable to numerous different spatial operators. Many of the ingredients are already known techniques, but in its generality it is a new result. In particular, we consider for a bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$, $d \geq 1$, and a final time $T > 0$ the general class of problems:

$$\left\{\begin{aligned} \partial_t u + Lu &= f && \text{in } \Omega \times (0,T), \\ Bu &= 0 && \text{on } \partial\Omega \times (0,T), \\ u(0) &= u_0 && \text{in } \Omega, \end{aligned}\right. \tag{1.1}$$

where L is a possibly singularly perturbed differential operator in space, B a boundary operator, and u_0 an initial value in Ω. For the notation we will drop in general the dependence on space, such that $u(0) = u(0,x)$ for $x \in \Omega$. Examples for the application of our analysis will be given in Section 4 but for convenience

Received by the editor January 25, 2016, and, in revised form, December 6, 2016, and March 22, 2017.

2010 *Mathematics Subject Classification.* Primary 65M12, 65M15, 65M60.

the reader may think of $Lu = -\varepsilon\Delta u + b\cdot\nabla u + cu$ and $Bu = u$ as a second-order elliptic convection-diffusion problem with homogeneous Dirichlet conditions or $Lu = \varepsilon^2\Delta^2 u + b\Delta u + cu$ and $Bu = (u, \partial_n u)^\top$ as a fourth-order problem with homogeneous Dirichlet and Neumann conditions. For each problem we will denote by V the corresponding Sobolev space wherein the problem is well defined. For the two given examples we would have $V = H_0^1(\Omega)$ for the first one and $V = H_0^2(\Omega)$ for the second one. Furthermore, let V_N denote a finite element space defined on Ω with N being a parameter associated with the number of elements.

We denote by $a(t;\cdot,\cdot)$ the bilinear form associated with the differential operator L evaluated at time t. In many cases the stability of the method can be increased by adding a stabilisation term $s_N(t;\cdot,\cdot)$ which may depend on time. We set $a_N := a + s_N$. Although a_N may be affine linear in the first function argument and linear in the second function argument, we refer to a_N as a bilinear form.

2. Assumptions and notation

We denote by $\|\cdot\|$ and $(\cdot,\cdot)$ the standard L^2-norm and the standard L^2-scalar product on Ω, respectively. The duality pairing between V and its dual space V' is given by $\langle\cdot,\cdot\rangle$. Furthermore, we denote by $W^{k,p}(\Omega)$ the usual Sobolev spaces and in the case $p = 2$ we write $H^m(\Omega)$ instead of $W^{m,2}(\Omega)$. Throughout the paper C denotes a generic constant that is independent of the meshes in time and space, and of a possible perturbation parameter ε of the spatial operator. Labelled constants like C_1 have a fixed value but are also independent of the meshes and the perturbation parameter.

In order to carry out our fairly general analysis, we need some assumptions. They are collected into several groups.

The first assumption deals with the relations between different norms and the stabilisation term s_N.

Assumption 1 (Norms)**.** There exists a norm $|||\cdot|||$ on $V + V_N$ such that

$$\|\xi\| \le C_1\, |||\xi||| \quad \text{and} \quad s_N(t;\xi,\xi) \le C_2\, |||\xi|||^2$$

hold true uniformly for all $\xi \in V + V_N$ and all $t \in [0,T]$.

Here and further on we call a function $g : \mathbb{N} \to [0,\infty)$ *ε-uniformly bounded* if there exists a function $\widetilde{g} : \mathbb{N} \to [0,\infty)$ independent of ε such that $g(N) \le \widetilde{g}(N)$ for all $N \in \mathbb{N}$ and for all $\varepsilon \in (0,1]$. This means in particular that the function $g(N)$ is for $\varepsilon \to 0$ always bounded.

The second assumption considers a spatial interpolation operator I_N.

Assumption 2 (Interpolation)**.** Given the norm $|||\cdot|||$ from Assumption 1, there are an integer $k \in \mathbb{N}$, an interpolation operator $I_N : V \to V_N$, and two ε-uniformly bounded, monotonically decreasing functions $g_{L^2} : \mathbb{N} \to [0,\infty)$ and $g_e : \mathbb{N} \to [0,\infty)$ such that

$$\|z - I_N z\| \le C\, g_{L^2}(N) \tag{2.1}$$

and

$$|||z - I_N z||| \le C\, g_e(N) \tag{2.2}$$

hold true, where $z \in V \cap H^k(\Omega)$ is the solution of the stationary problem $a(t;z,v) = \langle f(t), v\rangle$ for all $v \in V$ and $t \in [0,T]$.

Note that Assumptions 1 and 2 are fulfilled for second-order elliptic problems discretised on properly defined S-type meshes. Here, the spatial analysis relies on precise knowledge of the layer structure which is often given for the unit square or smooth domains; see, e.g., [26, Ch. III.1.4]. Using a simple quasi-uniform mesh results in bounds for the interpolation error which are usually *not* ε-uniformly bounded.

The third assumption deals with the bilinear form of the stationary problem.

Assumption 3 (Bilinear form)**.** Given the norm $|||\cdot|||$ from Assumption 1 and the interpolation operator I_N from Assumption 2, there exist an integer $k \in \mathbb{N}$, a constant $\gamma > 0$, and an ε-uniformly bounded, monotonically decreasing function $g_d : \mathbb{N} \to [0, \infty)$ such that

- The stabilised bilinear form a_N is uniformly coercive on $[0, T]$ with respect to $|||\cdot|||$, i.e.,

$$a_N(t; \varphi, \varphi) \geq \gamma \, |||\varphi|||^2 \qquad \text{for all } \varphi \in V + V_N, \ t \in [0, T]. \tag{2.3}$$

- The bilinear form a_N provides for all $v_N \in V_N$,

$$\left| a_N(t; z - I_N z, v_N) \right| \leq C g_d(N) \, |||v_N||| \, , \tag{2.4}$$

  uniformly in $[0, T]$, where $z \in V \cap H^k(\Omega)$ denotes the solution of the stationary problem $a(t; z, w) = \langle f(t), w \rangle$ for all $w \in V$.

The fourth assumption considers the used stabilisation method.

Assumption 4 (Stabilisation method)**.** Let the stabilisation term s_N either be consistent, i.e.,

$$s_N(t; u, v_N) = 0 \qquad \text{for all } v_N \in V_N, \tag{2.5}$$

where u denotes the solution of (1.1) or provide for all $t \in [0, T]$ the estimate

$$s_N(t; \varphi, \psi) \leq s_N(t; \varphi, \varphi)^{1/2} s_N(t; \psi, \psi)^{1/2} \quad \text{for all } \varphi, \psi \in V + V_N, \tag{2.6}$$

together with the bound

$$s_N(t; \varphi, \varphi) \leq C g_s^2(N) \quad \text{for all } \varphi \in V \cap H^k(\Omega), \tag{2.7}$$

where $g_s : \mathbb{N} \to [0, \infty)$ is an ε-uniformly bounded, monotonically decreansing function. In the case of a consistent stabilisation method we set $g_s(N) = 0$ for all $N \in \mathbb{N}$.

In Section 4, we present examples fulfilling the above assumptions.

The final assumption deals with the time derivatives of u.

Assumption 5 (Time behaviour)**.** For a given nonnegative integer q, we assume that

$$\|\partial_t^s \partial^k u\| \leq C \|\partial^k u\| \qquad \text{for all } 0 \leq s \leq q + 2$$

holds for any spatial derivative abbreviated by ∂^k. This means especially that time derivatives do not introduce negative powers of the perturbation parameter ε.

Remark 2.1. Assumption 5 restricts the applicability of our analysis to problems with data supporting this assumption. Investigations on ε-uniform bounds on time derivatives for a second-order example can be found in [5].

In the subsequent analysis of this paper, let the Assumptions 1–5 be fulfilled.

We have in most applications the relations

$$g_{L^2}(N) \le C g_d(N) \le C g_e(N)$$

for the different bounding functions. We will also use the abbreviation

$$G(N) := g_{L^2}(N) + g_d(N) + g_e(N) + g_s(N)$$

for the final bounding function of the spatial influence which is ε-uniformly bounded itself.

3. Numerical method and analysis

A weak formulation of (1.1) for a right-hand side $f \in L^2(0,T;V')$ reads:

Find $u \in L^2(0,T;V)$ such that $u' \in L^2(0,T;V')$, $u(0) = u_0$, and

$$\int_{I_m} \langle u', v\rangle \,\mathrm{d}t + \int_{I_m} a(t;u,v)\,\mathrm{d}t = \int_{I_m} \langle f, v\rangle \,\mathrm{d}t \quad \text{for all } v \in L^2(0,T;V). \tag{3.1}$$

Note that the initial condition $u(0) = u_0$ is well defined since the function $u \in L^2(0,T;V)$ with $u' \in L^2(0,T;V')$ is continuous in time. Moreover, the solution u of (1.1) solves (3.1). Note that the duality pairing $\langle u', v\rangle$ becomes the L^2-scalar product (u', v) for u being smooth enough in space.

