LAYER STRUCTURE AND THE GALERKIN FINITE ELEMENT METHOD FOR A SYSTEM OF WEAKLY COUPLED SINGULARLY PERTURBED CONVECTION-DIFFUSION EQUATIONS WITH MULTIPLE SCALES

Hans-Görg Roos[1] and Martin Schopf[1]

Abstract. We consider a system of weakly coupled singularly perturbed convection-diffusion equations with multiple scales. Based on sharp estimates for first order derivatives, Linß [T. Linß, *Computing* **79** (2007) 23–32.] analyzed the upwind finite-difference method on a Shishkin mesh. We derive such sharp bounds for second order derivatives which show that the coupling generates additional weak layers. Finally, we prove the first robust convergence result for the Galerkin finite element method for this class of problems on modified Shishkin meshes introducing a mesh grading to cope with the weak layers. Numerical experiments support our theory.

Mathematics Subject Classification. 34E15, 34A30, 65L10, 65L11, 65L60, 65L70.

1. Introduction

There is a growing interest in systems of singularly perturbed convection-reaction-diffusion equations, which have a variety of applications for instance in fluid-flow with large Reynolds number, semiconductor device modelling or problems related to heat transfer. A more exhaustive list of applications can be found in [1]. For a survey on such singularly perturbed systems we refer to [12] and to Section 5 of [16].

Consider the system of ℓ weakly coupled linear and singularly perturbed convection-diffusion equations: find $\mathbf{u} = (u_1, \ldots, u_\ell) \in \left(\mathrm{C}^2(0,1) \cap \mathrm{C}^0[0,1] \right)^\ell$ such that

$$
\left.
\begin{array}{r}
-\varepsilon_k u_k'' - b_k u_k' + \displaystyle\sum_{j=1}^{\ell} c_{kj} u_j = f_k \quad \text{on} \quad (0,1), \\[2mm]
u_k(0) = u_k(1) = 0
\end{array}
\right\}
\quad k = 1, \ldots, \ell,
\tag{1.1}
$$

where $\varepsilon_k \ll 1$ are small parameters and the functions b_k, c_{kj} and f_k $(k, j = 1, \ldots, \ell)$ are sufficiently smooth.

Since we require every solution component u_k to obey homogeneous boundary conditions it features an exponential boundary layer which interacts with other solution components through the weak coupling *via* the

[1] University of Technology Dresden, 01062 Dresden, Germany. `Hans-Goerg.Roos@tu-dresden.de; Martin.Schopf@outlook.de`

zero-order term. Note that for instance in 1D-semiconductor device modelling Dirichlet boundary conditions are often incorporated in order to prescribe values for concentrations and the electric potential.

Without loss of generality let us assume that the equations and unknowns are numbered such that

$$\varepsilon_1 \leq \varepsilon_2 \leq \ldots \leq \varepsilon_\ell \ll 1. \tag{1.2}$$

Moreover, we assume that

$$b_k \geq \beta_k \geq \beta > 0 \quad k = 1, \ldots, \ell. \tag{1.3}$$

Lemma 1.1. *Based on* (1.2) *and* (1.3) *we may assume without loss of generality that the diagonal entries of the coupling matrix* $\mathbf{C}$ *are sufficiently large, i.e. there exists a constant* $\lambda > 0$, *independent of all* ε_k, *such that*

$$c_{kk} + \frac{1}{2}b'_k \geq \sum_{\substack{j=1 \\ j \neq k}}^{\ell} |c_{kj}| + \lambda. \tag{1.4}$$

Proof. The change of variable $u_k(x) = \mathrm{e}^{-\alpha x}\tilde{u}_k(x)$ transforms the kth equation of system (1.1) into

$$-\varepsilon_k \tilde{u}''_k - (b_k - 2\alpha\varepsilon_k)\tilde{u}'_k + \sum_{\substack{j=1 \\ j \neq k}}^{\ell} c_{kj}\tilde{u}_j + (c_{kk} + \alpha b_k - \alpha^2\varepsilon_k)\tilde{u}_k = \tilde{f}_k$$

with $\tilde{f}_k := \exp(\alpha x)f(x)$. Hence, by (1.2) and (1.3) one can choose $\alpha > 0$ such that (1.4) holds true. $\qquad\square$

Up to now system (1.1) has been solved numerically only by finite difference schemes based on simple upwinding on layer adapted meshes, see [2, 3, 8, 21]. In [19, 20] two dimensional problems related to (1.1) are teated by the same approach.

Note that for $b_k \equiv 0$, $k = 1, \ldots, \ell$ problem (1.1) reduces to a system of singularly perturbed reaction-diffusion equations which is far better understood, see *e.g.* [4, 9–15].

It is well-known that the analysis of the finite element method for linear elements requires information on second order derivatives of the exact solution. In our case of multiple scales we need, moreover, detailed estimates which show the precise dependence of the derivatives up to the second order of $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_\ell$.

Bounds for the first order derivatives of the solution u to problem (1.1) have been obtained in [2, 8]. While the estimates of [2] are not sharp Linß [8] proved under certain assumptions (see Sect. 2)

$$\left| u_k^{(n)}(x) \right| \leq C \left(1 + \varepsilon_k^{-n} \exp\left(-\frac{\beta_k x}{\varepsilon_k} \right) \right), \tag{1.5}$$

for $x \in [0,1]$, $n = 0, 1$ and $k = 1, \ldots, \ell$. A problem closely related to (1.1) is studied in [17], namely

$$\left. \begin{array}{l} -\varepsilon u''_1 + b_1 u'_1 + c_{11} u_1 + c_{12} u_2 = f_1 \\ -\varepsilon u''_2 - b_2 u'_2 + c_{21} u_1 + c_{22} u_2 = f_2 \end{array} \right\} \text{ on } (0,1), \qquad \begin{array}{l} u_1(0) = u_1(1) = 0, \\ u_2(0) = u_2(1) = 0. \end{array}$$

That is (1.1) for $\ell = 2$ but with converse convection ($b_i \geq \beta_i > 0$, $i = 1, 2$) and $\varepsilon_1 = \varepsilon_2$. In that paper the existence of weak layers is observed, *i.e.* under certain assumptions on the data for instance the second order derivative of u_1 can be estimated by

$$\left| u''_1(x) \right| \leq C \left(1 + \varepsilon^{-2} \exp\left(-\frac{\beta_1(1-x)}{\varepsilon} \right) + \varepsilon^{-1} \exp\left(-\frac{\beta_2 x}{\varepsilon} \right) \right).$$

If β_1, β_2 are of moderate size, the weak boundary layer function $\varepsilon^{-1} \exp\left(-\frac{\beta_2 x}{\varepsilon}\right)$ is less significant compared to the strong boundary layer function $\varepsilon^{-2} \exp\left(-\frac{\beta_1(1-x)}{\varepsilon}\right)$. In particular, for aligned convection like in (1.1) all layers arise in a neighborhood of $x = 0$ and the weak boundary layer plays no role. However, the situation changes if $\varepsilon_1, \varepsilon_2$ are different small parameters: the approach of [17] fails as the quotients $\varepsilon_1/\varepsilon_2$ or $\varepsilon_2/\varepsilon_1$ grow out of bounds. Nonetheless the first main result of the present paper Theorem 2.7 in Section 2 shows that the second derivative of the kth solution component u_k of system (1.1) satisfies

$$|u_k''(x)| \leq C \left(1 + \varepsilon_k^{-2} \exp\left(-\frac{\beta x}{\varepsilon_k}\right) + \frac{1}{\delta} \sum_{j=k+1}^{\ell} \varepsilon_j^{-1} \exp\left(-\frac{(1-\delta)\beta x}{\varepsilon_j}\right) \right) \tag{1.6}$$

for $x \in [0,1]$, $k = 1, \ldots, \ell$ and where $\delta \in (0,1)$ is some arbitrary constant.

The estimate (1.6) shows that the kth solution component u_k features a strong boundary layer of width $\mathcal{O}(\varepsilon_k|\ln \varepsilon_k|)$ overlapped by weak layers of width $\mathcal{O}(\varepsilon_j|\ln \varepsilon_j|)$ at $x = 0$, $j > k$. The sharper layers in the other components $j < k$ do not carry over to the solution component u_k due to the increase in diffusion.

In order to prove (1.6) u is split into regular and irregular parts which are estimated separately. The bound of the regular component follows by an asymptotic expansion an L^∞-stability result for the differential operator involved and the sharp bounds (1.5) of Linß [8]. The key to analyze the irregular component is to consider it within the layer region first. There we use once more the aforementioned L^∞-stability estimate, (1.5) and additionally the differential equation. The information obtained in this manner enable a more careful analysis of the second derivative outside of the layer. A technique similar to [7] is applied on a subdomain away from the layer and a crucial application of partial integration finally puts the structure of the weak layers in perspective.

The second goal of the present paper is to provide a uniform error estimate for the linear Galerkin finite element approximation $\mathbf{u}^N := (u_1^N, \ldots, u_\ell^N)$ of problem (1.1) on suitable meshes. The importance of such estimates in the field of singularly perturbed problems cannot be stressed enough. In classical error estimates some norm of the error is bound by expressions that contain Sobolev norms of the solution. These constants depend on negative powers of the small perturbation parameters and blow up as these tend towards zero. Therefore, classical error estimates are of no practical use for singularly perturbed problems. The urge for meaningful results no matter how small the singular perturbation may be lead to the concept of uniform convergence or robustness. A uniform error estimate bounds an adequate norm of the error by a constant that is independent of any perturbation parameter times a measure for the work required, say the (ε-independent) number of mesh cells. Hence, with this concept one can guarantee that an increase in work results in an improved approximation.

In order to achieve this goal it is standard [10,18] to introduce a layer-adapted mesh based on the information on derivatives. We already referred to exponential terms like those of (1.6) as layer functions. Their presence requires some mesh condensing in those regions where their contribution to the second order derivative is significant. As an example let us consider the layer function $E : x \mapsto \varepsilon^{-j} e^{-\beta x/\varepsilon}$. Clearly, $E(x) \leq 1$ for $x \geq \tau := j\varepsilon/\beta |\ln \varepsilon|$, which is exactly why we say that the corresponding layer has the width $\mathcal{O}(\varepsilon|\ln \varepsilon|)$. This observation tempts one to define a mesh that is fine within $[0, \tau]$ and coarse in $[\tau, 1]$. The easiest implementation of this idea would be to divide each of these intervals into $N/2$ subintervals where N is an even integer — the total number of mesh intervals. Such piecewise uniform meshes are often called Asymptotic-mesh or A-mesh. While such meshes yield uniform convergence in the energy norm for reaction-diffusion problems, this is not the case for convection-diffusion problems. Therefore, other meshes are preferred for convection-diffusion problems.

If we alter the definition of τ by replacing the $|\ln \varepsilon|$ factor with $\ln N$, i.e. $\tau := j\varepsilon/\beta \ln N$ but stick to the same mesh construction process as before we arrive at the frequently-studied Shishkin meshes, see e.g. [10,18]. These meshes support the concept of uniform convergence but the (interpolation) error analysis is more complicated because the magnitude of the layer function E is still large on some part of the first coarse mesh interval.

In addition to these piecewise uniform meshes there are other layer-adapted but graded meshes, see e.g. [10,18]. Let us mention meshes of Shishkin type, Bakhvalov (type) meshes and Duran's meshes [5]. A modification of the latter will play an important role in the construction of our mesh, which we shall describe now.

