Robust exponential convergence of $hp$ FEM for singularly perturbed reaction–diffusion systems with multiple scales

J. M. Melenk

Institut für Analysis und Scientific Computing, Vienna University of Technology, Wiedner Hauptstrasse 8–10, A-1040 Wien, Austria
melenk@tuwien.ac.at

C. Xenophontos*

Department of Mathematics and Statistics, University of Cyprus, PO Box 20537, Nicosia 1678, Cyprus
*Corresponding author: xenophontos@ucy.ac.cy

AND

L. Oberbroekling

Department of Mathematics and Statistics, Loyola University of Maryland, 4501 N. Charles Street, Baltimore, MD 21210, USA
loberbro@loyola.edu

[Received on 14 April 2012]

We consider a coupled system of two singularly perturbed reaction–diffusion equations in one dimension. Associated with the two singular perturbation parameters $0 < \varepsilon \leq \mu \leq 1$ are boundary layers of length scales $O(\varepsilon)$ and $O(\mu)$. We propose and analyse an $hp$ finite element scheme which includes elements of size $O(\varepsilon^p)$ and $O(\mu^p)$ near the boundary, where $p$ is the degree of the approximating polynomials. We show that under the assumption of analytic input data, the method yields exponential rates of convergence, independently of $\varepsilon$ and $\mu$ and independently of the relative size of $\varepsilon$ to $\mu$. In particular, the full range $0 < \varepsilon \leq \mu \leq 1$ is covered by our analysis. Numerical computations supporting the theory are also presented.

Keywords: boundary layers; singularly perturbed systems; multiple scales; $p$-version FEM; uniform convergence.

1. Introduction

The numerical solution of singularly perturbed problems has been studied extensively over the last couple of decades (see, e.g., the books by Miller et al., 1996; Morton, 1996; Roos et al., 2008 and the references therein). Besides the question of stability of discretizations (e.g., in the treatment of convection-dominated problems), a main difficulty in these problems is the presence of boundary layers in the solution, whose accurate approximation, independently of the singular perturbation parameter(s), is of great importance for the overall quality of the approximate solution. In the context of the finite element method (FEM) and finite difference method (FDM), the robust approximation of boundary layers requires the use of layer-adapted, parameter-dependent meshes (cf. Bakhvalov, 1969; Shishkin, 1989 in an FDM setting); in the present work, a possible choice for the $p$ (and $hp$) version of the FEM...
is termed the *spectral boundary layer mesh* (Definition 3.1) and harks back to Schwab & Suri (1996). (See also Melenk (1997, 2002)).

In this article we consider a system of two coupled, singularly perturbed, linear, reaction–diffusion equations, which have two overlapping boundary layers. In contrast to equations with a single singular perturbation parameter, systems with multiple parameters (and correspondingly multiple layers) are much less studied and understood. The problem under consideration here was studied by Madden & Stynes (2003) and Linß & Madden (2003, 2004b) in the context of finite differences, and by Linß & Madden (2004a) in the context of the $h$ version of the FEM with piecewise linear basis functions. We also refer to Linß & Stynes (2009) for a survey on the numerical solution of systems of singularly perturbed differential equations and to Shishkin (1995) for a finite difference approximation for a problem with multiple scales. In Xenophontos & Oberbroeckling (2007) an $hp$ FEM was presented for a coupled system of reaction–diffusion equations, and its robust exponential convergence was demonstrated via several numerical experiments. The recent regularity results of Melenk et al. (2011) allow us to provide the mathematical justification of what was reported in Xenophontos & Oberbroeckling (2007), which is the purpose of this article. We also refer to Xenophontos & Oberbroeckling (2010) where the case $\varepsilon = \mu$ is analyzed.

The rest of the paper is organized as follows: in Section 2, we present the model problem and discuss the typical phenomena, along with the regularity of the solution as determined in Melenk et al. (2011). In Section 3 we prove our main result, which is the exponential convergence of the proposed FEM, and in Section 4 we present the results of some numerical computations verifying our theoretical findings. Finally, in Section 5 we give some closing remarks.