Let M be a positive integer and $0 = t_0 < t_1 < \cdots < t_M = T$. Using these time mesh points, we define intervals $I_m := (t_{m-1}, t_m]$, $m = 1,\dots,M$, with mesh sizes $\tau_m := t_m - t_{m-1}$ and the maximal mesh size $\tau := \max\limits_{m=1,\dots,M} \tau_m$.

For a piecewise smooth time-dependent function φ, we define at $t = t_m$ the one-sided limits $\varphi_m^\pm$ and the jump $[\![\varphi]\!]_m$ by

$$\varphi_m^\pm := \lim_{t\to t_m \pm 0} \varphi(t), \qquad [\![\varphi]\!]_m := \varphi_m^+ - \varphi_m^-.$$

Let q be the polynomial order of our elements in time. We define the discrete function space

$$V_N^\tau := \big\{ W \in L^2(0,T;V_N) \;:\; W|_{I_m} \in \mathcal{P}_q(I_m, V_N),\ 1 \le m \le M \big\},$$

where $\mathcal{P}_q(I_m, V_N)$ denotes the space of V_N-valued polynomials of degree at most q on the time interval I_m.

The dG(q) method reads as follows:

Given $U_0^- \in V_N$ as a suitable approximation of u_0, find $U \in V_N^\tau$ such that

$$\sum_{m=1}^{M} \left\{ \int_{I_m} (U', V)\,\mathrm{d}t + \int_{I_m} a_N(t;U,V)\,\mathrm{d}t + ([\![U]\!]_m, V_m^+) \right\} = \sum_{m=1}^{M} \int_{I_m} \langle f, V\rangle \,\mathrm{d}t.$$

Since the test functions are allowed to be discontinuous at the time points, the global problem decouples into a sequence of local problems on the intervals I_m. Such local problem reads:

Given $U_m^- \in V_N$, find $U|_{I_m} \in \mathcal{P}_q(I_m, V_N)$ such that

$$\int_{I_m} (U', V)\,\mathrm{d}t + \int_{I_m} a_N(t;U,V)\,\mathrm{d}t + ([\![U]\!]_m, V_m^+) = \int_{I_m} \langle f, V\rangle \,\mathrm{d}t,$$

where U_0^- is again a suitable approximation of u_0.

In order to evaluate the time integrals numerically, we use the right-sided Gauß–Radau quadrature formula. To this end, denote by $\hat\omega_i$ and $\hat t_i$, $i = 0,\dots,q$, the weights and nodes of the Gauß–Radau formula with $q+1$ nodes on the reference time interval $\widehat{I}=(-1,1]$. Using the transformation

$$T_m:\widehat{I}\to I_m,\ \hat{t}\mapsto \frac{t_{m-1}+t_m}{2}+\frac{\tau_m}{2}\hat{t},$$

we define by

$$Q_m\left[v\right]:=\frac{\tau_m}{2}\sum_{i=0}^{q}\hat{\omega}_i v(t_{m,i})$$

with the transformed Gauß–Radau points $t_{m,i}:=T_m(\hat{t}_i)$, $i=0,\dots,q$, a quadrature formula on I_m which is exact for polynomials of degree at most $2q$. For the same nodes $t_{m,i}$ we denote by $\varphi_{m,i}$ with $i=0,\dots,q$ the associated Lagrange basis functions and define for a function $v\in C([0,T],V+V_N)$ by

$$(Pv)\big|_{I_m}(t):=\sum_{i=0}^{q}v(t_{m,i})\varphi_{m,i}(t),\qquad m=1,\dots,M, \tag{3.2}$$

an interpolation operator. To complete the definition, we set

$$(Pv)(0)=v(0).$$

We will use in the Lemmas 3.3 and 3.4 an additional interpolation operator which utilises $t_{m,-1}=t_{m-1}$ in addition to $t_{m,i}$, $i=0,\dots,q$. Denoting the associated Lagrange basis functions by $\psi_{m,i}$, $i=-1,0,\dots,q$, this interpolation operator is given by

$$(\widehat{P}v)\big|_{I_m}(t):=\sum_{i=-1}^{q}v(t_{m,i})\psi_{m,i}(t),\qquad m=1,\dots,M. \tag{3.3}$$

Note that $\widehat{P}$ maps to functions which are continuous in time while the image of P is allowed to be discontinuous at the time mesh points.

The interpolation operators P and $\widehat{P}$ provide the error estimates

$$\sup_{t\in I_m}\|(v-Pv)(t)\|\le C\tau_m^{q+1}\sup_{t\in I_m}\|v^{(q+1)}(t)\|,\quad v\in W^{q+1,\infty}(0,T;L^2(\Omega)), \tag{3.4a}$$

$$\sup_{t\in I_m}\|(v-\widehat{P}v)'(t)\|\le C\tau_m^{q+1}\sup_{t\in I_m}\|v^{(q+2)}(t)\|,\quad v\in W^{q+2,\infty}(0,T;L^2(\Omega)), \tag{3.4b}$$

which follow directly from polynomial interpolation theory.

We will use

$$\begin{aligned}\|v\|_\infty &:=\sup_{t\in[0,T]}\|v(t)\|,\\ \|v\|_{\infty,d} &:=\sup_{i=1,\dots,M}\|v(t_i^-)\|,\\ |||v|||_Q &:=\left(\sum_{m=1}^{M}Q_m\left[|||v(t)|||^2\right]\right)^{1/2}\end{aligned} \tag{3.5}$$

in the subsequent analysis and assume in the remainder of this paper that $f\in C(0,T;V')$ and $u\in C^1(0,T;V)$.

The fully discrete problem on I_m then reads:

Given $U_m^-\in V_N$, find $U\in V_N^\tau$ such that

$$Q_m\left[(U',V)\right]+Q_m\left[a_N\big(t;U,V\big)\right]+\big([\![U]\!]_{m-1},V_{m-1}^+\big)=Q_m\left[\langle f,V\rangle\right] \tag{3.6}$$

for all $V \in V_N^\tau$. Note that $U_0^- \in V_N$ is a suitable approximation of the initial condition u_0.

The solution $u \in C^1(0,T;V)$ of (1.1) solves the collocation formulation:

$$Q_m\left[(u',v)\right] + Q_m\left[a(t;u,v)\right] = Q_m\left[\langle f,v\rangle\right] \quad \text{for all } v \in \{v : v(t_{m,i}) \in V\}. \tag{3.7}$$

Lemma 3.1. *Let u and U denote the solutions of* (1.1) *and* (3.6)*. Then the relation*

$$Q_m\left[(U'-u',V)\right] + Q_m\left[a_N(t;U-u,V)\right] + \left([\![U-u]\!]_{m-1}, V_{m-1}^+\right) = Q_m\left[s_N(t,u,V)\right]$$

holds true for all $V \in V_N^\tau$.

Proof. This equality follows directly from the difference of (3.7) and (3.6). For a consistent stabilisation term s_N, the right-hand side is zero and the relation is the Galerkin orthogonality. ☐

Let us split the error $U - u = \xi - \eta$ with $\xi := U - PI_N u$ and $\eta := u - PI_N u$. Then from Lemma 3.1 follows the error equation

$$\begin{aligned} &Q_m\left[(\xi',V)\right] + Q_m\left[a_N(t;\xi,V)\right] + \left([\![\xi]\!]_{m-1}, V_{m-1}^+\right) \\ &\qquad = Q_m\left[(\eta',V)\right] + Q_m\left[a_N(t;\eta,V)\right] + \left([\![\eta]\!]_{m-1}, V_{m-1}^+\right) + Q_m\left[s_N(t,u,V)\right] \end{aligned} \tag{3.8}$$

for all $V \in V_N^\tau$.

In order to estimate the left-hand side of (3.8) from below, we state the following result.