The estimate (1.5) leads us to the idea to use a (different) Shishkin mesh for every solution component. That means: Setting

$$\tau_k := \min\{\nu\varepsilon_k/\beta \ln N, \ 1/2\}$$

where $\nu > 2$ is some constant, the mesh for the kth component is fine and equidistant in $[0, \tau_k]$, coarse and equidistant in $[\tau_k, 1]$ and consists of N mesh cells (assuming again that N is an even integer). However, it turns out that the additional term in (1.6) which reflects the existence of weak layers requires a modification of the mesh. We propose in Section 3 such a modification based on ideas from [5], which we want to refer to as weak layer grading. More precisely, we choose the first interval in $[\tau_k, 1]$ to have the diameter N^{-2}. Diameters of successive mesh cells $[x_{i-1}, x_i]$, are set iteratively according to

$$x_i - x_{i-1} = \sigma N^{-1} \ln N x_{i-1}, \quad i > N/2 + 1$$

where $\sigma > 0$ is a user chosen constant. Some further technicalities are required to ensure that $x = 1$ is a mesh point and that the last two mesh intervals are comparable in size. For details see Section 3, in particular Figure 1 for some illustration. Note that the generation of the mesh ω_k^N for the kth solution component is independent of any other component which simplifies our uniform error analysis considerably. We shall prove in Lemma 3.1 that the number of mesh intervals of our mesh ω_k^N is equivalent to N.

In the subsequent Section 4 we study the interpolation error on our mesh.

Finally we arrive in Section 5 at the second main result of our paper which is the uniform error estimate

$$\|\mathbf{u} - \mathbf{u}^N\|_{\boldsymbol{\varepsilon}} \leq C N^{-1} \ln N$$

for the linear finite element solution $\mathbf{u}^N$ on our modified mesh in the energy norm $\|\cdot\|_{\boldsymbol{\varepsilon}}$, *i.e.*

$$\|\mathbf{v}\|_{\boldsymbol{\varepsilon}}^2 := \sum_{k=1}^{\ell} \|v_k\|_{\varepsilon_k}^2 := \sum_{k=1}^{\ell} \left(\varepsilon_k \|v_k'\|_2^2 + \|v_k\|_2^2 \right), \tag{1.7}$$

where $\|\cdot\|_2$ denotes the $L^2(0,1)$ norm. See the the next but one paragraph for other main notation.

We support our theoretical findings by numerical experiments in Section 6. These also indicate that the linear Galerkin finite element method converges of almost second order in the L^∞ norm $\|\cdot\|_\infty$. Note that in Lemma 4.1 we prove

$$\left\|\mathbf{u} - \mathbf{u}^I\right\|_\infty \leq C N^{-2} (\ln N)^2,$$

where $\mathbf{u}^I$ is the nodal interpolant to $\mathbf{u}$ on the meshes $\{\omega_k^N\}_{k=1,\ldots,\ell}$. Hence, the convergence in the L^∞ norm can possibly be shown. Furthermore, our numerical experiments suggest a supercloseness property of $\|\mathbf{u}^I - \mathbf{u}^N\|_{\boldsymbol{\varepsilon}}$, *i.e.* the difference between the approximate solution $\mathbf{u}^N$ and the nodal interpolant $\mathbf{u}^I$ of the exact solution $\mathbf{u}$ measured in the energy norm is of higher order than the error itself.

1.1. Main notation

We use bold face symbols to denote vector valued functions with ℓ components, for instance $\mathbf{f} = (f_1, \ldots, f_\ell)$. Following this convention, system (1.1) may be written in the form

$$\mathcal{L}\mathbf{u} := -\operatorname{diag}(\boldsymbol{\varepsilon})\mathbf{u}'' - \operatorname{diag}(\mathbf{b})\mathbf{u}' + \mathbf{C}\mathbf{u} = \mathbf{f} \quad \text{on} \quad (0,1), \tag{1.8a}$$

$$\mathbf{u}(0) = \mathbf{u}(1) = \mathbf{0}, \tag{1.8b}$$

where $\operatorname{diag}(\boldsymbol{\varepsilon})$ is the $\ell \times \ell$ diagonal matrix whose diagonal entries starting in the upper left corner are $\varepsilon_1, \ldots, \varepsilon_\ell$.

Throughout the paper C will denote a generic constant that is independent of $\boldsymbol{\varepsilon}$ and any mesh. Note however that C may take different values in different places. We shall work in the context of standard Sobolev spaces $W_p^k(0,1)$, in particular $H^k(0,1) := W_2^k(0,1)$. We follow the standard notation $H_0^1(0,1)$ for the Hilbert space of $H^1(0,1)$-functions with vanishing trace. For some subinterval $D \subset (0,1)$ and $p \in \{1,2,\infty\}$ we use $\| \cdot \|_{p,D}$ to denote the standard norm in $L_p(D)$. Similarly, the symbol $\|v^{(k)}\|_{p,D}$ stands for the standard seminorm of a function $v \in W_p^k(D)$. For the sake of readability we shall drop the subscript D whenever $D = (0,1)$. The $W_p^k(D)$ norm of a vector valued function $\mathbf{v} = (v_1, \dots, v_\ell) \in W_p^k(D)^\ell$ is defined by the Euclidean norm of the $W_p^k(D)$ norms of its components, *i.e.* $\|\mathbf{v}^{(k)}\|_{p,D} := \sqrt{\sum_{j=1}^\ell \|v_j^{(k)}\|_{p,D}^2}$. We write $(\cdot,\cdot)$ for the standard inner product in $L^2(0,1)$ and extend this symbol to vector valued functions $\mathbf{v}, \mathbf{w}$ with ℓ components by $(\mathbf{v},\mathbf{w}) := \sum_{j=1}^\ell (v_j, w_j)$.

2. *A PRIORI* ESTIMATES

First we sketch Linß [8] approach to bound first order derivatives. He moves the coupling terms in (1.1) to the right hand side and applies a one dimensional stability result to obtain an L^∞-stability result for the operator $\mathcal{L}$. The following slightly improved result is taken from [12].

Theorem 2.1 (Thm. 4.1 in [12])**.** *Assume that the matrix* $\mathbf{C}$ *has non-negative diagonal entries and that the constant* $\ell \times \ell$ *matrix* $\Gamma = \Gamma(\mathbf{b}, \mathbf{C}) = (\gamma_{ij})_{i,j=1,\dots,\ell}$ *defined by*

$$\gamma_{ii} = 1 \qquad and \qquad \gamma_{ij} = -\min\left\{\left\|\frac{c_{ij}}{c_{ii}}\right\|_\infty, \left\|\frac{c_{ij}}{b_i}\right\|_\infty\right\} \quad for \quad i \neq j$$

is inverse-monotone. Then any function $\mathbf{v} = (v_1, \dots, v_\ell) \in \left(C^2(0,1) \cap C^0[0,1]\right)^\ell$ *with* $\mathbf{v}(0) = \mathbf{v}(1) = \mathbf{0}$ *satisfies*

$$\|v_i\|_\infty \leq \sum_{k=1}^\ell \left(\Gamma^{-1}\right)_{ik} \min\left\{\left\|\frac{(\mathcal{L}\mathbf{v})_k}{c_{kk}}\right\|_\infty, \left\|\frac{(\mathcal{L}\mathbf{v})_k}{b_k}\right\|_\infty\right\}. \tag{2.1}$$

Note that in the setting of Theorem 2.1 the operator $\mathcal{L}$ defined in (1.8) in general does not satisfy a maximum principle. Consequently, an L^∞-stability estimate for $\mathcal{L}$ is a weaker assumption compared to a maximum principle for this operator.

Corollary 2.2. *By Lemma 1.1 the matrix* Γ *in Theorem 2.1 is an M-matrix and in particular inverse-monotone. Thus, the operator* $\mathcal{L}$ *from* (1.8) *satisfies an* L^∞*-stability estimate of the form*

$$\|v_i\|_\infty \leq C \max_{k=1,\dots,\ell} \|(\mathcal{L}\mathbf{v})_k\|_\infty \tag{2.2}$$

for all $\mathbf{v} = (v_1, \dots, v_\ell) \in \left(C^2(0,1) \cap C^0[0,1]\right)^\ell$ *with* $\mathbf{v}(0) = \mathbf{v}(1) = \mathbf{0}$ *where* C *is independent of* ε_k, $k = 1, \dots, \ell$.

Given (2.2) problem (1.1) possesses a unique solution $\mathbf{u}$ that satisfies $\|\mathbf{u}\|_\infty \leq C \max_{k=1,\dots,\ell} \|f_k\|_\infty$, see [12]. Next, Linß applies the technique of [7] to obtain bounds for the first order derivatives.

Theorem 2.3 (Thm. 2 in [8])**.** *The solution* $\mathbf{u} = (u_1, \dots, u_\ell)$ *to problem* (1.1) *satisfies* (1.5), *i.e.*

$$\left|u_k^{(n)}(x)\right| \leq C \left(1 + \varepsilon_k^{-n} \exp\left(-\frac{\beta_k x}{\varepsilon_k}\right)\right),$$

where $x \in [0,1]$, $n = 0,1$ *and* $k = 1, \dots, \ell$.

These bounds and the differential equation (1.1) immediately yield

$$\left\|u_k''\right\|_\infty \le C\varepsilon_k^{-1}\left(1 + \varepsilon_k^{-1}\exp\left(-\frac{\beta_k x}{\varepsilon_k}\right)\right) \le C\varepsilon_k^{-2}. \tag{2.3}$$

In the following we shall improve this crude bound without assuming a maximum principle for the operator $\mathcal{L}$. For this purpose we decompose the solution $\mathbf{u}$ to problem (1.1) into a regular and irregular part and bound them separately. Hence, $\mathbf{u} = \mathbf{S} + \mathbf{E}$, where for some suitable $\boldsymbol{\alpha} \in \mathbb{R}^\ell$ the components $\mathbf{S}$ and $\mathbf{E}$ are given by (2.4) and (2.8), respectively. Since the irregular components associated with the boundary layers lie in the null space of the operator $\mathcal{L}$ this splitting is a so-called Shishkin decomposition. We start off with the regular solution component.

Lemma 2.4.

(i) *For any $\boldsymbol{\alpha} \in \mathbb{R}^\ell$ there exists a unique solution $\mathbf{S} = (S_1, \ldots, S_\ell)$ of the boundary value problem*

$$\mathcal{L}\mathbf{S} = \mathbf{f} \quad on \quad (0,1), \qquad \mathbf{S}(0) = \boldsymbol{\alpha}, \quad \mathbf{S}(1) = 0. \tag{2.4}$$

(ii) *There is a vector $\boldsymbol{\alpha} \in \mathbb{R}^\ell$ with $\|\boldsymbol{\alpha}\|_\infty \le C$ such that for the solution of (2.4) the estimates*

$$\|S_k^{(n)}\|_\infty \le C \tag{2.5}$$

for $n = 0, 1, 2$ and $k = 1, \ldots, \ell$ hold true.

Proof.

(i) The first assertion is a consequence of (2.2), c.p. [12].