We will utilize the usual Sobolev space notation $H^k(I)$ to denote the space of functions on $I$ with $0, 1, 2, \ldots, k$ generalized derivatives in $L^2(I)$, equipped with the norm $\| \cdot \|_{k, I}$ and seminorm $| \cdot |_{k, I}$. For vector functions $U := (u_1, u_2)^T$, we will write

$$\| U \|_{k, I}^2 = \| u_1 \|_{k, I}^2 + \| u_2 \|_{k, I}^2.$$

We will also use the space $H^1_0(I) = \{ u \in H^1(I) : u|_{\partial I} = 0 \}$, where $\partial I$ denotes the boundary of $I$. The norm of the space $L^\infty(I)$ of essentially bounded functions is denoted $\| \cdot \|_{\infty, I}$. Finally, the letter $C$ will be used to denote a generic positive constant, independent of any discretization or singular perturbation parameters and possibly having different values in each occurrence.

### 2. The model problem and its regularity

We consider the following model problem: find a pair of functions $(u, v)$ such that

$$\begin{cases}
-\varepsilon^2 u''(x) + a_{11}(x)u(x) + a_{12}(x)v(x) = f(x) & \text{in } I = (0, 1), \\
-\mu^2 v''(x) + a_{21}(x)u(x) + a_{22}(x)v(x) = g(x) & \text{in } I = (0, 1),
\end{cases}$$

(2.1a)

along with the boundary conditions

$$u(0) = u(1) = 0, \quad v(0) = v(1) = 0.$$  

(2.1b)

With the abbreviations

$$U = \begin{pmatrix} u \\ v \end{pmatrix}, \quad E^{\varepsilon, \mu} := \begin{pmatrix} \varepsilon^2 & 0 \\ 0 & \mu^2 \end{pmatrix}, \quad A(x) := \begin{pmatrix} a_{11}(x) & a_{12}(x) \\ a_{21}(x) & a_{22}(x) \end{pmatrix}, \quad F = \begin{pmatrix} f \\ g \end{pmatrix},$$

we have

$$\begin{pmatrix} \varepsilon^2 U'' + E^{\varepsilon, \mu} U' + A(x) U \end{pmatrix} = F.$$
the boundary value problem (2.1a) and (2.1b) may also be written in the following, more compact form:

\[ L_{\varepsilon,a} U := -\varepsilon^2 U''(x) + A(x) U = F, \quad U(0) = U(1) = 0. \]  

(2.2)

The parameters \(0 < \varepsilon \leq \mu \leq 1\) are given, as are the functions \(f, g\) and \(a_{ij}, i,j \in \{1, 2\}\), which are assumed to be analytic on \(I = [0, 1]\). Moreover we assume that there exist constants \(C_f, \gamma_f, C_g, \gamma_g, C_a, \gamma_a > 0\) such that

\[
\begin{align*}
\|f^{(n)}\|_{\infty, I} \leq C_f \gamma_f^n n! & \quad \forall n \in \mathbb{N}_0, \\
\|g^{(n)}\|_{\infty, I} \leq C_g \gamma_g^n n! & \quad \forall n \in \mathbb{N}_0, \\
\|a_{ij}^{(n)}\|_{\infty, I} \leq C_a \gamma_a^n n! & \quad \forall n \in \mathbb{N}_0, \quad i,j \in \{1, 2\}.
\end{align*}
\]

(2.3)

The variational formulation of (2.1a) and (2.1b) reads: find \(U := (u, v) \in [H^1_0(I)]^2\) such that

\[ B(U, V) = F(V) \quad \forall V := (\bar{u}, \bar{v}) \in [H^1_0(I)]^2, \]

(2.4)

where, with \((\cdot, \cdot)\) being the usual \(L^2(I)\) inner product,

\[
B(U, V) = \varepsilon^2 (u', \bar{u}')_I + \mu^2 (v', \bar{v}')_I + (a_{11} u + a_{12} v, \bar{u})_I + (a_{21} u + a_{22} v, \bar{v})_I, \\
F(V) = (f, \bar{u})_I + (g, \bar{v})_I.
\]

(2.5)

(2.6)

The matrix-valued function \(A\) is assumed to be pointwise positive definite (but not necessarily symmetric), i.e., for some fixed \(\alpha > 0\),

\[
\xi^T A(x) \xi \geq \alpha^2 \xi^T \xi \quad \forall \xi \in \mathbb{R}^2, \quad \forall x \in \bar{I}.
\]

(2.7)

It follows that the bilinear form \(B(\cdot, \cdot)\) given by (2.5) is coercive with respect to the energy norm

\[
\|U\|_{E,I}^2 := (u, v)_{_{E,I}}^2 := \varepsilon^2 \|u\|_{_{1,I}}^2 + \mu^2 \|v\|_{_{1,I}}^2 + \alpha^2 (\|u\|_{_{0,I}}^2 + \|v\|_{_{0,I}}^2),
\]