Lemma 3.2. *For $\chi \in V_N^\tau$ it holds that*

$$\begin{aligned} &Q_m\left[(\chi',\chi)\right] + Q_m\left[a_N(t;\chi,\chi)\right] + \left([\![\chi]\!]_{m-1}, \chi_{m-1}^+\right) \\ &\qquad \geq \frac{1}{2}\left(\|\chi_m^-\|^2 - \|\chi_{m-1}^-\|^2 + \|[\![\chi]\!]_{m-1}\|^2\right) + \gamma Q_m\left[|\!|\!|\chi|\!|\!|^2\right]. \end{aligned} \tag{3.9}$$

Proof. Let us start with the first and the last term on the left-hand side of (3.9). Here we obtain

$$\begin{aligned} Q_m\left[(\chi',\chi)\right] + \left([\![\chi]\!]_{m-1}, \chi_{m-1}^+\right) &= \int_{I_m} (\chi',\chi)\,\mathrm{d}t + \left(\chi_{m-1}^+ - \chi_{m-1}^-, \chi_{m-1}^+\right) \\ &= \frac{1}{2}\left(\int_{I_m} \frac{\mathrm{d}}{\mathrm{d}t}\|\chi\|^2\,\mathrm{d}t + \|\chi_{m-1}^+\|^2 - \|\chi_{m-1}^-\|^2 + \|[\![\chi]\!]_{m-1}\|^2\right) \\ &= \frac{1}{2}\left(\|\chi_m^-\|^2 - \|\chi_{m-1}^-\|^2 + \|[\![\chi]\!]_{m-1}\|^2\right), \end{aligned}$$

where the exactness of the quadrature rule Q_m for polynomials up to order $2q$ was used. For the remaining term on the left-hand side of (3.9), it follows that

$$\begin{aligned} Q_m\left[a_N(t;\chi,\chi)\right] &= \frac{\tau_m}{2}\sum_{i=0}^{q} \hat{\omega}_i a_N\big(t_{m,i};\chi(t_{m,i}),\chi(t_{m,i})\big) \\ &\geq \frac{\tau_m}{2}\sum_{i=0}^{q} \hat{\omega}_i \gamma\, |\!|\!|\chi(t_{m,i})|\!|\!|^2 = \gamma Q_m\left[|\!|\!|\chi|\!|\!|^2\right] \end{aligned}$$

from assumption (2.3). ☐

In the next lemma we use $\widehat{P}$ to replace P and obtain a higher order interpolation error.

Lemma 3.3. *We have*

$$Q_m\left[(\eta',V)\right] + \left(\llbracket\eta\rrbracket_{m-1}, V_{m-1}^+\right) = Q_m\left[\left((u-\widehat{P}I_Nu)',V\right)\right]$$

for all $V\in V_N^\tau$.

Proof. Recall that $\eta = u - PI_Nu$ and note that $\widehat{P}I_Nu$ interpolates additionally in t_{m-1} on each time interval; see (3.3). Since $\llbracket u\rrbracket_{m-1}=0$, $m=1,\dots,M-1$, we have to consider only the discrete contributions. Hence, we obtain on each interval

$$\begin{aligned}
Q_m&\left[((PI_Nu)',V)\right] + ((PI_Nu)_{m-1}^+, V_{m-1}^+) - ((PI_Nu)_{m-1}^-, V_{m-1}^+)\\
&= \int_{I_m} ((PI_Nu)',V)\ \mathrm{d}t + ((PI_Nu)_{m-1}^+, V_{m-1}^+) - (I_Nu_{m-1}, V_{m-1}^+)\\
&= -\int_{I_m} (PI_Nu,V')\ \mathrm{d}t + ((PI_Nu)_{m}^-, V_{m}^-) - (I_Nu_{m-1}, V_{m-1}^+)\\
&= -Q_m\left[(PI_Nu,V')\right] + ((PI_Nu)_{m}^-, V_{m}^-) - (I_Nu_{m-1}, V_{m-1}^+)\\
&= -Q_m\left[(\widehat{P}I_Nu,V')\right] + ((\widehat{P}I_Nu)_{m}^-, V_{m}^-) - (I_Nu_{m-1}, V_{m-1}^+)\\
&= -\int_{I_m} (\widehat{P}I_Nu,V')\ \mathrm{d}t + ((\widehat{P}I_Nu)_{m}^-, V_{m}^-) - (I_Nu_{m-1}, V_{m-1}^+)\\
&= \int_{I_m} ((\widehat{P}I_Nu)',V)\ \mathrm{d}t + ((\widehat{P}I_Nu)_{m-1}^+, V_{m-1}^+) - (I_Nu_{m-1}, V_{m-1}^+)\\
&= Q_m\left[\left((\widehat{P}I_Nu)',V\right)\right],
\end{aligned}$$

where we used the definition of the interpolators P and $\widehat{P}$ in $t_{m,i}$ (see (3.2) and (3.3)), integration by parts and the exactness of the quadrature rule for polynomials of degree $2q$ multiple times. ☐

Now let us bound the right-hand side of (3.8).

Lemma 3.4. *Let* $V\in V_N^\tau$ *arbitrary and let* $\alpha,\beta>0$ *be chosen such that* $\alpha\beta=\frac14$. *Provided* $u\in W^{q+2,\infty}\big(0,t;L^2(\Omega)\big)\cap C^1\big(0,T;H^k(\Omega)\big)$, *we have the estimates:*

$$|Q_m\left[a_N(t;\eta,V)\right]| \le C\alpha\tau_m g_d^2(N) + \beta Q_m\left[|||V|||^2\right], \tag{3.10}$$

$$\left|Q_m\left[(\eta',V)\right] + \left(\llbracket\eta\rrbracket_{m-1},V_{m-1}^+\right)\right| \le C\alpha\tau_m(\tau_m^{2q+2}+g_{L^2}^2(N)) + \beta Q_m\left[\|V\|^2\right]. \tag{3.11}$$

If the stabilisation term is not consistent, then the bound

$$|Q_m\left[s_N(t;u,V)\right]| \le C\alpha\tau_m g_s^2(N) + \beta Q_m\left[s_N(t;V,V)\right] \tag{3.12}$$

holds true.

Proof. Let us start with (3.10). Here we have with assumption (2.4):

$$\begin{aligned}
\left|Q_m\left[a_N(t;\eta,V)\right]\right| &= \frac{\tau_m}{2}\left|\sum_{i=0}^{q}\hat{\omega}_i a_N\left(t_{m,i};u(t_{m,i})-PI_Nu(t_{m,i}),V(t_{m,i})\right)\right| \\
&= \frac{\tau_m}{2}\left|\sum_{i=0}^{q}\hat{\omega}_i a_N\left(t_{m,i};u(t_{m,i})-I_Nu(t_{m,i}),V(t_{m,i})\right)\right| \\
&\le C\frac{\tau_m}{2}\sum_{i=0}^{q}\hat{\omega}_i g_d(N)\,|||V(t_{m,i})||| \\
&\le C\alpha\tau_m g_d^2(N) + \beta Q_m\left[|||V|||^2\right],
\end{aligned}$$

where Young's inequality was used with $\alpha\beta = \frac{1}{4}$ in the last step.

In order to prove (3.11), we use Lemma 3.3 and obtain

$$\begin{aligned}
&\left|Q_m\left[(\eta',V)\right] + \left([\![\eta]\!]_{m-1},V_{m-1}^+\right)\right| \\
&\quad\le \frac{\tau_m}{2}\sum_{i=0}^{q}\hat{\omega}_i\left|\left((u-\widehat{P}I_Nu)'(t_{m,i}),V(t_{m,i})\right)\right| \\
&\quad\le \frac{\tau_m}{2}\sum_{i=0}^{q}\hat{\omega}_i\|(u-\widehat{P}u)'(t_{m,i})\|\,\|V(t_{m,i})\| \\
&\qquad+\frac{\tau_m}{2}\sum_{i=0}^{q}\hat{\omega}_i\|(\widehat{P}u)'(t_{m,i})-I_N(\widehat{P}u)'(t_{m,i})\|\,\|V(t_{m,i})\| \\
&\quad\le C\frac{\tau_m}{2}\sum_{i=0}^{q}\hat{\omega}_i\tau_m^{q+1}\|V(t_{m,i})\| + C\frac{\tau_m}{2}\sum_{i=0}^{q}\hat{\omega}_i g_{L^2}(N)\|V(t_{m,i})\| \\
&\quad\le C\alpha\tau_m(\tau_m^{2q+2}+g_{L^2}^2(N)) + \beta Q_m\left[\|V\|^2\right],
\end{aligned}$$

where (3.4b), assumption (2.1), and Young's inequality were used. Note that $(\widehat{P}u)'(t_{m,i})$ can be written as a linear combination of solutions of stationary problems such that assumption (2.1) can be applied.

Now let us come to the consistency error (3.12). Assuming a nonconsistent method, we have with assumptions (2.6) and (2.7):

$$\begin{aligned}
\left|Q_m\left[s_N(t;u,V)\right]\right| &\le Q_m\left[s_N(t;u,u)\right]^{1/2}Q_m\left[s_N(t;V,V)\right]^{1/2} \\
&\le \alpha Q_m\left[s_N(t;u,u)\right] + \beta Q_m\left[s_N(t;V,V)\right] \\
&\le C\alpha\tau_m g_s^2(N) + \beta Q_m\left[s_N(t;V,V)\right],
\end{aligned}$$

where again $\alpha\beta = \frac{1}{4}$ was used. ☐

Recall the definition of $|||v|||_Q$; see (3.5).