(ii) Substituting the ansatz

$$\mathbf{S} = \mathbf{S}^0 + \operatorname{diag}(\boldsymbol{\varepsilon})\mathbf{S}^1 + \operatorname{diag}(\boldsymbol{\varepsilon})^2\mathbf{S}^2 \tag{2.6}$$

into (2.4) and comparing the coefficients of ε_k, $k = 1, \ldots, \ell$ we see that $\mathbf{S}^i$, $i = 1, 2, 3$ can be chosen as the unique solutions of

$$-\mathbf{b}(\mathbf{S}^0)' + \mathbf{C}\mathbf{S}^0 = \mathbf{f} \qquad \text{on} \quad (0,1), \qquad \mathbf{S}^0(1) = 0, \tag{2.7a}$$
$$-\mathbf{b}(\mathbf{S}^1)' + \mathbf{C}\mathbf{S}^1 = (\mathbf{S}^0)'' \quad \text{on} \quad (0,1), \qquad \mathbf{S}^1(1) = 0, \tag{2.7b}$$
$$\mathcal{L}\mathbf{S}^2 = (\mathbf{S}^1)'' \quad \text{on} \quad (0,1), \qquad \mathbf{S}^2(0) = \mathbf{S}^2(1) = 0. \tag{2.7c}$$

Clearly, the subproblems (2.7a) and (2.7b) defining $\mathbf{S}^0$ and $\mathbf{S}^1$ are independent of all ε_k, $k = 1, \ldots, \ell$. Hence, the derivatives of these vector-valued functions are bounded uniformly, *i.e.* $\|(S_k^i)^{(n)}\|_\infty \le C$ for $i \in \{0, 1\}$, $k = 1, \ldots, \ell$ and $n = 0, 1, 2$. This also gives a uniform bound for the value of $\boldsymbol{\alpha} := \mathbf{S}^0(0) + \operatorname{diag}(\boldsymbol{\varepsilon})\mathbf{S}^1(0)$. Finally, the bounds (1.5) and (2.3) for problem (2.7c) yield

$$\varepsilon_k^2\left\|(S_k^2)^{(n)}\right\|_\infty \le C \quad \text{for} \quad n = 0, 1, 2 \quad \text{and} \quad k = 1, \ldots, \ell.$$

Consequently, all summands of (2.6) are uniformly bounded and the result follows by a triangle inequality. $\square$

It remains to study the solution components associated with the boundary layers.

Lemma 2.5.

(i) *For any $\boldsymbol{\alpha} \in \mathbb{R}^\ell$ there exists a unique solution $\mathbf{E} = (E_1, \ldots, E_\ell)$ of the boundary value problem*

$$\mathcal{L}\mathbf{E} = \mathbf{0} \quad on \quad (0,1), \qquad \mathbf{E}(0) = -\boldsymbol{\alpha}, \quad \mathbf{E}(1) = 0. \tag{2.8}$$

(ii) *Let $\delta \in (0,1)$ be some arbitrary constant. For $\boldsymbol{\alpha} \in \mathbb{R}^{\ell}$ as in Lemma* 2.4 *the estimates*

$$|E_k^{(n)}(x)| \leq C\left(1 + \varepsilon_k^{-n} \exp\left(-\frac{\beta_k x}{\varepsilon_k}\right)\right) \qquad x \in [0,1], \; n \in \{0,1\}, \tag{2.9a}$$

$$|E_k''(x)| \leq C\left(1 + \varepsilon_k^{-2} \exp\left(-\frac{\beta x}{\varepsilon_k}\right) + \frac{1}{\delta} \sum_{j=k+1}^{\ell} \varepsilon_j^{-1} \exp\left(-\frac{(1-\delta)\beta x}{\varepsilon_j}\right)\right), \tag{2.9b}$$

for $x \in [0,1]$, $k = 1, \dots, \ell$ hold true.

Proof.

(i) Again, the assertion follows from the existence of the L^{∞}-stability result (2.2) for the operator $\mathcal{L}$.

(ii) Clearly, $\mathbf{E} = \mathbf{u} - \mathbf{S}$. Hence (2.9a) follows from Theorem 2.3 and Lemma 2.4. For instance,

$$|E_k'(x)| \leq |u_k'(x)| + |S_k'(x)| \leq C\left(1 + \varepsilon_k^{-1} \exp\left(-\frac{\beta_k x}{\varepsilon_k}\right)\right).$$

Next, we turn our attention to the second order derivative of the boundary layer function $\mathbf{E} = (E_1, \dots, E_{\ell})$. Let $x_k^{\star} := \varepsilon_k/\beta |\ln \varepsilon_k|$ denote the formal size of the boundary layer of the kth component, $k = 1, \dots, \ell$. By the first estimate of (2.3) and the definition of $x_k^{\star}$ one sees that

$$|E_k''(x)| \leq |u_k''(x)| + |S_k''(x)| \leq C\varepsilon_k^{-1}\left(1 + \varepsilon_k^{-1} \exp\left(-\frac{\beta_k x}{\varepsilon_k}\right)\right)$$

$$\leq C\varepsilon_k^{-2} \exp\left(-\frac{\beta_k x}{\varepsilon_k}\right) \leq C\varepsilon_k^{-2} \exp\left(-\frac{\beta x}{\varepsilon_k}\right) \quad \text{for} \quad x \in [0, x_k^{\star}].$$

In particular,

$$|E_k''(x_k^{\star})| \leq C\varepsilon_k^{-1}. \tag{2.10}$$

Outside of the layer region, *i.e.* in $[x_k^{\star}, 1]$ we use a technique similar to that of [7]. Note that (2.9a) and the definition of $x_k^{\star}$ give

$$|E_k'(x)| \leq C \qquad \text{for } x \in [x_k^{\star}, 1]. \tag{2.11}$$

In contrast to this the second order derivative is still large in part of this subdomain. By differentiating the differential equation for E_k we obtain

$$-\varepsilon_k E_k''' - b_k E_k'' - b_k' E_k' + \sum_{j=1}^{\ell} \left(c_{kj} E_j' + c_{kj}' E_j\right) = 0 \quad \text{in} \quad (x_k^{\star}, 1).$$

We are interested in $z_k := E_k''$. Rewriting the differential equation yields

$$-\varepsilon_k z_k' - b_k z_k = b_k' E_k' - \sum_{j=1}^{\ell} \left(c_{kj} E_j' + c_{kj}' E_j\right) =: h_k \quad \text{in} \quad (x_k^{\star}, 1).$$

Hence, z_k can be represented as

$$z_k(x) = z_k(x_k^{\star}) \exp\left(-\frac{\int_{x_k^{\star}}^{x} b_k(s)\,\mathrm{d}s}{\varepsilon_k}\right) - \varepsilon_k^{-1} \int_{x_k^{\star}}^{x} h_k(t) \exp\left(-\frac{\int_{t}^{x} b_k(s)\,\mathrm{d}s}{\varepsilon_k}\right) \mathrm{d}t, \tag{2.12}$$

for $x \in [x_k^\star, 1]$. Based on $b_k \geq \beta_k$ one easily sees that

$$\exp\left(-\frac{\int_t^x b_k(s)\,\mathrm{d}s}{\varepsilon_k}\right) \leq \exp\left(-\frac{\beta_k(x-t)}{\varepsilon_k}\right), \qquad x \geq t. \tag{2.13}$$

The definition of $x_k^\star$ and the estimates (2.9a) and (2.11) yield

$$|h_k(x)| \leq C\left(1 + \sum_{j=1}^{k} \varepsilon_j^{-1} \exp\left(-\frac{\beta_j x}{\varepsilon_j}\right) + \sum_{j=k+1}^{\ell} \varepsilon_j^{-1} \exp\left(-\frac{\beta_j x}{\varepsilon_j}\right)\right)$$

$$\leq C\left(1 + \sum_{j=k+1}^{\ell} \varepsilon_j^{-1} \exp\left(-\frac{\beta_j x}{\varepsilon_j}\right)\right) \tag{2.14}$$

for $x \in [x_k^\star, 1]$ since in the first sum the layer functions have declined already because of the definition of $x_k^\star$, in particular $x_j^\star \leq x_k^\star \leq x$ for $j \leq k$ and

$$\varepsilon_j^{-1} \exp\left(-\frac{\beta_j x}{\varepsilon_j}\right) \leq \varepsilon_j^{-1} \exp\left(-\frac{\beta x}{\varepsilon_j}\right) \leq \exp\left(-\frac{\beta(x-x_j^\star)}{\varepsilon_j}\right) \quad \text{for} \quad j \leq k. \tag{2.15}$$

Collecting (2.10), (2.12), (2.13), (2.14) and (2.15) gives

$$|z_k(x)| \leq C\varepsilon_k^{-1} \exp\left(-\frac{\beta_k(x-x_k^\star)}{\varepsilon_k}\right) + C\varepsilon_k^{-1} \int_{x_k^\star}^{x} \exp\left(-\frac{\beta_k(x-t)}{\varepsilon_k}\right) \mathrm{d}t$$

$$+ C\varepsilon_k^{-1} \int_{x_k^\star}^{x} \sum_{j=k+1}^{\ell} \varepsilon_j^{-1} \exp\left(-\frac{\beta_j t}{\varepsilon_j}\right) \exp\left(-\frac{\beta_k(x-t)}{\varepsilon_k}\right) \mathrm{d}t$$

$$=: C\varepsilon_k^{-1} \exp\left(-\frac{\beta_k(x-x_k^\star)}{\varepsilon_k}\right) + I_k + J_k. \tag{2.16}$$

For the first term, we use $\beta_k \geq \beta$ and the definition of $x_k^\star := \varepsilon_k/\beta |\ln \varepsilon_k|$ to obtain

$$\varepsilon_k^{-1} \exp\left(-\frac{\beta_k(x-x_k^\star)}{\varepsilon_k}\right) \leq \varepsilon_k^{-1} \exp\left(-\frac{\beta(x-x_k^\star)}{\varepsilon_k}\right) \leq \varepsilon_k^{-1} \exp\left(-\frac{\beta x}{\varepsilon_k}\right) \exp\left(\frac{\beta x_k^\star}{\varepsilon_k}\right) \leq \varepsilon_k^{-2} \exp\left(-\frac{\beta x}{\varepsilon_k}\right). \tag{2.17}$$

The term I_k is easily bounded:

$$|I_k| \leq C \quad \text{for} \quad x \geq x_k^\star. \tag{2.18}$$

Finally, we estimate J_k using integration by parts, $\varepsilon_j \geq \varepsilon_k$ for $j > k$ and $\beta_k \geq \beta$ for all $k = 1, \ldots, \ell$.

$$
\begin{aligned}
J_k &\leq C \left[\frac{\varepsilon_k}{\beta_k \varepsilon_k} \sum_{j=k+1}^{\ell} \varepsilon_j^{-1} \exp\left(-\frac{\beta_j t}{\varepsilon_j} \right) \exp\left(-\frac{\beta_k(x-t)}{\varepsilon_k} \right) \right]_{t=x_k^\star}^{x} \\
&\quad + C \int_{x_k^\star}^{x} \frac{\varepsilon_k}{\beta_k \varepsilon_k} \sum_{j=k+1}^{\ell} \frac{\beta_j}{\varepsilon_j^2} \exp\left(-\frac{\beta_j t}{\varepsilon_j} \right) \exp\left(-\frac{\beta_k(x-t)}{\varepsilon_k} \right) \, dt \\
&\leq C \sum_{j=k+1}^{\ell} \varepsilon_j^{-1} \exp\left(-\frac{\beta x}{\varepsilon_j} \right) + C \sum_{j=k+1}^{\ell} \varepsilon_j^{-1} \exp\left(-\frac{\beta x_k^\star}{\varepsilon_j} \right) \exp\left(-\frac{\beta(x-x_k^\star)}{\varepsilon_j} \right) \\
&\quad + C \int_{x_k^\star}^{x} \sum_{j=k+1}^{\ell} \varepsilon_j^{-2} \exp\left(-\frac{\beta t}{\varepsilon_j} \right) \exp\left(-\frac{\beta(x-t)}{\varepsilon_j} \right) \, dt \\
&\leq C \sum_{j=k+1}^{\ell} \varepsilon_j^{-1} \left(1 + \varepsilon_j^{-1}(x-x_k^\star) \right) \exp\left(-\frac{\beta x}{\varepsilon_j} \right).
\end{aligned}
\tag{2.19}
$$

In the last inequality we carried out the multiplication of the exponentials in the second and third summand. By doing so the second summand equals the first one. Additionally, the last summand becomes independent of t and is easily integrated.

We want to estimate J_k in a more intuitive way. Therefore we rewrite its second inner summand:

$$
\begin{aligned}
\frac{x - x_k^\star}{\varepsilon_j} \exp\left(-\frac{\beta x}{\varepsilon_j} \right) &= \frac{x - x_k^\star}{\varepsilon_j} \exp\left(-\frac{\delta \beta x}{\varepsilon_j} \right) \exp\left(-\frac{(1-\delta)\beta x}{\varepsilon_j} \right) \\
&= \frac{1}{\beta} \exp\left(-\frac{\delta \beta x_k^\star}{\varepsilon_j} \right) \frac{\beta(x - x_k^\star)}{\varepsilon_j} \exp\left(-\delta \frac{\beta(x-x_k^\star)}{\varepsilon_j} \right) \exp\left(-\frac{(1-\delta)\beta x}{\varepsilon_j} \right)
\end{aligned}
$$

Since $z\exp(-\delta z)$ is bounded for any fixed $\delta \in (0,1)$, in fact

$$
\max_{z \in [0,\infty)} |z\exp(-\delta z)| \leq \frac{1}{\delta} \exp\left(-\delta^2 \right) \leq \frac{1}{\delta}
$$

we can estimate

$$
\begin{aligned}
\frac{x - x_k^\star}{\varepsilon_j} \exp\left(-\frac{\beta x}{\varepsilon_j} \right) &\leq C \frac{\beta(x - x_k^\star)}{\varepsilon_j} \exp\left(-\delta \frac{\beta(x-x_k^\star)}{\varepsilon_j} \right) \exp\left(-\frac{(1-\delta)\beta x}{\varepsilon_j} \right) \\
&\leq C \frac{1}{\delta} \exp\left(-\frac{(1-\delta)\beta x}{\varepsilon_j} \right) \quad \text{for} \quad x \geq x_k^\star.
\end{aligned}
$$

Hence, (2.19) is simplified to

$$
J_k \leq \frac{C}{\delta} \sum_{j=k+1}^{\ell} \varepsilon_j^{-1} \exp\left(-\frac{(1-\delta)\beta x}{\varepsilon_j} \right),
\tag{2.20}
$$

where $\delta \in (0,1)$ is arbitrary.

Summarizing (2.16), (2.17), (2.18) and (2.20) we verified (2.9b) so the proof is completed. $\qquad\square$

Remark 2.6. The proof of Lemma 2.5 does not require that the layer components satisfy the homogeneous differential equation, *i.e.* the argument works for **u** instead of **E**. However, then one would require more smoothness of the right hand side f (recall that the differential equation was derived in order to obtain (2.9b)).

From the assertions of Lemmas 2.4 and 2.5 we can obtain certain *a priori* estimates which are the first main result of this paper.

Theorem 2.7. *Let* **u** *denote the solution to problem* (1.1) *with* (1.3) *and* (1.4).

(i) *The second derivative of* **u** *satisfies* (1.6).

(ii) *There exists a constant C such that*

$$\|u_k''\|_\infty \le C\varepsilon_k^{-2}, \qquad \|u_k''\|_2 \le C\varepsilon_k^{-\frac{3}{2}} \qquad and \qquad \|u_k''\|_1 \le C\varepsilon_k^{-1}. \tag{2.21}$$

Proof.

(i) A triangle inequality in combination with Lemmas 2.4 and 2.5 gives the first assertion.
(ii) By (1.6) and since $\varepsilon_j^{-1} \le \varepsilon_k^{-1}$ for $j \ge k$ one obtains

$$\|u_k''\|_\infty \le C\left(\varepsilon_k^{-2} + \sum_{j=k+1}^{\ell} \varepsilon_j^{-1}\right) \le C\varepsilon_k^{-2}.$$

Similarly, for $p \in \{1, 2\}$ the estimate (1.6) yields

$$\|u_k''\|_p^p \le C\left(1 + \int_0^1 \varepsilon_k^{-2p} \exp\left(-\frac{p\beta x}{\varepsilon_k}\right) \mathrm{d}x + \sum_{j=k+1}^{\ell} \varepsilon_j^{-p} \int_0^1 \mathrm{d}x\right) \le C\varepsilon_k^{-2p+1}$$

which completes the proof. Alternatively, one may use interpolation to conclude an L^p estimate from the bounds in L^1 and L^∞.

$\square$

Remark 2.8. An inspection of the proof of Theorem 2.7 shows that condition (1.4) is stronger than needed. In fact, (2.2) is sufficient for (1.6).

In Corollary 2.2 we assumed (1.4) to obtain the L^∞-stability estimate (2.2) using the results of [8, 12]. An alternative to this approach is needed whenever the transformation of Lemma 1.1 can not be applied, for instance if not all convection coefficients b_k are of the same sign.

The concept of a so-called Green's matrix is well elaborated for divergence form, second order strongly elliptic systems with smooth coefficients in bounded two dimensional domains, see *e.g.* [6] and the references therein. More precisely, the Green's matrix $\left(G_{ij}(x,\xi)\right)_{i,j=1}^{\ell}$ associated to (1.1) grants

$$u_i(x) = \int_0^1 \sum_{j=1}^{\ell} G_{ij}(x,\xi)f_j(\xi)\,\mathrm{d}\xi.$$

Consequently, the entries of the Green's matrix solve the coupled adjoint system

$$-\varepsilon_k\partial_{\xi\xi}G_{ik}(x,\xi) + \partial_\xi\left(b_k(\xi)G_{ik}(x,\xi)\right) + \sum_{j=1}^{\ell} G_{ij}(x,\xi)c_{jk} = \delta_{ik}\delta(\xi - x), \tag{2.22a}$$

$$G_{ik}(x,0) = G_{ik}(x,1) = 0, \tag{2.22b}$$

where $\delta(\cdot)$ is the Dirac delta function and δ_{ik} its discrete analog – namely Kronecker's delta, *i.e.*

$$\delta_{ik} := 0 \quad \text{for } i \ne k \quad \text{and} \quad \delta_{ii} := 1.$$

Leaving the context of distributions we may characterize $\big(G_{ij}(x,\xi)\big)_{i,j=1}^{\ell}$ in a more classical fashion. For any fixed $x \in (0,1)$ the functions $G_{ik}(x,\cdot) \in C^2\big((0,x) \cup (x,1)\big)$ are continuous on $[0,1]$ and satisfy

$$-\varepsilon_k \partial_{\xi\xi} G_{ik}(x,\xi) + \partial_\xi\big(b_k(\xi)G_{ik}(x,\xi)\big) + \sum_{j=1}^{\ell} G_{ij}(x,\xi)c_{jk} = 0 \quad \text{in } (0,1)\setminus\{x\} \tag{2.23a}$$

$$G_{ik}(x,0) = G_{ik}(x,1) = 0, \tag{2.23b}$$

$$-\varepsilon_k \big[\partial_\xi G_{ik}(x,\cdot)\big](x) = \delta_{ik}, \tag{2.23c}$$

where $[v](x) := \lim_{t\downarrow x} v(t) - \lim_{t\uparrow x} v(t)$ denotes the jump of the function v at x.

The stability estimate (2.2) could be deduced from an ε-uniform bound for

$$\max_{s\in[0,1]} \int_0^1 \sum_{j=1}^{\ell} \big|G_{ij}(s,\xi)\big|\, d\xi.$$

Remark 2.9. If there is some constant $0 < \gamma < 1$ such that

$$\varepsilon_j \leq \gamma \varepsilon_{j+1} \quad j = k+1, \ldots, \ell$$

then one can show (1.6) without the multiplier $1/\delta$ even for $\delta = 0$ using a more involved analysis.

Remark 2.10. The boundary layer interaction is of a weaker nature in comparison with coupled systems of singularly perturbed reaction-diffusion equations

$$\mathcal{L}\mathbf{u} := -\operatorname{diag}(\boldsymbol{\varepsilon})^2 \mathbf{u}'' + \mathbf{C}\mathbf{u} = \mathbf{f} \quad \text{on} \quad (0,1), \tag{2.24a}$$

$$\mathbf{u}(0) = \mathbf{u}(1) = \mathbf{0}. \tag{2.24b}$$

Under certain conditions on the coupling matrix $\mathbf{C}$ it is shown in [9] that there exists a positive constant $\kappa = \kappa(\mathbf{C})$ such that the solution to problem (2.24) can be decomposed into $\mathbf{u} = \mathbf{w} + \mathbf{v}$ with

$$\big|w_k^{(n)}(x)\big| \leq C\left(1 + \sum_{j=k}^{\ell} \varepsilon_j^{-n} \exp\left(-\frac{\kappa x}{\varepsilon_j}\right)\right),$$

$$\big|v_k^{(n)}(x)\big| \leq C\left(1 + \sum_{j=k}^{\ell} \varepsilon_j^{-n} \exp\left(-\frac{\kappa(1-x)}{\varepsilon_j}\right)\right),$$

for $n = 0,1,2$ and $k = 1,\ldots,\ell$.

3. A MODIFIED SHISHKIN MESH WITH ADDITIONAL WEAK LAYER GRADING

In order to discretize (1.1) let N denote an even integer. This discretization parameter will later be associated with the number of subintervals within the mesh. We want to employ for every component of (1.1) a different mesh of Shishkin flavor with some additional grading to resolve the weak layers. For this purpose let

$$\tau_k := \min\{\nu\varepsilon_k/\beta \ln N,\, 1/2\}$$

denote a transition point where ν is some constant with $\nu > 2$. If $\tau_k = 1/2$ then N is exponentially large compared to $1/\varepsilon_k$ and we choose our mesh ω_k^N for the kth component of system (1.1) to be uniform with N mesh intervals.
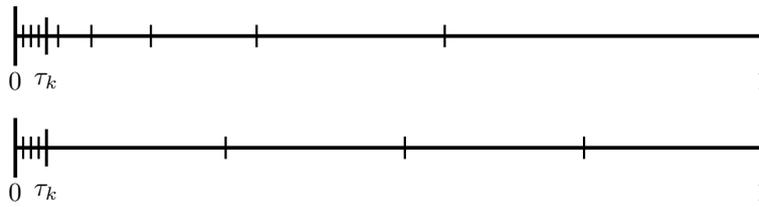
FIGURE 1. Shishkin mesh (below) and its variant with additional weak layer grading (top), $\varepsilon_k = 10^{-2}$, $N = 8$, $\beta = 1$, $\sigma = 3$.

If on the other hand $\tau_k < 1/2$ then in $[0, \tau_k]$ we obtain our mesh ω_k^N by dividing the interval $[0, \tau_k]$ into $N/2$ equidistant subintervals. Hence, close to $x = 0$ the mesh intervals are of the small size $h_k = \mathcal{O}(\varepsilon_k N^{-1} \ln N)$. In the remaining interval $[\tau_k, 1]$ we adopt some ideas of [5]. More precisely, the mesh points are given by the set $\omega_k^N = \{x_i, i = 1, \ldots, M\}$ defined by

$$x_i = \begin{cases} 2i\tau_k/N & \text{for } i = 0, \ldots, N/2, \\ \tau_k + N^{-2} & \text{for } i = N/2 + 1, \\ x_{i-1} + \sigma N^{-1} \ln N x_{i-1} & \text{for } i = N/2 + 2, \ldots, M - 1, \\ 1 & \text{for } i = M, \end{cases} \tag{3.1}$$

where M is such that $x_{M-1} < 1 \leq x_{M-1} + \sigma N^{-1} \ln N x_{M-1}$ and $\sigma > 0$ is a user chosen constant. Similar to [5] we assume without loss of generality that the last interval $(x_{M-1}, 1)$ is not to small in comparison with the previous one (x_{M-2}, x_{M-1}) as otherwise one can move the node x_{M-1} to the midpoint of the interval (x_{M-2}, x_M). See Figure 1 for some illustration.