Theorem 3.5. *Let $U \in V_N^\tau$ and $u \in W^{q+2,\infty}\big(0,t;L^2(\Omega)\big)\cap C^1\big(0,T;H^k(\Omega)\big)$ be the solutions of* (3.6) *and* (1.1)*, respectively. Then it holds that*

$$\|\xi_M^-\|^2 + \gamma\,|||\xi|||_Q^2 + \sum_{m=1}^{M}\|[\![\xi]\!]_{m-1}\|^2 \le CT\left(\tau^{q+1}+g_d(N)+g_{L^2}(N)+g_s(N)\right)^2$$

for the discrete error $\xi = U - PI_Nu$.

Proof. Let us assume a nonconsistent stabilisation method. Then combining the error representation (3.8) with Lemma 3.2 and the estimates of Lemma 3.4 leads to

$$\begin{aligned}(3.13)\quad \frac{1}{2}\left(\|\xi_m^-\|^2 - \|\xi_{m-1}^-\|^2 + \|[\![\xi]\!]_{m-1}\|^2\right) + \gamma Q_m\left[|||\xi|||^2\right] \\ \leq C\alpha\tau_m\left(g_d(N) + \tau_m^{q+1} + g_{L^2}(N) + g_s(N)\right)^2 \\ + \beta\left(Q_m\left[|||\xi|||^2\right] + Q_m\left[\|\xi\|^2\right] + Q_m\left[s_N(t;\xi,\xi)\right]\right).\end{aligned}$$

In order to absorb the ξ-contribution on the left-hand side, we set $\beta = \gamma/(2(1 + C_1^2 + C_2))$ with C_1, C_2 from Assumption 1 and get

$$\begin{aligned}\|\xi_m^-\|^2 - \|\xi_{m-1}^-\|^2 + \|[\![\xi]\!]_{m-1}\|^2 + \gamma Q_m\left[|||\xi|||^2\right] \\ \leq C\tau_m\left(\tau_m^{q+1} + g_d(N) + g_{L^2}(N) + g_s(N)\right)^2.\end{aligned}$$

If the stabilisation is consistent, the estimate (3.13) holds true without the s_N-term. We have in this case $g_s(N) = 0$ and obtain the same result.
Summation of (3.13) over $m = 1, \ldots, k$ yields

$$\begin{aligned}(3.14)\quad \|\xi_k^-\|^2 + \sum_{m=1}^{k}\left(\gamma Q_m\left[|||\xi|||^2\right] + \|[\![\xi]\!]_{m-1}\|^2\right) \\ \leq \|\xi_0^-\|^2 + C\sum_{m=1}^{k}\tau_m\left(\tau_m^{q+1} + g_d(N) + g_{L^2}(N) + g_s(N)\right)^2.\end{aligned}$$

With $\xi_0^- = 0$, $\tau_m^{q+1} \leq \tau^{q+1}$, and $\sum_{m=1}^{M}\tau_m = T$, the proof is done upon setting $k = M$. ☐

Let us now come to the interpolation error.

Lemma 3.6. *We have the bounds*

$$\begin{aligned}\|(u - I_N P u)_M^-\| &\leq C g_{L^2}(N), \\ \|(u - I_N P u)(t)\| &\leq C(\tau_m^{q+1} + g_{L^2}(N)), \quad t \in I_m, \\ |||u - I_N P u|||_Q^2 &\leq C T g_e^2(N), \\ \sum_{m=1}^{M}\|[\![u - I_N P u]\!]_{m-1}\|^2 &\leq C\tau^{2q+1}\end{aligned}$$

for the interpolation error of $u \in W^{q+1,\infty}\big(0,t;L^2(\Omega)\big) \cap C^1\big(0,T;H^k(\Omega)\big)$.

Proof. The first estimate is a direct consequence of assumption (2.1). Indeed, we have

$$\|(u - I_N P u)_M^-\| = \|(u - I_N u)_M^-\| \leq C g_{L^2}(N).$$

The second estimate follows with the triangle inequality and assumption (2.1)

$$\|(u - I_N P u)(t)\| \leq \|(u - Pu)(t)\| + \|(Pu - I_N P u)(t)\| \leq C(\tau^{q+1} + g_{L^2}(N)),$$

where we have used that $(Pu)(t)$ can be written as a linear combination of stationary solutions. With assumption (2.2) we have

$$Q_m\left[|||u - I_N P u|||^2\right] = \frac{\tau_m}{2}\sum_{i=0}^{q}\hat{\omega}_i\,|||u(t_{m,i}) - I_n u(t_{m,i})|||^2 \leq C\tau_m g_e^2(N),$$

which upon summation over m yields the third estimate.

Finally, we have for $\chi = I_N u - PI_N u$:

$$-\int_0^T (\chi, \chi')\,\mathrm{d}t = -\sum_{m=1}^M \int_{t_{m-1}}^{t_m} (\chi, \chi')\,\mathrm{d}t = -\frac{1}{2}\sum_{m=1}^M (\|\chi_m^-\|^2 - \|\chi_{m-1}^+\|^2)$$
$$= \frac{1}{2}\sum_{m=1}^M \|[\![\chi]\!]_{m-1}\|^2,$$

where $\chi_m^- = 0$, $m = 0, \ldots, M$, was used several times. Thus, we obtain

$$\sum_{m=1}^M \|[\![u - I_N P u]\!]_{m-1}\|^2 \le 2\sum_{m=1}^M \|[\![u - I_N u]\!]_{m-1}\|^2 + 2\sum_{m=1}^M \|[\![I_N u - PI_N u]\!]_{m-1}\|^2$$
$$= 2\sum_{m=1}^M \|[\![I_N u - PI_N u]\!]_{m-1}\|^2 = -4\int_0^T (\chi, \chi')\,\mathrm{d}t$$
$$\le C\tau^{2q+1},$$

because $u - I_N u$ is continuous at t_{m-1}, $m = 1, \ldots, M-1$, and χ is an interpolation error in time. ☐

A direct consequence of Theorem 3.5 and Lemma 3.6 is the following theorem.

Theorem 3.7. *Let $U \in V_N^\tau$ be the discrete solution of* (3.6) *and*

$$u \in W^{q+2,\infty}\big(0, t; L^2(\Omega)\big) \cap C^1\big(0, T; H^k(\Omega)\big)$$

be the solution of (1.1). *Then the error $u - U$ can be bounded uniformly in ε by*

$$\|(u - U)(T)\|^2 + \gamma\,|||u - U|||_Q^2 \le CT\left(\tau^{q+1} + G(N)\right)^2,$$
$$\sum_{m=1}^M \|[\![u - U]\!]_{m-1}\|^2 \le CT\left(\tau^{q+1/2} + G(N)\right)^2.$$

Remark 3.8. Note that in the case of supercloseness of the spatial method w.r.t. I_N, the space-time method inherits this property. In these cases the rate $g_d(N)$ is better than the rate $g_e(N)$. Examples for such methods can be found in [6, 7, 9–12, 14, 17, 20, 21, 25, 34]. Thus, Theorem 3.5 predicts a higher rate of convergence with respect to the spatial discretisation for the discrete error than Theorem 3.7 for the actual error. Such a property can be used to improve the numerical solution by means of a local interpolation into a higher order polynomial space on a macro-mesh (in space); see, e.g., [8, 27] for more details. The resulting postprocessed numerical solution is then superconvergent in space.

The result of Theorem 3.7 can be improved. Recall for this the definition of $\|v\|_\infty$; see (3.5).

Theorem 3.9. *Let $U \in V_N^\tau$ be the discrete solution of* (3.6) *and let $u \in W^{q+2,\infty}\big(0, t; L^2(\Omega)\big) \cap C^1\big(0, T; H^k(\Omega)\big)$ be the solution of* (1.1). *Then the error $u - U$ can be bounded uniformly in ε by*

$$\|u - U\|_\infty^2 \le CT\left(\tau^{q+1} + G(N)\right)^2.$$

Proof. The proof imitates the one of [30]. Let us define for $\xi = U - PI_N u \in V_N^\tau$ the auxiliary function

$$\varphi|_{I_m} := P\left(\frac{\tau_m}{t - t_{m-1}}\xi\right)$$

which yields

$$\varphi(t_{m,i}) = \frac{\tau_m}{t_{m,i} - t_{m-1}} \xi(t_{m,i}).$$

Using [30, Lemma 5] we have

$$Q_m \left[(\xi', 2\varphi) \right] + (\xi_{m-1}^+, 2\varphi_{m-1}^+) \geq \frac{1}{\tau_m} Q_m \left[\|\varphi\|^2 \right]. \tag{3.15}$$

Furthermore, we have

$$\begin{aligned}
0 \leq \gamma Q_m \left[|||\xi|||^2 \right] &= \gamma \frac{\tau_m}{2} \sum_{i=0}^{q} \hat{\omega}_i \, |||\xi(t_{m,i})|||^2 \\
&\leq \frac{\tau_m}{2} \sum_{i=0}^{q} \hat{\omega}_i a_N \big(t_{m,i}; \xi(t_{m,i}), \xi(t_{m,i}) \big) \\
&\leq \frac{\tau_m}{2} \sum_{i=0}^{q} \hat{\omega}_i \frac{2\tau_m}{t_{m,i} - t_{m-1}} a_N \big(t_{m,i}; \xi(t_{m,i}), \xi(t_{m,i}) \big) \\
&= \frac{\tau_m}{2} \sum_{i=0}^{q} \hat{\omega}_i a_N \big(t_{m,i}; \xi(t_{m,i}), 2\varphi(t_{m,i}) \big) = Q_m \left[a_N(t; \xi, 2\varphi) \right],
\end{aligned}$$

where $2\tau_m/(t_{m,i} - t_{m-1}) \geq 2$ was used. With ξ being a discrete function, we have the norm equivalence

$$\sup_{t \in I_m} \|\xi(t)\|^2 \leq K \frac{1}{\tau_m} Q_m \left[\|\xi\|^2 \right] \leq K \frac{1}{\tau_m} Q_m \left[\|\varphi\|^2 \right]$$

with a constant $K > 0$ independent of τ_m.

Now, with (3.15) it holds that

$$\begin{aligned}
\sup_{t \in I_m} \|\xi(t)\|^2 + K\gamma Q_m \left[|||\xi|||^2 \right] &\leq K \left(\frac{1}{\tau_m} Q_m \left[\|\varphi\|^2 \right] + Q_m \left[a_N(t; \xi, 2\varphi) \right] \right) \\
&\leq K \Big(Q_m \left[(\xi', 2\varphi) \right] + Q_m \left[a_N(t; \xi, 2\varphi) \right] \\
&\qquad + (\xi_{m-1}^-, 2\varphi_{m-1}^+) + \big([\![\xi]\!]_{m-1}, 2\varphi_{m-1}^+ \big) \Big).
\end{aligned}$$

Using the error equation (3.8) and Lemma 3.4, we continue estimating

$$\begin{aligned}
&\sup_{t \in I_m} \|\xi(t)\|^2 + K\gamma Q_m \left[|||\xi|||^2 \right] \\
&\leq K \Big(Q_m \left[(\eta', 2\varphi) \right] + Q_m \left[a_N(t; \eta, 2\varphi) \right] + \big([\![\eta]\!]_{m-1}, 2\varphi_{m-1}^+ \big) \\
&\qquad + Q_m \left[s_N(t; u, 2\varphi) \right] + (\xi_{m-1}^-, 2\varphi_{m-1}^+) \Big) \\
&\leq C \left(\alpha_1 \tau_m \big(\tau_m^{2q+2} + G^2(N) \big) + \alpha_2 \|\xi_{m-1}^-\|^2 \right) + \beta_1 Q_m \left[|||2\varphi|||^2 \right] + \beta_2 \|2\varphi_{m-1}^+\|^2,
\end{aligned}$$

where Young's inequality with $\alpha_i \beta_i = \frac{1}{4}$, $i = 1, 2$, was used. With

$$\|2\varphi_{m-1}^+\|^2 \leq L \sup_{t \in I_m} \|\xi(t)\|^2 \quad \text{and} \quad Q_m \left[|||2\varphi|||^2 \right] \leq L Q_m \left[|||\xi|||^2 \right],$$

where L is a constant depending only on the polynomial degree q, we obtain

$$\sup_{t \in I_m} \|\xi(t)\|^2 \leq C \big(\|\xi_{m-1}^-\|^2 + \tau_m (\tau_m^{2q+2} + G^2(N)) \big)$$

by setting $\beta_1 = \frac{K\gamma}{L}$ and $\beta_2 = \frac{1}{2L}$. Using (3.14), we have

$$\|\xi_{m-1}^-\|^2 \le C \sum_{k=1}^{m} \tau_k \big(\tau_k^{2q+2} + G^2(N)\big)$$

which results in

$$\|\xi\|_\infty^2 = \sup_{t\in[0,T]} \|\xi(t)\|^2 = \max_{m=1,\dots,M} \sup_{t\in I_m} \|\xi(t)\|^2 \le C \sum_{m=1}^{M} \tau_m \big(\tau_m^{2q+2} + G^2(N)\big).$$

The proof is completed upon exploiting the error estimate for $\|\eta\| = \|u - I_N Pu\|$ of Lemma 3.6. ☐

4. Applications of the general approach

Let us look at some applications of our general analysis. We will give in each case:

- the differential equation, the domain Ω, and the space V;
- the type of mesh used in literature;
- the convergence bound $G(N)$ in the stationary case.

Then with Theorems 3.7 and 3.9 we have

$$\|u - U\|_\infty + \gamma \, |||u - U|||_Q \le C(\tau^{q+1} + G(N)).$$

The polynomial spaces on Ω will in general be piecewise polynomials of order p. We denote in this section by $\bar{u}$ the solution of the stationary problem

$$\begin{cases} L\bar{u} = f & \text{in } \Omega, \\ B\bar{u} = 0 & \text{on } \partial\Omega. \end{cases}$$

4.1. **Reaction-diffusion problems.** Let the problem be given by

$$\begin{aligned} \partial_t u - \varepsilon\Delta u + cu &= f && \text{in } \Omega \times (0,T), \\ u &= 0 && \text{on } \partial\Omega \times (0,T), \\ u(0) &= u_0 && \text{in } \Omega \end{aligned}$$

on a domain $\Omega \subset \mathbb{R}^2$, where $c \ge c_0 > 0$ uniformly in $\overline{\Omega} \times [0,T]$. Thus, the bilinear form associated with the stationary differential equation is given for fixed t by

$$a(t; u, v) := \varepsilon(\nabla u, \nabla v) + (c(t)u, v),$$

the suitable solution space is $V = H_0^1(\Omega)$ and $|||v|||^2 = \varepsilon\|\nabla v\|^2 + c_0\|v\|^2$. The solution u has along the whole boundary of Ω boundary layers which are of characteristic type. Therefore, a layer-adapted mesh has to be refined along all boundaries.

Assuming $\partial\Omega$ being smooth and $\bar{u} \in H^{p+1}(\Omega)$, [31] gives convergence results on standard Shishkin meshes for higher-order elements and the Galerkin method. The results can easily be extended to S-type meshes and give

$$G(N) = (h + N^{-1} \max |\psi'|)^p,$$

where ψ is a mesh-characterising function (see [24]), and h the maximal mesh size orthogonal to the boundary in the layer region. Using standard assumptions on the solution decomposition, the proof can also be given on the unit square. In [1], so-called A-meshes are used on the unit square. They are constructed similar to

S-type meshes with the exception that the transition points use $|\ln\varepsilon|$ instead of $\ln N$. Then we have for $\bar{u} \in H^{p+1}(\Omega)$ the bound

$$G(N) = N^{-p}(\varepsilon^{1/2}|\ln\varepsilon|^{p+1/2} + N^{-1}).$$

Note that in this case G is *not* independent of ε, but ε-uniformly bounded.

4.2. **Convection-diffusion problems.** In this subsection we consider the singularly perturbed convection-diffusion problem given on $\Omega = (0,1)^2$ by

$$\begin{aligned} \partial_t u - \varepsilon\Delta u - b\cdot\nabla u + cu &= f \quad && \text{in } \Omega\times(0,T),\\ u &= 0 && \text{on } \partial\Omega\times(0,T),\\ u(0) &= u_0 && \text{in } \Omega, \end{aligned}$$

where $c + \frac{1}{2}\mathrm{div}\, b \geq c_0 > 0$ holds uniformly in $\overline{\Omega}\times[0,T]$. Then the bilinear form associated with the stationary differential equation is given by

$$a(t;u,v) := \varepsilon(\nabla u, \nabla v) + (c(t)u - b(t)\cdot\nabla u, v),$$

the suitable solution space is $V = H_0^1(\Omega)$ and $|||v|||^2 = \varepsilon\|\nabla v\|^2 + c_0\|v\|^2$ for standard Galerkin analysis. Depending on the components of $b = (b_1, b_2)$ the solution has different layer structures. Assuming $b \geq (\beta_1, \beta_2) > (0,0)$, we expect exponential outflow boundary layers along $x = 0$ and $y = 0$. If $b = (b_1, 0)$ with $b_1 > \beta_1 > 0$, then there is still an exponential outflow layer at $x = 0$, but in addition there are characteristic layers along $y = 0$ and $y = 1$. Furthermore, we have in both cases corner layers as interaction of the two meeting layers.