Let us estimate the number of mesh intervals M in terms of N.

Lemma 3.1. *For the number M of mesh intervals of ω_k^N it holds $M \sim N$.*

Proof. If $\tau_k = 1/2$ one has $M = N$. For $\tau_k < 1/2$ we follow the spirit of ([5], proof of Cor. 2.3).

$$M = \frac{N}{2} + 1 + \sum_{i=N/2+2}^{M} 1 = \frac{N}{2} + 1 + \sum_{i=N/2+2}^{M} (x_i - x_{i-1})^{-1} \int_{x_{i-1}}^{x_i} dx.$$

Based on the definition of the grid one finds

$$\frac{1}{x_i - x_{i-1}} = \frac{N}{\sigma \ln N x_{i-1}} \leq \frac{N}{\sigma \ln N} \frac{1}{x}$$

for $x \in [x_{i-1}, x_i]$. Hence,

$$M - \frac{N}{2} - 1 \leq \frac{N}{\sigma \ln N} \sum_{i=N/2+2}^{M} \int_{x_{i-1}}^{x_i} \frac{1}{x} \, dx = \frac{N}{\sigma \ln N} \int_{x_{N/2+1}}^{1} \frac{1}{x} \, dx.$$

Since $1 \geq x_{N/2+1} = \tau_k + N^{-2} > N^{-2}$ one has

$$0 \leq -\ln(x_{N/2+1}) \leq -\ln(N^{-2}) \leq 2 \ln N.$$

Consequently,

$$M \leq \frac{N}{2} + 1 + \frac{N}{\sigma \ln N} 2 \ln N = N \left(\frac{1}{2} + \frac{2}{\sigma} \right) + 1$$

and $M \sim N$ because $N/2 \leq M$ is trivial. □

4. The interpolation error on the modified mesh

Let $V_k^N \subset H_0^1(0,1)$ denote the space of linear finite element functions based on the triangulation ω_k^N and set $\mathbf{V}^N := V_1^N \times \ldots \times V_\ell^N$.

Next, let $\mathbf{u}^I \in \mathbf{V}^N$ denote the linear nodal interpolant to $\mathbf{u}$.

Lemma 4.1. *For the interpolation error it holds that*

$$\left\| u_k - u_k^I \right\|_\infty \le CN^{-2}(\ln N)^2, \tag{4.1a}$$

$$\varepsilon_k^{1/2} \left\| \left(u_k - u_k^I \right)' \right\|_2 \le CN^{-1} \ln N, \tag{4.1b}$$

for $k = 1 \ldots, \ell$.

Proof. Let us consider the case $\tau_k < 1/2$ as in the other case the assertion follows easily by classical arguments. Using the well-known estimate

$$\left\| u_k - u_k^I \right\|_{\infty,T} \le C h_T^2 \|u_k''\|_{\infty,T}$$

over all elements $T \in [0, \tau_k]$ with $h_T := \operatorname{diam} T = h_k$ one obtains

$$\left\| u_k - u_k^I \right\|_{\infty,[0,\tau_k]} \le C h_k^2 \|u_k''\|_{\infty,[0,\tau_k]} \le C(\varepsilon_k N^{-1} \ln N)^2 \|u_k''\|_\infty,$$

from which

$$\left\| u_k - u_k^I \right\|_{\infty,[0,\tau_k]} \le CN^{-2}(\ln N)^2 \tag{4.2}$$

follows due to the first estimate of (2.21).

Interchanging the order of integration and estimating the representation

$$\left(u_k^I - u_k \right)(x) = \frac{1}{x_i - x_{i-1}} \int_{x_{i-1}}^{x_i} \int_{x_{i-1}}^x \int_\xi^s u_k''(t) \, \mathrm{d}t \, \mathrm{d}\xi \, \mathrm{d}s$$

for the interpolation error in $[x_{i-1}, x_i]$, cp. [10, page 152] one obtains

$$\left| \left(u_k - u_k^I \right)(x) \right| \le \int_{x_{i-1}}^{x_i} (\xi - x_{i-1}) |u_k''(\xi)| \, \mathrm{d}\xi.$$

Let $x \in [x_{i-1}, x_i] \subset [\tau_k, 1]$. Based on (2.5) and (2.9) we can estimate

$$\left| \left(u_k - u_k^I \right)(x) \right| \le C \int_{x_{i-1}}^{x_i} (\xi - x_{i-1}) \left(1 + \varepsilon_k^{-2} \exp\left(-\frac{\beta\xi}{\varepsilon_k} \right) \right) \mathrm{d}\xi$$

$$+ C \int_{x_{i-1}}^{x_i} (\xi - x_{i-1}) \sum_{j=k+1}^\ell \varepsilon_j^{-1} \exp\left(-\frac{(1-\delta)\beta\xi}{\varepsilon_j} \right) \mathrm{d}\xi =: I_1(x) + I_2(x). \tag{4.3}$$

For any positive integer p and any positive monotonically decreasing function g over some interval $[a, b]$ it holds true that

$$\int_a^b (\xi - a)^{p-1} g(\xi) \, \mathrm{d}\xi \le \frac{1}{p} \left(\int_a^b g(\xi)^{1/p} \, \mathrm{d}\xi \right)^p,$$

which can easily be verified by considering both integrals as functions of the upper integration limit, cp. ([10], Lem. 4.16). Hence,

$$
\begin{aligned}
I_1(x) &\le C \left(\int_{x_{i-1}}^{x_i} 1 + \varepsilon_k^{-1} \exp\left(-\frac{\beta\xi}{2\varepsilon_k}\right) \, \mathrm{d}\xi \right)^2 \\
&\le C \left((x_i - x_{i-1}) + \int_{\tau_k}^{\infty} \varepsilon_k^{-1} \exp\left(-\frac{\beta\xi}{2\varepsilon_k}\right) \, \mathrm{d}\xi \right)^2 \\
&\le C \left(\max\{N^{-2}, \sigma N^{-1} \ln N x_{i-1}\} + N^{-1} \right)^2 \le C N^{-2} (\ln N)^2.
\end{aligned}
\tag{4.4}
$$

Next, we estimate $I_2(x)$ and consider the first mesh interval in $[\tau_k, 1]$ separately. For $\xi \in [x_{N/2}, x_{N/2+1}]$ the trivial estimate $\xi - x_{N/2} \le \xi$ yields

$$
(\xi - x_{N/2})\varepsilon_j^{-1} \exp\left(-\frac{(1-\delta)\beta\xi}{\varepsilon_j}\right) \le \frac{\xi}{\varepsilon_j} \exp\left(-\frac{(1-\delta)\beta\xi}{\varepsilon_j}\right) \le C,
\tag{4.5}
$$

because $z \exp(-(1-\delta)\beta z)$ remains bounded for $z \in [0, \infty)$. Hence for any $x \in [x_{N/2}, x_{N/2+1}] = [\tau_k, \tau_k + N^{-2}]$ we find that

$$
I_2(x) \le C \int_{x_{N/2}}^{x_{N/2+1}} \ell \, \mathrm{d}\xi \le C N^{-2}.
\tag{4.6}
$$

Now let $x \in [x_{i-1}, x_i]$ with $i > N/2 + 1$. For $\xi \in [x_{i-1}, x_i]$ one can estimate

$$
\xi - x_{i-1} \le x_i - x_{i-1} \le \sigma N^{-1} \ln N x_{i-1} \le \sigma N^{-1} \ln N \xi.
\tag{4.7}
$$

Hence and by $\frac{\xi}{\varepsilon_j} \exp\left(-\frac{(1-\delta)\beta\xi}{\varepsilon_j}\right) \le C$ according to (4.5) and $x_i - x_{i-1} \le C N^{-1} \ln N$, as in (4.7),

$$
\begin{aligned}
I_2(x) &\le C N^{-1} \ln N \int_{x_{i-1}}^{x_i} \xi \sum_{j=k+1}^{\ell} \varepsilon_j^{-1} \exp\left(-\frac{(1-\delta)\beta\xi}{\varepsilon_j}\right) \, \mathrm{d}\xi \\
&\le C \ell N^{-1} \ln N (x_i - x_{i-1}) \le C N^{-2} (\ln N)^2.
\end{aligned}
\tag{4.8}
$$

Collecting (4.3), (4.4), (4.6) and (4.8) we arrive at

$$
\left\| u_k - u_k^I \right\|_{\infty, [\tau_k, 1]} \le C N^{-2} (\ln N)^2.
\tag{4.9}
$$

Summarizing (4.2) and (4.9) the estimate (4.1a) is verified.

Finally, let us consider the H^1 norm of the interpolation error. We use integration by parts to obtain

$$
\begin{aligned}
\left\| (u_k - u_k^I)' \right\|_2^2 &= \sum_{i=1}^{M} \int_{x_{i-1}}^{x_i} \left((u_k - u_k^I)'(x) \right)^2 \, \mathrm{d}x \\
&= -\sum_{i=1}^{M} \int_{x_{i-1}}^{x_i} (u_k - u_k^I)(x) u_k''(x) \, \mathrm{d}x.
\end{aligned}
$$

Hence, by a Hölder's inequality (4.1a) and the last estimate of (2.21)

$$
\left\| (u_k - u_k^I)' \right\|_2^2 \le \| u_k - u_k^I \|_\infty \| u_k'' \|_1 \le C \varepsilon_k^{-1} N^{-2} (\ln N)^2.
$$

Thus, the estimate (4.1b) is proven. $\square$

In $[\tau_k + N^{-2}, 1]$ we need an H^1 norm estimate without the ε_k multiplier.

Lemma 4.2. *The interpolation error satisfies*

$$\left\|\left(u_k - u_k^I\right)'\right\|_{2,[\tau_k+N^{-2},1]} \leq CN^{-1}\ln N$$

for N sufficiently large independently of ε (depending only on ν and β).

Proof. We begin by splitting the subdomain $[\tau_k + N^{-2}, 1]$ into mesh intervals and applying the classical interpolation error estimate $\left\|\left(u_k - u_k^I\right)'\right\|_{2,T} \leq Ch_T\|u_k''\|_{2,T}$. Thus,

$$\left\|\left(u_k - u_k^I\right)'\right\|^2_{2,[\tau_k+N^{-2},1]} \leq \sum_{i=N/2+2}^{M} \left\|\left(u_k - u_k^I\right)'\right\|^2_{2,[x_{i-1},x_i]}$$

$$\leq \sum_{i=N/2+2}^{M} (x_i - x_{i-1})^2 \|u_k''\|^2_{2,[x_{i-1},x_i]}.$$

Next, we use that $x_i - x_{i-1} = \sigma N^{-1}\ln N x_{i-1} \leq CN^{-1}\ln N x$ for $x \geq x_{i-1}$. Hence,

$$\left\|\left(u_k - u_k^I\right)'\right\|^2_{2,[\tau_k+N^{-2},1]} \leq CN^{-2}(\ln N)^2 \sum_{i=N/2+2}^{M} \|xu_k''\|^2_{2,[x_{i-1},x_i]} \tag{4.10}$$

$$\leq CN^{-2}(\ln N)^2 \|xu_k''\|^2_{2,[\tau_k+N^{-2},1]}.$$

With (1.6) we estimate the weighted H^2 seminorm of u_k

$$\|xu_k''\|^2_{2,[\tau_k+N^{-2},1]} \leq C + C\int_{\tau_k+N^{-2}}^{1} x^2\varepsilon_k^{-4}\exp\left(-\frac{2\beta x}{\varepsilon_k}\right)\ \mathrm{d}x$$

$$+ C\sum_{j=k+1}^{\ell}\int_{\tau_k+N^{-2}}^{1} x^2\varepsilon_j^{-2}\exp\left(-\frac{2(1-\delta)\beta x}{\varepsilon_j}\right)\ \mathrm{d}x$$

$$=: C + I_1 + I_2.$$

The term I_2 is easily seen to be bounded, *i.e.* $|I_2| \leq C$ because

$$\left(\frac{x}{\varepsilon_j}\right)^2\exp\left(-\frac{2(1-\delta)\beta x}{\varepsilon_j}\right) \leq C$$

as $z^2\exp\left(-2(1-\delta)\beta z\right)$ is bounded for $z \in [0,\infty)$.