In [13, 28], the convergence of the standard Galerkin method on Shishkin meshes for higher order FEM in the stationary case was considered for each case. The error bound

$$G(N) = (h + N^{-1}\max|\psi'|)^p$$

for $\bar{u} \in H^{p+1}(\Omega)$ can be found therein or be obtained with standard techniques.

For convection-diffusion problems many stabilisation methods are known. The norm $|||\cdot|||$ includes in each case additional terms controlling the error further. The local projection stabilisation LPS was analysed in [12, 13, 22, 23]. Under restrictions on the stabilisation parameters, the above bound $G(N)$ is given. Note that LPS is not consistent, but it fulfills Assumption 4. Other stabilisation methods can also be included by slight changes. The continuous interior penalty method CIP [2–4, 7, 32] can be treated by setting $V = H_0^1(\Omega)\cap H^2(\Omega)$ because then it is consistent. For including discontinuous Galerkin methods in space (see, e.g., [25, 32, 33]), we have to modify the norm to

$$|||v|||^2 = \varepsilon\sum_{K\in T^N}\|\nabla v\|_K^2 + \|v\|^2 + \text{dG-terms}.$$

Concerning supercloseness for both layer cases, there are results for bilinear ($p = 1$) elements [10, 17, 34] with a bound

$$g_d = (h + N^{-1}\max|\psi'|)^2$$

for the standard Galerkin method and $\bar{u} \in H^3(\Omega)$. For higher order elements and exponential layers, we have for the Galerkin method on Shishkin meshes

$$g_d = N^{-(p+1/4)}$$

for odd polynomial degrees $p \geq 3$ and $\bar{u} \in H^{p+2}(\Omega)$; see [14]. Nevertheless, numerical studies [8] indicate a full order improvement for Galerkin and $p \geq 3$.

4.3. **Fourth-order problems.** Let us consider the singularly perturbed clamped-plate problem

$$\begin{aligned} \partial_t u + \varepsilon^2 \Delta^2 u - \Delta u + cu &= f && \text{in } \Omega \times (0,T), \\ \frac{\partial u}{\partial n} = u &= 0 && \text{on } \partial\Omega \times (0,T), \\ u(0) &= u_0 && \text{in } \Omega, \end{aligned}$$

where $c \geq \underline{c}^2 > 0$ holds uniformly in $\overline{\Omega} \times [0,T]$. With $V = H_0^2(\Omega)$ and

$$(D^2 u, D^2 v) = (u_{xx}, v_{xx}) + 2(u_{xy}, v_{xy}) + (u_{yy}, v_{yy}),$$

the bilinear form associated with the stationary problem is given by

$$a(t; u, v) := \varepsilon^2 (D^2 u, D^2 v) + (\nabla u, \nabla v) + (c(t)u, v);$$

see [15]. A conforming finite element method needs C^1-elements which are costly. Thus in [15], a nonconforming method using standard $\mathcal{Q}_p$-elements on a Shishkin mesh and a continuous interior penalty method (CIP) is applied. In the corresponding norm

$$|||v|||^2 = \varepsilon^2 \sum_{K \in T^N} \|D^2 v\|_K^2 + \|\nabla v\|^2 + \|\underline{c} v\|^2 + \text{CIP-terms},$$

an error estimate with the bound

$$G(N) = \varepsilon^{1/2} (N^{-1} \ln N)^{p-1} + N^{-p} \tag{4.1}$$

for $p \geq 2$ and $\bar{u} \in H^{p+1}(\Omega)$ is given where again G is ε-uniformly bounded.

5. Numerical examples

Let us apply in 2d a standard Galerkin FEM with $\mathcal{Q}_p$-elements on a Bakhvalov–Shishkin mesh with N cells in each dimension for the spatial approximation; see [24] for a precise definition. The expected convergence rates on such meshes are

$$\|u - U\|_\infty + \gamma\, |||u - U|||_Q \leq C(\tau^{q+1} + N^{-r}),$$

where r depends on the polynomial order p and the considered problem. Furthermore, ε is assumed to be small enough to hold $h \leq CN^{-1}$. We use an equidistant mesh with M intervals in time, thus $\tau = T/M$. Assuming a constant ratio between N and M, we obtain a balancing of the two error components for

$$r = q + 1,$$

which gives a condition on the polynomial degree p. For nonsingularly-perturbed problems, a superconvergence result at the endpoints of the intervals is known [29, Ch. 12]. Thus we also check the error in the discrete maximum norm in time

$$\|u - U\|_{\infty,d} \leq C(\tau^{2q+1} + N^{-r}).$$

Here a balancing occurs for

$$r = 2q + 1.$$

All computations were performed in $\mathbb{SOFE}$, a Matlab-FEM suite initiated by Lars Ludwig [19]. We look for a solution in $t \in [0, 2]$ and set the perturbation parameter ε to 10^{-6}.

TABLE 1. Convergence results for several pairs (p, q) for Example 1.

p / q	N	M	$\|u-U\|_\infty$		$\|\|\|u-U\|\|\|_Q$	
2 / 1	16	8	3.739e-02	1.88	3.413e-02	1.98
	32	16	1.016e-02	1.97	8.682e-03	1.99
	64	32	2.600e-03	1.99	2.184e-03	2.00
	128	64	6.543e-04		5.473e-04	
3 / 2	16	8	3.111e-03	2.97	2.219e-03	2.99
	32	16	3.959e-04	3.03	2.793e-04	2.99
	64	32	4.854e-05	2.99	3.504e-05	3.00
	128	64	6.117e-06		4.389e-06	
4 / 3	16	8	1.666e-04	3.89	1.125e-04	3.97
	32	16	1.126e-05	3.97	7.189e-06	3.98
	64	32	7.176e-07	3.99	4.548e-07	3.99
	128	64	4.507e-08		2.861e-08	

TABLE 2. Superconvergence for several pairs (p, q) for Example 1.

p / q	N	M	$\|u-U\|_\infty$		$\|u-U\|_{\infty,d}$	
3 / 1	16	8	3.737e-02	1.88	2.614e-03	2.79
	32	16	1.016e-02	1.97	3.774e-04	2.86
	64	32	2.600e-03	1.99	5.181e-05	2.94
	128	64	6.543e-04		6.741e-06	
5 / 2	16	8	3.113e-03	2.98	8.446e-05	4.71
	32	16	3.959e-04	3.03	3.231e-06	4.92
	64	32	4.854e-05	2.99	1.070e-07	4.97
	128	64	6.117e-06		3.425e-09	

Example 1: We consider the two-dimensional convection-reaction-diffusion problem

$$\partial_t u - \varepsilon \Delta u - \begin{pmatrix} (1+x)(2+\cos(\pi t)) \\ 2+y+t \end{pmatrix} \cdot \nabla u + u = f$$

in $\Omega = (0,1)^2$ with $u = 0$ on $\partial\Omega$ and f chosen such that

$$u = \left(\cos(\pi x/2) - \frac{e^{-x/\varepsilon} - e^{-1/\varepsilon}}{1 - e^{-1/\varepsilon}}\right)\left((1-y) - \frac{e^{-2y/\varepsilon} - e^{-2/\varepsilon}}{1 - e^{-2/\varepsilon}}\right)\cos(\pi t)$$

is the solution. Here we have

$$|||v|||^2 = \varepsilon \|\nabla v\|^2 + \|v\|^2 \quad \text{and} \quad r = p.$$

Tables 1 and 2 give the computed errors together with the computed rates of convergence for several pairs of polynomial degrees. Clearly we see in Table 1 the predicted rates $p = q + 1$ of Theorems 3.7 and 3.9. Furthermore, we observe in Table 2 for the discrete error in the nodes t_i^- an improved order of $p = 2q + 1$, supporting the assumption on superconvergence in the end-points.

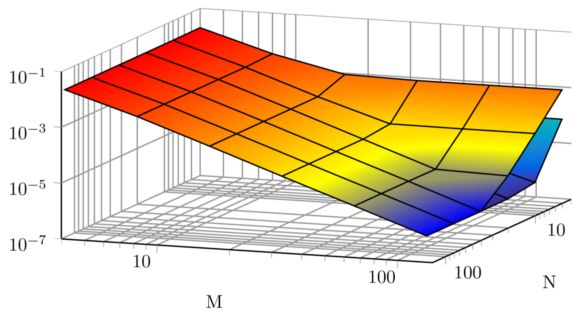

FIGURE 1. $\|u-U\|_\infty$ over N and M for $q=2$ and $p=3$ (upper surface) and $p=5$ (lower surface).