It remains to study the term I_1 induced by the strong boundary layer. Let again $x_k^\star := \varepsilon_k/\beta|\ln\varepsilon_k|$ denote the formal size of the boundary layer of the kth component, $k = 1,\ldots,\ell$. In ([5], proof of Lem 2.1.) the authors show by a straight forward calculation that

$$\int_{\kappa x_k^\star}^{1} x^2\left(1 + \varepsilon_k^{-4}\exp\left(-\frac{2\beta x}{\varepsilon_k}\right)\right)\ \mathrm{d}x \leq C \tag{4.11}$$

for some constant $\kappa > 1/2$. Hence, we would obtain

$$\|xu_k''\|^2_{2,[\tau_k+N^{-2},1]} \leq C \tag{4.12}$$

if we proved

$$\kappa x_k^\star \leq \tau_k + N^{-2}.$$

Since we set the constant ν within the definition of the transition point τ_k according to $\nu > 2$ we can find a value for κ such that

$$\frac{\nu}{\kappa} > 4 \quad \text{with} \quad \kappa > 1/2, \tag{4.13}$$

hence, this κ is a valid choice for (4.11).

For $\varepsilon_k^{-1} \leq N^{\nu/\kappa}$ we estimate

$$\kappa x_k^\star = \kappa \frac{\varepsilon_k}{\beta} \ln \frac{1}{\varepsilon_k} \leq \kappa \frac{\varepsilon_k}{\beta} \ln(N^{\nu/\kappa}) = \frac{\nu \varepsilon_k}{\beta} \ln N = \tau_k < \tau_k + N^{-2}.$$

Otherwise we have $\varepsilon_k < N^{-\nu/\kappa}$ and proceed as follows

$$\kappa x_k^\star = \kappa \frac{\varepsilon_k}{\beta} \ln \frac{1}{\varepsilon_k} \leq \frac{\kappa}{\beta} \sqrt{\varepsilon_k} < \frac{\kappa}{\beta} N^{-\frac{\nu}{2\kappa}} \leq N^{-2}$$

because of (4.13) for N sufficiently large, independently of $\boldsymbol{\varepsilon}$, *i.e.* for

$$N \geq C_{\nu,\beta},$$

where $C_{\nu,\beta}$ depends only on ν and β (and κ which is however determined by the particular value of ν). With (4.12) verified the assertion of the Lemma follows from (4.10). □

5. ERROR ANALYSIS OF THE GALERKIN FEM

A standard weak formulation of the kth equation of the weakly coupled system (1.1) reads: find $u_k \in H_0^1(0,1)$, such that

$$\varepsilon_k(u_k', v_k') - (b_k u_k', v_k) + \sum_{j=1}^{\ell}(c_{kj} u_j, v_k) = (f_k, v_k) \tag{5.1}$$

for each $v_k, \in H_0^1(0,1)$ and $k = 1, \ldots, \ell$. Summing up (5.1) for $k = 1, \ldots, \ell$ yields a weak formulation for the entire system (1.1):

$$\mathbf{a}(\mathbf{u}, \mathbf{v}) = (\mathbf{f}, \mathbf{v}) \tag{5.2}$$

for each $\mathbf{v} \in \mathbf{V} := \left(H_0^1(0,1)\right)^\ell$ where

$$\mathbf{a}(\mathbf{u}, \mathbf{v}) := \sum_{k=1}^{\ell} \left(\varepsilon_k(u_k', v_k') + (c_{kk} u_k - b_k u_k', v_k) + \sum_{\substack{j=1 \\ j \neq k}}^{\ell}(c_{kj} u_j, v_k) \right) \tag{5.3}$$

$$= (\,\text{diag}(\varepsilon)\mathbf{u}', \mathbf{v}') - (\,\text{diag}(\mathbf{b})\mathbf{u}', \mathbf{v}) + (\mathbf{C}\mathbf{u}, \mathbf{v}).$$

We want to solve (5.2) numerically with linear finite elements. Consequently, we have to define a finite element space $\mathbf{V}^N \subset \mathbf{V}$ and the finite element approximation $\mathbf{u}^N := (u_1^N, \ldots, u_\ell^N) \in \mathbf{V}^N$ then solves the linear system

$$\mathbf{a}(\mathbf{u}^N, \mathbf{v}^N) = (\mathbf{f}, \mathbf{v}^N) \quad \text{for all } \mathbf{v}^N \in \mathbf{V}^N. \tag{5.4}$$

In the next lemma we prove coercivity of the bilinear form $\mathbf{a}(\cdot, \cdot)$ with respect to the energy norm $\|\cdot\|_{\boldsymbol{\varepsilon}}$. Recall that

$$\|\mathbf{v}\|_{\boldsymbol{\varepsilon}}^2 := \sum_{k=1}^{\ell} \|v_k\|_{\varepsilon_k}^2 := \sum_{k=1}^{\ell} \left(\varepsilon_k \|v_k'\|_2^2 + \|v_k\|_2^2 \right).$$

This property implies that (5.2) and (5.4) have unique solutions $\mathbf{u} \in \mathbf{V}$ and $\mathbf{u}^N \in \mathbf{V}^N$. Moreover, it will be vital in our error analysis.

Lemma 5.1. *The bilinear form* $\mathbf{a}(\cdot,\cdot)$ *is coercive in* $\left(H_0^1(0,1)\right)^\ell$ *with respect to the energy norm,* i.e. *there exists* $\mu > 0$, *independent of* $\boldsymbol{\varepsilon}$, *such that*

$$\mu\|\mathbf{v}\|_{\boldsymbol{\varepsilon}}^2 \le \mathbf{a}(\mathbf{v},\mathbf{v}).$$

Proof. Using $-(b_k v_k', v_k) = \frac{1}{2}(b_k' v_k, v_k)$ for $v_k \in H_0^1(0,1)$, $k = 1,\ldots,\ell$ and (1.4) one can show that there exists some constant $\tilde\mu > 0$ that is independent of $\boldsymbol{\varepsilon}$ such that

$$\sum_{k=1}^\ell \left(\left(\left(c_{kk} + \frac{1}{2} b_k' \right) v_k, v_k \right) + \sum_{\substack{j=1 \\ j\neq k}}^\ell (c_{kj} v_j, v_k) \right) \ge \tilde\mu \sum_{k=1}^\ell (v_k, v_k).$$

In fact, this is equivalent to

$$\left(\left(\mathbf{C} + \frac{1}{2}\operatorname{diag}(\mathbf{b}') \right) \mathbf{v}, \mathbf{v} \right) \ge \tilde\mu(\mathbf{v},\mathbf{v})$$

and (1.4) yields

$$\frac{\mathbf{v}^T \left(\mathbf{C}(x) + \frac{1}{2}\operatorname{diag}\left(\mathbf{b}'(x)\right) \right)\mathbf{v}}{\mathbf{v}^T\mathbf{v}} \ge \tilde\mu > 0$$

for all $x \in [0,1]$ and $\mathbf{v} \in \mathbb{R}^\ell \setminus \{\mathbf{0}\}$. Setting $\mu := \min\{1,\tilde\mu\}$ completes the proof. $\square$

Based on $\mathbf{V}^N \subset \left(H_0^1(0,1)\right)^\ell$, (5.2) and (5.4) the error features the Galerkin orthogonality property

$$\mathbf{a}(\mathbf{u} - \mathbf{u}^N, \mathbf{v}^N) = 0 \qquad \text{for each } \mathbf{v}^N \in \mathbf{V}^N. \tag{5.5}$$

We have now all the tools available to prove the second main result of this paper.

Theorem 5.2. *Let* $\mathbf{u}$ *denote the solution to problem* (1.1) *with* (1.3) *and* (1.4). *Then the linear Galerkin finite element solution* $\mathbf{u}^N$ *on the modified Shishkin meshes* $(\omega_k^N)_{k=1,\ldots,\ell}$ *with* $\mathcal{O}(N)$ *mesh intervals defined in* (3.1) *satisfies the uniform error estimate*

$$\|\mathbf{u} - \mathbf{u}^N\|_{\boldsymbol{\varepsilon}} \le C N^{-1} \ln N \tag{5.6}$$

for N *sufficiently large independently of* $\boldsymbol{\varepsilon}$ (*depending only on* ν *and* β).

Proof. Let $\mathbf{e}^N := \mathbf{u}^N - \mathbf{u}^I \in \mathbf{V}^N$. In view of Lemma 4.1 it remains to establish an energy norm bound for $\mathbf{e}^N$. In a standard approach we abbreviate $\boldsymbol{\eta} := \mathbf{u} - \mathbf{u}^I$ and start off with coercivity (Lem. 5.1) and Galerkin orthogonality (5.5):

$$\mu\|\mathbf{e}^N\|_{\boldsymbol{\varepsilon}}^2 \le \mathbf{a}(\mathbf{e}^N, \mathbf{e}^N) = \mathbf{a}(\boldsymbol{\eta}, \mathbf{e}^N)$$

$$= \sum_{k=1}^\ell \left(\varepsilon_k \left(\eta_k', (e_k^N)' \right) + (c_{kk}\eta_k - b_k\eta_k', e_k^N) + \sum_{\substack{j=1 \\ j\neq k}}^\ell \left(c_{kj}\eta_j, e_k^N \right) \right)$$

$$\le \left(1 + \left(\sum_{j,k=1}^\ell \|c_{jk}\|_\infty^2 \right)^{1/2} \right) \|\boldsymbol{\eta}\|_{\boldsymbol{\varepsilon}} \|\mathbf{e}^N\|_{\boldsymbol{\varepsilon}} + \sum_{k=1}^\ell \left| (b_k\eta_k', e_k^N) \right|.$$

Using Lemma 4.1 we arrive at

$$\mu\|\mathbf{e}^N\|_{\boldsymbol{\varepsilon}}^2 \le CN^{-1}\ln N\|\mathbf{e}^N\|_{\boldsymbol{\varepsilon}} + \sum_{k=1}^{\ell} \left|(b_k\eta_k', e_k^N)\right|. \tag{5.7}$$

Let us examine a remaining summand. The triangle inequality and integration by parts give

$$\begin{aligned}
\left|(b_k\eta_k', e_k^N)\right| &\le \left|\int_0^{x_{N/2+1}} b_k(x)\eta_k'(x)e_k^N(x)\ \mathrm{d}x\right| + \left|\int_{x_{N/2+1}}^1 b_k(x)\eta_k'(x)e_k^N(x)\ \mathrm{d}x\right| \\
&\le \left|\int_0^{x_{N/2+1}} b_k(x)\eta_k(x)(e_k^N)'(x)\ \mathrm{d}x\right| + \left|\int_0^{x_{N/2+1}} b_k'(x)\eta_k(x)e_k^N(x)\ \mathrm{d}x\right| \\
&\quad + \left|\int_{x_{N/2+1}}^1 b_k(x)\eta_k'(x)e_k^N(x)\ \mathrm{d}x\right| =: I_1 + I_2 + I_3.
\end{aligned} \tag{5.8}$$

The second term I_2 is easily estimated using Hölder's inequality

$$I_2 \le \|b_k'\|_\infty\|\eta_k\|_2\|e_k^N\|_2 \le CN^{-1}\ln N\|e_k^N\|_{\varepsilon_k}. \tag{5.9}$$

For the first term we benefit from the smallness of the subdomain $[0, x_{N/2}] = [0, \tau_k]$ and the adjacent element:

$$\begin{aligned}
I_1 &\le \|b_k\|_\infty\Big(\|\eta_k\|_\infty\|(e_k^N)'\|_{1,(0,\tau_k)} + \|\eta_k\|_{\infty,[\tau_k,\tau_k+N^{-2}]}\|(e_k^N)'\|_{1,(\tau_k,\tau_k+N^{-2})}\Big) \\
&\le C\Big(N^{-2}(\ln N)^2\tau_k^{1/2}\|(e_k^N)'\|_{2,[0,\tau_k]} + N^{-2}N^{-1}\|(e_k^N)'\|_{2,[\tau_k,\tau_k+N^{-1}]}\Big) \\
&\le C\left(N^{-2}(\ln N)^{5/2}\varepsilon_k^{1/2}\|(e_k^N)'\|_2 + N^{-1}\|e_k^N\|_2\right) \le CN^{-1}\ln N\|e_k^N\|_{\varepsilon_k},
\end{aligned} \tag{5.10}$$

where we also used an inverse estimate on the element $[x_{N/2}, x_{N/2+1}] = [\tau_k, \tau_k + N^{-2}]$.

For the final third term I_3 Lemma 4.2 and another Hölder's inequality give

$$I_3 \le \|b\|_\infty\|\eta_k'\|_{2,[\tau_k+N^{-2},1]}\|e_k^N\|_2 \le CN^{-1}\ln N\|e_k^N\|_{\varepsilon_k}. \tag{5.11}$$

Collecting (5.8), (5.9), (5.10) and (5.11) we arrive at

$$\sum_{k=1}^{\ell}\left|(b_k\eta_k', e_k^N)\right| \le CN^{-1}\ln N\|\mathbf{e}^N\|_{\boldsymbol{\varepsilon}}.$$

A combination of this estimate with (5.7) completes the proof. $\qquad\square$

Remark 5.3. Assume compatibility conditions in all corners of a rectangular domain and a Shishkin decomposition of the solution into regular and boundary layer components such that the latter are exponentially small in points that are sufficiently remote to the critical boundary. Then it should be possible to extend the result (5.6) to Q_1 elements on tensor product meshes based on our modified Shishkin mesh.

6. NUMERICAL RESULTS

In comparison with a finite element method based on solely one mesh for all components an implementation featuring the support of multiple meshes is more demanding. The additional flexibility comes at the price of having to compute stiffness matrix entries involving integrals of basis functions defined on different meshes. These integrals are calculated efficiently if both basis functions are reinterpreted as finite element functions over the same mesh which can be seen as a suitable refinement of both meshes. On this common refinement mesh

assembling the stiffness matrix $\tilde{M}$ is easy and existing code can be reused. In order to obtain the desired two-mesh stiffness matrix M the matrix $\tilde{M}$ has to be multiplied on both sides with sparse transformation matrices describing the aforementioned reinterpretation process of the basis functions.

We want to examine the performance of the proposed method if applied to the following test problem

$$-\varepsilon_1 u_1''(x) - u_1'(x) + 2u_1(x) - u_2(x) = \mathrm{e}^x \qquad x \in (0, 1), \tag{6.1a}$$
$$-\varepsilon_2 u_2''(x) - \theta u_2'(x) - u_1(x) + 4u_2(x) = \cos(x) \quad x \in (0, 1), \tag{6.1b}$$
$$u_1(0) = u_1(1) = 0, \tag{6.1c}$$
$$u_2(0) = u_2(1) = 0, \tag{6.1d}$$

with $\theta \in \{2, -2\}$. For $\theta = 2$ (*i.e.* aligned convection) this problem was also considered in [8]. We identify $\ell = 2$, $\beta = 1$ and set $\nu := 2.5$, $\sigma := 3$.

We used MAPLE to compute an exact solution with 64 digits. For instance for aligned convection, more precisely $\theta = 2$, $\varepsilon_1 = 10^{-8}$ and $\varepsilon_2 = 10^{-4}$ we obtain

$$u_1(x) = \sum_{i=1}^{4} a_i \exp(r_i x) + a_5 \cos(x) + a_6 \sin(x) + a_7 \exp(x),$$

where

$$r_1 \approx -100000001.99999996000100179995991200440532971367762,$$
$$r_2 \approx -20001.999775044987191192529194240288306208889271114,$$
$$r_3 \approx 1.2928514226019047729222749864154238627852279334615,$$
$$r_4 \approx 2.7069235823862882195668312582782120597816670413239,$$
$$a_1 \approx -0.73630993627099030006639499165518728673719879021 28,$$
$$a_2 \approx -0.17080524288334069552263866119053855416318406007 17 \times 10^{-4},$$
$$a_3 \approx -1.2685632431331649784592768861625809741096802997002,$$
$$a_4 \approx -0.51389540072523197710485779722395110344076792974 88 \times 10^{-1},$$
$$a_5 \approx 0.56179749995965410115667914804972204004664919952 70 \times 10^{-1},$$
$$a_6 \approx -0.89885127689057014658552346743703052530358406051 77 \times 10^{-1},$$
$$a_7 \approx 2.0001000500050014001900420066013102210417072813412.$$

As expected, we see that u_1 features a strong exponential layer of width $\mathcal{O}(\varepsilon_1)$ and a weak layer of width $\mathcal{O}(\varepsilon_2)$ at $x = 0$ which is induced by the weak coupling with the second component.

Similarly, one can obtain such a representation for u_2.

All integrals (in the assembly routines of both methods and for the calculation of L^2-based errors) are approximated by a standard five point Gaussian quadrature formula over each element of the mesh introduced. The numerical rates of convergence p are computed from successive pairs $(N_1, err(N_1))$, $(N_2, err(N_2))$ by eliminating the constant C in

$$err(N_1) = C \left(N_1^{-1} \ln N_1 \right)^p \qquad \text{and} \qquad err(N_2) = C \left(N_2^{-1} \ln N_2 \right)^p.$$

and solving for p.

First, we study the performance of the Galerkin method applied to the test problem with aligned convection, *i.e.* $\theta = 2$. In Table 1 we use the Galerkin FEM on the modified Shishkin meshes for $\varepsilon_1 = 10^{-8}$ and $\varepsilon_2 = 10^{-4}$ as in the above example. Moving to lower rows in this table corresponds to a finer discretization. In fact, the first column denotes the discretization parameter N which is doubled in every row. The second column depicts

TABLE 1. Errors and convergence rates for the Galerkin FEM on the modified Shishkin mesh applied to the test problem with $\theta = 2$, $\varepsilon_1 = 10^{-8}$ and $\varepsilon_2 = 10^{-4}$.

N	$\max\limits_{k=1,2} M$	$\|\mathbf{u} - \mathbf{u}^N\|_\varepsilon$		$\|\mathbf{u} - \mathbf{u}^N\|_\infty$		$\|\mathbf{u}^I - \mathbf{u}^N\|_\varepsilon$	
		error	rate	error	rate	error	rate
128	156	4.691e-2	1.00	7.575e-3	1.55	3.131e-3	1.92
256	306	2.685e-2	1.00	3.178e-3	1.85	1.070e-3	2.01
512	604	1.511e-2	1.00	1.093e-3	1.78	3.359e-4	1.94
1024	1195	8.393e-3	1.00	3.844e-4	1.87	1.071e-4	1.97
2048	2346	4.616e-3	1.00	1.257e-4	1.96	3.293e-5	2.01
4096	4541	2.518e-3	1.00	3.840e-5	1.96	9.722e-6	1.97
8192	8724	1.364e-3	1.00	1.157e-5	1.98	2.903e-6	2.01
16 384	16 765	7.344e-4	1.00	3.403e-6	1.99	8.379e-7	1.97
32 768	32 318	3.934e-4	1.00	9.824e-7	1.99	2.452e-7	2.03
65 536	62 515	2.098e-4	1.00	2.806e-7	2.00	6.840e-8	1.72
131 072	121 301	1.115e-4	–	7.938e-8	–	2.303e-8	–

TABLE 2. Errors and convergence rates for the Galerkin FEM on the modified Shishkin mesh applied to the test problem with $\theta = 2$.

N	$\max\limits_{\substack{k=1,2 \\ \varepsilon\in\mathcal{E}}} M$	$\max\limits_{\varepsilon\in\mathcal{E}} \|\mathbf{u} - \mathbf{u}^N\|_\varepsilon$		$\max\limits_{\varepsilon\in\mathcal{E}} \|\mathbf{u} - \mathbf{u}^N\|_\infty$		$\max\limits_{\varepsilon\in\mathcal{E}} \|\mathbf{u}^I - \mathbf{u}^N\|_\varepsilon$	
		error	rate	error	rate	error	rate
128	156	4.750e-2	1.00	9.378e-3	1.81	3.154e-3	1.85
256	306	2.722e-2	0.99	3.413e-3	1.68	1.119e-3	2.04
512	604	1.537e-2	1.00	1.295e-3	2.04	3.461e-4	1.51
1024	1195	8.537e-3	1.00	3.916e-4	1.89	1.429e-4	2.03
2048	2346	4.700e-3	1.00	1.263e-4	1.95	4.252e-5	2.33
4096	4541	2.565e-3	1.00	3.865e-5	1.95	1.035e-5	1.60
8192	8724	1.390e-3	1.00	1.166e-5	1.98	3.873e-6	2.45
16 384	16 765	7.486e-4	1.00	3.417e-6	1.99	8.488e-7	1.49
32 768	32 768	4.010e-4	1.00	9.860e-7	1.99	3.355e-7	2.07
65 536	65 536	2.139e-4	1.00	2.814e-7	2.00	9.140e-8	1.80
131 072	131 072	1.136e-4	–	7.961e-8	–	2.924e-8	–

the maximum of the number of mesh intervals M over both meshes. Inspecting the third and fourth column of Table 1 we see that the Galerkin finite element method converges of first order which is in total agreement with Theorem 5.2. The successive columns indicate second order convergence in the L$^\infty$-norm and the presence of some supercloseness phenomena, *i.e.* the energy norm rate at which the numerical solution $\mathbf{u}^N$ converges to the nodal interpolant $\mathbf{u}^I$ of the true solution $\mathbf{u}$ exceeds one. Presumably the supercloseness error is a second order term. One might benefit from this property by improving the performance of the method with some adequate post processing operation.

As we are mainly interested in studying the robustness of the method we need errors for various values of $\boldsymbol{\varepsilon}$. Hence, we set

$$\mathcal{E} := \left\{ (\varepsilon_1, \varepsilon_2) \,:\, \varepsilon_k = 10^{-j}, \, k = 1, 2, \, j \in \{0, 1, \ldots, 8\} \right\}.$$

Note that both, $\mathbf{u}$ as well as $\mathbf{u}^N$ depend on $\boldsymbol{\varepsilon}$.