Comparing the L^∞-L^2 errors of Tables 1 and 2 belonging to the same polynomial order in time, it is noticeable that they are almost identical. Figure 1 shows these errors as surfaces over N and M for the temporal polynomial degree $q=2$ and spatial polynomial degrees $p=3$ and $p=5$. As one can see, the temporal error dominates the total error behaviour and an improvement of spatial accuracy does not pay off.

Example 2: We look at the fourth-order problem following [15],

$$\partial_t u + \varepsilon^2\Delta^2 u - \Delta u = f$$

in $\Omega=(0,1)^2$, $u=\partial_n u = 0$ on $\partial\Omega$, where f and u_0 are chosen such that

$$\begin{aligned}
u(x,y,t) =& \frac{1}{2}\left(\sin(\pi x) + \frac{\pi\varepsilon}{1-\mathrm{e}^{-1/\varepsilon}}\left(\mathrm{e}^{-x/\varepsilon}+\mathrm{e}^{(x-1)/\varepsilon}-1-\mathrm{e}^{-1/\varepsilon}\right)\right)\\
&\cdot\left(2y(1-y^2)+\varepsilon\left(\ell d(1-2y)-3\frac{s}{\ell}+\left(\frac{3}{\ell}-d\right)\mathrm{e}^{-y/\varepsilon}\right.\right.\\
&\left.\left.+\left(\frac{3}{\ell}+d\right)\mathrm{e}^{(y-1)/\varepsilon}\right)\right)\cdot(1+2t\cos(2\pi t))
\end{aligned}$$

with $\ell = 1-\mathrm{e}^{-1/\varepsilon}$, $s=2-\ell$ and $d=1/(s-2\varepsilon\ell)$ is the solution. Here we have

$$|||v|||^2 = \varepsilon^2\left(\sum_{K\in T^N}\|D^2 v\|_K^2 + \sum_{e\in\mathcal{E}}\sigma_e\|[\![\partial_n v]\!]\|_{L^2(e)}^2\right)+\|\nabla v\|^2,$$

with the mesh-dependent constants σ_e as given in [15]. For $\varepsilon\le CN^{-2}$ we expect $r=p$ while for large ε the rate decreases to $r=p-1$.

The two error components behave slightly different. We observe for $\|u-U\|_\infty$ a behaviour comparable to Figure 1 with errors in the range 10^{-5} to 10^{-1}. Therefore, we do not present the results here.

Figure 2 shows the error $|||u-U|||_Q^2$ as surface over N and M. There are three regions visible. For large N and small M the time error dominates and converges with an order of approximately $3=q+1$. For large M and small N the spatial error dominates and converges with order approximately $3=p$. For large N and M there is almost a stagnation of the error. This corresponds very well with the behaviour of the stationary problem and the given error bound $G(N)$ in (4.1). For

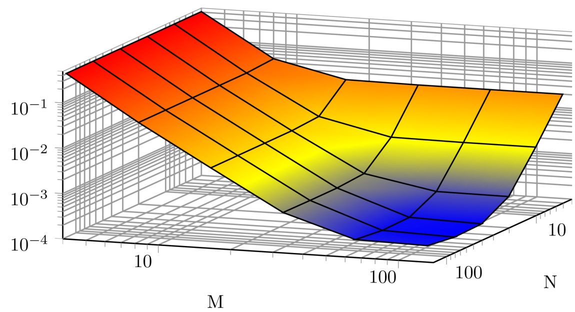

FIGURE 2. $\|u-U\|_Q$ over N and M for $q=2$ and $p=3$.

a further refinement in space we expect a convergence of order $2 = p-1$ as here N^{-1} is getting small compared with $\varepsilon^{1/2} = 10^{-3}$.

Example 3: As a final example we come back to the second-order convection-reaction-diffusion problem

$$\partial_t u - \varepsilon\Delta u - \begin{pmatrix}(1+x)(2+\cos(\pi t))\\ 2+y+t\end{pmatrix}\cdot\nabla u + u = f$$

in $\Omega = (0,1)^2$ with $u=0$ on $\partial\Omega$ considered already in Example 1, but this time we do not force a given solution. Instead we prescribe the initial data $u(x,y,0)=0$ for $(x,y)\in\Omega$ and the right-hand side

$$f(x,y,t) = -4(2xy-x-y)(2xy-x-y+1)\sin(t)$$

which is zero in the corners of the domain and at $t=0$. Thus, f fulfills for each fixed t the first compatibility conditions given in [18] and we can expect the solution u not to have strong corner singularities. Nevertheless, it is an open question whether Assumption 5 will be fulfilled for this example.

We chose this example as here it makes sense for the numerical procedure to change the mesh over time in order to capture the varying width of the boundary layers. The analysis of the dG-method presented in this article holds also in this setting. In order to handle different discrete spaces in time, the space V_N^τ is replaced by

$$\{W\in L^2(0,T;V)\,:\,W|_{I_m}\in\mathcal{P}_q(I_m,V_N^m\},$$

where V_N^m, $m=1,\ldots,M$, may differ in the meshes of the spatial discretisation. However, if all of these spaces V_N^m have similar approximation properties and the jump contribution from the previous time step is properly implemented, the theoretical results still hold.

In our numerical experiment we do not know the exact solution. Therefore, we use the concept of a numerical reference solution u_{ref} to replace the exact solution u in the computation of errors. In our case, the reference solution has polynomial degrees $p=3$, $q=2$ and is computed on a mesh with $N=128$ and $M=64$. For the numerical solution we use the polynomial pair $p=2$, $q=1$ and compute the solutions on a fixed mesh and a time-adapted mesh, where the numerical layer width is determined using

$$\beta_1^m = \inf_{t\in I_m}\inf_{(x,y)\in\Omega} b_1(x,y,t),\quad \beta_2^m = \inf_{t\in I_m}\inf_{(x,y)\in\Omega} b_2(x,y,t)$$

TABLE 3. Comparison of convergence results for Example 3.

N	M	$\|u_{ref} - U\|_\infty$		$\|\|\|u_{ref} - U\|\|\|_Q$	
fixed mesh, $\varepsilon = 10^{-3}$					
16	8	3.383e-03	1.37	6.149e-03	1.88
32	16	1.307e-03	1.54	1.668e-03	1.95
64	32	4.494e-04	1.81	4.322e-04	1.98
128	64	1.280e-04		1.095e-04	
fixed mesh, $\varepsilon = 10^{-6}$					
16	8	3.393e-03	1.37	6.156e-03	1.88
32	16	1.309e-03	1.54	1.670e-03	1.95
64	32	4.500e-04	1.81	4.330e-04	1.98
128	64	1.281e-04		1.099e-04	
fixed mesh, $\varepsilon = 10^{-8}$					
16	8	3.393e-03	1.37	6.156e-03	1.88
32	16	1.309e-03	1.54	1.670e-03	1.95
64	32	4.500e-04	1.81	4.330e-04	1.98
128	64	1.281e-04		1.099e-04	
time-adapted mesh, $\varepsilon = 10^{-6}$					
16	8	3.396e-03	1.37	4.250e-03	1.90
32	16	1.309e-03	1.54	1.137e-03	1.96
64	32	4.500e-04	1.81	2.932e-04	1.98
128	64	1.281e-04		7.434e-05	

instead of

$$\beta_1 = \inf_{t\in[0,T],(x,y)\in\Omega} b_1(x,y,t), \quad \beta_2 = \inf_{t\in[0,T],(x,y)\in\Omega} b_2(x,y,t).$$

Table 3 shows the various results obtained. We can observe uniformity in ε of the numerical results, although Assumption 5 might be violated, by comparing the first, second and third block of results on the fixed meshes. All show an almost identical behaviour and the $|||\cdot|||_Q$-component of the error shows second-order convergence like in Example 1. But here the other norm-component is slightly worse, only converging toward second-order. This might be due to f not fulfilling all compatibility conditions, such that a solution decomposition exists supporting second-order convergence. Another possibility might lay in the usage of a reference solution instead of the (not available) exact solution.

Comparing the second and the fourth block of results we see for the same value of $\varepsilon = 10^{-6}$ the results for the two choices of mesh adaptation. Indeed, the adaptation of the meshes in time yields a slightly better result, visible in the $|||\cdot|||_Q$-component of the error that is better by a factor of about 0.75. The second error component is unchanged by the treatment of meshes. Again this might be due to problems discussed above.