Table 2 has a similar structure as Table 1. However the second column now lists the maximum number of mesh intervals M over both meshes and all $\boldsymbol{\varepsilon} \in \mathcal{E}$. Note that for sufficiently large N this maximal number of mesh intervals appears to converge towards N. Also, in the third, fifth and seventh column only the maximum energy

TABLE 3. Errors and convergence rates for the Galerkin FEM on the standard Shishkin mesh (*i.e.* without weak layer grading) applied to the test problem with $\theta = 2$.

N	$\max\limits_{\varepsilon \in \mathcal{E}} \|\mathbf{u} - \mathbf{u}^N\|_\varepsilon$		$\max\limits_{\varepsilon \in \mathcal{E}} \|\mathbf{u} - \mathbf{u}^N\|_\infty$		$\max\limits_{\varepsilon \in \mathcal{E}} \|\mathbf{u}^I - \mathbf{u}^N\|_\varepsilon$	
	error	rate	error	rate	error	rate
128	4.721e-2	1.00	5.423e-3	1.81	3.142e-3	1.98
256	2.703e-2	1.00	1.975e-3	1.88	1.039e-3	1.89
512	1.522e-2	1.00	6.678e-4	1.93	3.500e-4	1.82
1024	8.455e-3	1.00	2.143e-4	1.83	1.199e-4	1.86
2048	4.651e-3	1.00	7.184e-5	1.83	3.955e-5	1.14
4096	2.537e-3	1.00	2.369e-5	1.09	1.980e-5	0.98
8192	1.374e-3	1.00	1.217e-5	1.31	1.084e-5	1.65
16 384	7.399e-4	1.00	5.392e-6	1.43	3.890e-6	0.82
32 768	3.964e-4	1.00	2.208e-6	0.77	2.326e-6	0.77
65 536	2.114e-4	1.00	1.363e-6	1.26	1.431e-6	1.48
131 072	1.123e-4	–	6.141e-7	–	5.617e-7	–

norm error, maximum pointwise error and maximum supercloseness error over all $\varepsilon \in \mathcal{E}$ is shown, respectively. In agreement with Theorem 5.2 the third and fourth column of Table 2 show uniform first order convergence of the Galerkin FEM on the modified Shishkin meshes, *i.e.* the magnitude of the energy norm error as well as the rate of convergence are independent of ε_1 and ε_2. The fifth and sixth column indicate that the second order convergence of the L^∞ error is uniform. The same can be said for the supercloseness error which is depicted in the seventh column with corresponding rates of convergence in the eighth column. These are still larger than one on the entire horizon of computation.

We continue with Table 3 which has the same structure as Table 2. It shows the performance of the Galerkin FEM on standard Shishkin meshes, *i.e.* we neglect the presence of weak layers and discard the weak layer grading replacing it with a uniform mesh composed of $N/2$ mesh intervals. Consequently, the overall number of mesh elements is N on these meshes independently of ε. On standard Shishkin meshes the Galerkin FEM yields approximate solutions of similar quality if we measure the degree of exactness with the energy norm. The energy norm error of this method on both types of meshes is of the same magnitude. However, the L^∞-error indicates that the weak layer grading is better suited to achieve a high degree of accuracy. In fact, for the problem considered the convergence rate on standard Shishkin meshes breaks down and attains values smaller than 1.5. A similar order reduction can be observed also for the L^2-error and appears to be mentioned in [17] for the first time. In that paper a system of two equations and converse convection with $\varepsilon_1 = \varepsilon_2$ is considered. Our numerical experiments show that in the multiple parameter case a similar order reduction is evident even for aligned convection. The last columns of Table 3 show that ignoring the weak layers is also likely to diminish a supercloseness property with respect to the nodal interpolant in the energy norm.

Next, we consider (6.1) with converse convection, *i.e.* $\theta = -2$. This problem is not covered by our theory and we have to adapt our mesh again. For the second component the mesh needs to be fine at $x = 1$ where the strong boundary layer arises, because the convection in (6.1b) is positive. Moreover, we use weak layer grading at the transition point as well as at the opposite boundary point in both meshes. This approach may not be necessary for the problem studied. However, it will be necessary for problems with more than two equations and multiple (possibly overlapping) weak layers on both sides of the interval $[0, 1]$. Moreover, it follows the spirit of generating a mesh for a certain component independently of the other components which might be useful in an error analysis for this method applied to such problems. The performance of the Galerkin FEM on Shishkin meshes with and without weak layer grading is documented in Tables 4 and 5, respectively. Comparing both tables, we observe similar values for the energy norm error, which is again on both types of meshes of first order. In this respect, Table 4 raises hope for an extension of Theorem 5.2 that includes converse-convection problems.

Table 4. Errors and convergence rates for the Galerkin FEM on the modified Shishkin mesh applied to the test problem with $\theta = -2$.

N	$\max\limits_{\substack{k=1,2 \\ \varepsilon \in \mathcal{E}}} M$	$\max\limits_{\varepsilon \in \mathcal{E}} \|\mathbf{u} - \mathbf{u}^N\|_\varepsilon$		$\max\limits_{\varepsilon \in \mathcal{E}} \|\mathbf{u} - \mathbf{u}^N\|_\infty$		$\max\limits_{\varepsilon \in \mathcal{E}} \|\mathbf{u}^I - \mathbf{u}^N\|_\varepsilon$	
		error	rate	error	rate	error	rate
128	234	4.085e-2	1.00	4.602e-3	1.81	2.569e-3	1.98
256	462	2.339e-2	1.00	1.667e-3	1.89	8.487e-4	1.99
512	916	1.317e-2	1.00	5.619e-4	1.94	2.698e-4	1.98
1024	1826	7.318e-3	1.00	1.800e-4	1.96	8.412e-5	2.00
2048	3480	4.025e-3	1.00	5.562e-5	1.98	2.542e-5	2.01
4096	7300	2.196e-3	1.00	1.675e-5	1.99	7.518e-6	1.79
8192	14618	1.189e-3	–	4.945e-6	–	2.512e-6	–

Table 5. Errors and convergence rates for the Galerkin FEM on the standard Shishkin mesh (*i.e.* without weak layer grading) applied to the test problem with $\theta = -2$.

N	$\max\limits_{\varepsilon \in \mathcal{E}} \|\mathbf{u} - \mathbf{u}^N\|_\varepsilon$		$\max\limits_{\varepsilon \in \mathcal{E}} \|\mathbf{u} - \mathbf{u}^N\|_\infty$		$\max\limits_{\varepsilon \in \mathcal{E}} \|\mathbf{u}^I - \mathbf{u}^N\|_\varepsilon$	
	error	rate	error	rate	error	rate
128	4.080e-2	1.00	4.597e-3	1.81	2.613e-3	1.95
256	2.335e-2	1.00	1.666e-3	1.89	8.786e-4	1.83
512	1.314e-2	1.00	5.618e-4	1.94	3.060e-4	1.84
1024	7.302e-3	1.00	1.800e-4	1.82	1.037e-4	1.93
2048	4.016e-3	1.00	6.067e-5	1.35	3.274e-5	1.21
4096	2.190e-3	1.00	2.682e-5	1.04	1.570e-5	1.27
8192	1.186e-3	1.00	1.416e-5	–	7.204e-6	–

Finally, a comparison of Tables 4 and 5 gives again reason to believe that for robust second order convergence of the Galerkin FEM in the L$^\infty$-norm some mesh adaptation in the weak layer region seems necessary. Yet, without such a mesh adaptation also certain supercloseness phenomena cease to exist. Further insight into this observation and the verification of the uniform second order L$^\infty$ convergence as well as the supercloseness property on the modified Shishkin meshes are topics of future research.

References

[1] M. Bause and K. Schwegler, Higher order finite element approximation of systems of convection-diffusion-reaction equations with small diffusion. *J. Comput. Appl. Math.* **246** (2013) 52–64.

[2] Z. Cen, Parameter-uniform finite difference scheme for a system of coupled singularly perturbed convection-diffusion equations. *J. Syst. Sci. Complex.* **18** (2005) 498–510.

[3] Z. Cen, Parameter-uniform finite difference scheme for a system of coupled singularly perturbed convection-diffusion equations. *Int. J. Comput. Math.* **82** (2005) 177–192.

[4] C. Clavero and J.L. Gracia, An improved uniformly convergent scheme in space for 1D parabolic reaction-diffusion systems. *Appl. Math. Comput.* **243** (2014) 57–73.

[5] R.G. Durán and A.L. Lombardi, Finite element approximation of convection-diffusion problems using graded meshes. *Appl. Numer. Math.* **56** (2006) 1314–1325.

[6] K. Kang and S. Kim, Global pointwise estimates for Green's matrix of second order elliptic systems. *J. Differ. Eqs.* **249** (2010) 2643–2662.

[7] R.B. Kellogg and A. Tsan, Analysis of some difference approximations for a singular perturbation problem without turning points. *Math. Comput.* **32** (1978) 1025–1039.

[8] T. Linß, Analysis of an upwind finite-difference scheme for a system of coupled singularly perturbed convection-diffusion equations. *Computing* **79** (2007) 23–32.

[9] T. Linß, Analysis of a FEM for a coupled system of singularly perturbed reaction-diffusion equations. *Numer. Algor.* **50** (2009) 283–291.

[10] T. Linß, Layer-Adapted Meshes for Reaction-Convection-Diffusion Problems, in vol. 1985 of *Lect. Notes Math.* Springer, Berlin (2010).

[11] T. Linß and N. Madden, A finite element analysis of a coupled system of singularly perturbed reaction-diffusion equations. *Appl. Math. Comput.* **148** (2004) 869–880.

[12] T. Linß and M. Stynes, Numerical solution of systems of singularly perturbed differential equations. *Comput. Meth. Appl. Math.* **9** (2009) 165–191.

[13] N. Madden and M. Stynes, A uniform convergent numerical method for a coupled system of two singularly perturbed linear reaction-diffusion problems. *IMA J. Numer. Anal.* **23** (2003) 627–644.

[14] S. Matthews, E. O'Riordan and G. I. Shishkin, A numerical method for a system of singularly perturbed reaction-diffusion equations. *J. Comput. Appl. Math.* **145** (2002) 151–166.

[15] J. M. Melenk, C. Xenophontos and L. Oberbroeckling, Analytic regularity for a singularly perturbed system of reaction-diffusion equations with multiple scales. *Adv. Comput. Math.* **39** (2013) 367–394.

[16] H.-G. Roos, Robust numerical methods for singularly perturbed differential equations: a survey covering 2008–2012. *ISRN Appl. Math.* (2012). Doi: 10.5402/2012/379547.

[17] H.-G. Roos and C. Reibiger, Numerical analysis of a system of singularly perturbed convection-diffusion equations related to optimal control. *Numer. Math.: Theory Meth. Appl.* **4** (2011) 562–575.

[18] H.-G. Roos, M. Stynes and L. Tobiska, Robust Numerical Methods for Singularly Perturbed Differential Equations. Springer, Berlin (2008).

[19] G.I. Shishkin, Grid approximations of singularly perturbed systems for parabolic convection-diffusion equations with counter-flow. *Sib. Zh. Vychisl. Mat.* **1** (1998) 281–297.

[20] G.I. Shishkin, Approximation of elliptic convection-diffusion equations with parabolic boundary layers. *Zh. Vychisl. Mat. Mat. Fiz.* **40** (2000) 1648–1661.

[21] A. Tamilselvana and N. Ramanujam, A parameter uniform numerical method for a system of singularly perturbed convection-diffusion equations with discontinuous convection coefficients. *Int. J. Comput. Math.* **87** (2010) 1374–1388.