References

[1] T. Apel and G. Lube, *Anisotropic mesh refinement for a singularly perturbed reaction diffusion model problem*, Appl. Numer. Math. **26** (1998), no. 4, 415–433, DOI 10.1016/S0168-9274(97)00106-2. MR1612364

[2] E. Burman, *A unified analysis for conforming and nonconforming stabilized finite element methods using interior penalty*, SIAM J. Numer. Anal. **43** (2005), no. 5, 2012–2033, DOI 10.1137/S0036142903437374. MR2192329

[3] E. Burman, J. Guzmán, and D. Leykekhman, *Weighted error estimates of the continuous interior penalty method for singularly perturbed problems*, IMA J. Numer. Anal. **29** (2009), no. 2, 284–314, DOI 10.1093/imanum/drn001. MR2491428

[4] E. Burman and P. Hansbo, *Edge stabilization for Galerkin approximations of convection-diffusion-reaction problems*, Comput. Methods Appl. Mech. Engrg. **193** (2004), no. 15-16, 1437–1453, DOI 10.1016/j.cma.2003.12.032. MR2068903

[5] C. Clavero, J. C. Jorge, F. Lisbona, and G. I. Shishkin, *A fractional step method on a special mesh for the resolution of multidimensional evolutionary convection-diffusion problems*, Appl. Numer. Math. **27** (1998), no. 3, 211–231, DOI 10.1016/S0168-9274(98)00014-2. MR1634346

[6] R. G. Durán, A. L. Lombardi, and M. I. Prieto, *Supercloseness on graded meshes for $\mathcal{Q}_1$ finite element approximation of a reaction-diffusion equation*, J. Comput. Appl. Math. **242** (2013), 232–247, DOI 10.1016/j.cam.2012.10.004. MR2997439

[7] S. Franz, *Continuous interior penalty method on a Shishkin mesh for convection-diffusion problems with characteristic boundary layers*, Comput. Methods Appl. Mech. Engrg. **197** (2008), no. 45-48, 3679–3686, DOI 10.1016/j.cma.2008.02.019. MR2458107

[8] S. Franz, *Convergence phenomena of $\mathcal{Q}_p$-elements for convection-diffusion problems*, Numer. Methods Partial Differential Equations **29** (2013), no. 1, 280–296, DOI 10.1002/num.21709. MR3003111

[9] S. Franz, *Superconvergence using pointwise interpolation in convection-diffusion problems*, Appl. Numer. Math. **76** (2014), 132–144, DOI 10.1016/j.apnum.2013.07.007. MR3131867

[10] S. Franz and T. Linß, *Superconvergence analysis of the Galerkin FEM for a singularly perturbed convection-diffusion problem with characteristic layers*, Numer. Methods Partial Differential Equations **24** (2008), no. 1, 144–164, DOI 10.1002/num.20245. MR2371352

[11] S. Franz, T. Linß, H.-G. Roos, and S. Schiller, *Uniform superconvergence of a finite element method with edge stabilization for convection-diffusion problems*, J. Comput. Math. **28** (2010), no. 1, 32–44, DOI 10.4208/jcm.2009.09-m1005. MR2603579

[12] S. Franz and G. Matthies, *Local projection stabilisation on S-type meshes for convection-diffusion problems with characteristic layers*, Computing **87** (2010), no. 3-4, 135–167, DOI 10.1007/s00607-010-0079-y. MR2640011

[13] S. Franz and G. Matthies, *Convergence on layer-adapted meshes and anisotropic interpolation error estimates of non-standard higher order finite elements*, Appl. Numer. Math. **61** (2011), no. 6, 723–737, DOI 10.1016/j.apnum.2011.02.001. MR2774865

[14] S. Franz and H.-G. Roos, *Superconvergence of a Galerkin FEM for higher-order elements in convection-diffusion problems*, Numer. Math. Theory Methods Appl. **7** (2014), no. 3, 356–373. MR3250058

[15] S. Franz, H.-G. Roos, and A. Wachtel, *A* C^0 *interior penalty method for a singularly-perturbed fourth-order elliptic problem on a layer-adapted mesh*, Numer. Methods Partial Differential Equations **30** (2014), no. 3, 838–861, DOI 10.1002/num.21839. MR3190340

[16] L. Kaland and H.-G. Roos, *Parabolic singularly perturbed problems with exponential layers: robust discretizations using finite elements in space on Shishkin meshes*, Int. J. Numer. Anal. Model. **7** (2010), no. 3, 593–606. MR2644293

[17] T. Linß, *Uniform superconvergence of a Galerkin finite element method on Shishkin-type meshes*, Numer. Methods Partial Differential Equations **16** (2000), no. 5, 426–440, DOI 10.1002/1098-2426(200009)16:5⟨426::AID-NUM2⟩3.3.CO;2-I. MR1778398

[18] T. Linß and M. Stynes, *Asymptotic analysis and Shishkin-type decomposition for an elliptic convection-diffusion problem*, J. Math. Anal. Appl. **261** (2001), no. 2, 604–632, DOI 10.1006/jmaa.2001.7550. MR1853059

[19] L. Ludwig, $\mathbb{SOFE}$, `https://github.com/SOFE-Developers/SOFE`, last accessed: 03/21/17.

[20] L. Ludwig and H.-G. Roos, *Finite element superconvergence on Shishkin meshes for convection-diffusion problems with corner singularities*, IMA J. Numer. Anal. **34** (2014), no. 2, 782–799, DOI 10.1093/imanum/drt027. MR3194808

[21] L. Ludwig and H.-G. Roos, *Convergence and supercloseness of a finite element method for a singularly perturbed convection-diffusion problem on an L-shaped domain*, IMA J. Numer. Anal. **36** (2016), no. 3, 1261–1280, DOI 10.1093/imanum/drv018. MR3523584

[22] G. Matthies, *Local projection methods on layer-adapted meshes for higher order discretisations of convection-diffusion problems*, Appl. Numer. Math. **59** (2009), no. 10, 2515–2533, DOI 10.1016/j.apnum.2009.05.008. MR2553152

[23] G. Matthies, *Local projection stabilisation for higher order discretisations of convection-diffusion problems on Shishkin meshes*, Adv. Comput. Math. **30** (2009), no. 4, 315–337, DOI 10.1007/s10444-008-9070-y. MR2495789

[24] H.-G. Roos and T. Linß, *Sufficient conditions for uniform convergence on layer-adapted grids*, Computing **63** (1999), no. 1, 27–45, DOI 10.1007/s006070050049. MR1702159

[25] H.-G. Roos and H. Zarin, *A supercloseness result for the discontinuous Galerkin stabilization of convection-diffusion problems on Shishkin meshes*, Numer. Methods Partial Differential Equations **23** (2007), no. 6, 1560–1576, DOI 10.1002/num.20241. MR2355174

[26] H.-G. Roos, M. Stynes, and L. Tobiska, *Robust numerical methods for singularly perturbed differential equations*, 2nd ed., Springer Series in Computational Mathematics, vol. 24, Springer-Verlag, Berlin, 2008. Convection-diffusion-reaction and flow problems. MR2454024

[27] M. Stynes and L. Tobiska, *The SDFEM for a convection-diffusion problem with a boundary layer: optimal error analysis and enhancement of accuracy*, SIAM J. Numer. Anal. **41** (2003), no. 5, 1620–1642, DOI 10.1137/S0036142902404728. MR2035000

[28] M. Stynes and T. Tobiska, *Using rectangular q_p elements in the SDFEM for a convection-diffusion problem with a boundary layer*, Appl. Numer. Math. **58** (2008), no. 12, 1709–1802.

[29] V. Thomée, *Galerkin finite element methods for parabolic problems*, Springer Series in Computational Mathematics, vol. 25, Springer-Verlag, Berlin, 1997. MR1479170

[30] M. Vlasak and H.-G. Roos, *An optimal uniform a priori error estimate for an unsteady singularly perturbed problem*, Int. J. Numer. Anal. Model. **11** (2014), no. 1, 24–33. MR3126504

[31] C. Xenophontos and S. R. Fulton, *Uniform approximation of singularly perturbed reaction-diffusion problems by the finite element method on a Shishkin mesh*, Numer. Methods Partial Differential Equations **19** (2003), no. 1, 89–111, DOI 10.1002/num.10034. MR1946805

[32] H. Zarin, *Continuous-discontinuous finite element method for convection-diffusion problems with characteristic layers*, J. Comput. Appl. Math. **231** (2009), no. 2, 626–636, DOI 10.1016/j.cam.2009.04.010. MR2549728

[33] H. Zarin and H.-G. Roos, *Interior penalty discontinuous approximations of convection-diffusion problems with parabolic layers*, Numer. Math. **100** (2005), no. 4, 735–759, DOI 10.1007/s00211-005-0598-1. MR2194592

[34] Z. Zhang, *Finite element superconvergence on Shishkin mesh for 2-D convection-diffusion problems*, Math. Comp. **72** (2003), no. 243, 1147–1177, DOI 10.1090/S0025-5718-03-01486-8. MR1972731

Institute of Scientific Computing, Technische Universität Dresden, 01062 Dresden, Germany

Email address: sebastian.franz@tu-dresden.de

Institute of Numerical Mathematics, Technische Universität Dresden, 01062 Dresden, Germany

Email address: gunar.matthies@tu-dresden.de