(2.8)

i.e.,

\[
B(V, V) \geq \|V\|_{E,I}^2 \quad \forall V \in [H^1_0(I)]^2.
\]

(2.9)

This, along with the continuity of \(B(\cdot, \cdot)\) and \(F(\cdot)\), implies the unique solvability of (2.4). We also have the following standard \textit{a priori} estimate:

\[
\|U\|_{E,I} \leq \alpha^{-1} \sqrt{\|f\|_{0,I}^2 + \|g\|_{0,I}^2}.
\]

(2.10)

The finite element approximation of (2.1a) and (2.1b) reads: find \(U_N := (u_N, v_N)^T \in [S_N]^2 \subset [H^1_0(I)]^2\) such that

\[ B(U_N, V) = F(V) \quad \forall V := (\bar{u}, \bar{v})^T \in [S_N]^2, \]

(2.11)

where \([S_N]^2\) is an appropriately chosen finite-dimensional subspace of \([H^1_0(I)]^2\). The unique solvability of the discrete problem (2.11) follows from (2.7), and by the well-known Galerkin orthogonality we have

\[
\|U - U_N\|_{E,I} \leq \alpha^{-1} \inf_{V \in [S_N]^2} \|U - V\|_{E,I}.
\]

(2.12)

As is well known for singularly perturbed problems, the choice of the space \(S_N\) must be made carefully and in a way that depends on the layer structure of the solution \(U\) in order for the approximation to be \textit{robust}, i.e., for convergence to be independent of \(\varepsilon\) and \(\mu\). As we will formalize in Theorem 2.2, the
solution \( U \) has features on up to three different length scales, namely, \( \mathcal{O}(1) \), \( \mathcal{O}(\mu) \) and \( \mathcal{O}(\varepsilon) \), with the features on the \( \mathcal{O}(\varepsilon) \) and \( \mathcal{O}(\mu) \) scales being of a boundary layer type. These three different length scales have to be incorporated into the approximation space, and we will do this with the spectral boundary layer mesh below in Definition 3.1.

Our design of the spectral boundary layer mesh hinges on the regularity theory of Melenk et al. (2011), which we will discuss in more detail in Theorem 2.2 ahead. Essentially, Theorem 2.2 derives from asymptotic expansions of the solution. Such expansions rely on scale separation assumptions. For the present context of length scales \( \mathcal{O}(1) \), \( \mathcal{O}(\mu) \) and \( \mathcal{O}(\varepsilon) \), the following cases may occur.

(I) The ‘no-scale separation case’ which occurs when neither \( \mu/1 \) nor \( \varepsilon/\mu \) is small.

(II) The ‘three-scale case’ in which all scales are separated and which occurs when \( \mu/1 \) is small and \( \varepsilon/\mu \) is small.

(III) The first ‘two-scale case’ which occurs when \( \mu/1 \) is not small and \( \varepsilon/\mu \) is small.

(IV) The second ‘two-scale case’ which occurs when \( \mu/1 \) is small and \( \varepsilon/\mu \) is not small.

The concept of ‘small’ (or ‘not small’) mentioned above is tied in two ways to the regularity theory in terms of asymptotic expansions. First, on the level of constructing asymptotic expansions, the decision about which parameters are deemed small determines the ansatz to be made and thus the form of the expansion. Second, on the level of using asymptotic expansions for approximation purposes or the design of approximation spaces, the decision about which parameters are deemed small depends on the desired accuracy, i.e., whether the remainder resulting from the asymptotic expansion can be regarded as small. We mention that a similar classification of the possible relations between the arising scales has already been used in Shishkin (1995) for the design of a robust finite difference scheme with multiple scales.

In order to be able to describe the regularity assertions for the solution \( U \), we need to introduce some notation.

**Definition 2.1** 1. We say that a function \( w \) is analytic with length scale \( \nu > 0 \) (and analyticity parameters \( C_w, \gamma_w \)), abbreviated \( w \in \mathcal{A}(\nu, C_w, \gamma_w) \), if

\[
\|w^{(n)}\|_{\infty, I} \leq C_w \gamma_w^n \max\{n, \nu^{-1}\}^n \quad \forall n \in \mathbb{N}_0.
\]

2. We say that an entire function \( w \) is of \( L^\infty \) boundary layer type with length scale \( \nu > 0 \) (and analyticity parameters \( C_w, \gamma_w \)), abbreviated \( w \in \mathcal{B} L^\infty(\nu, C_w, \gamma_w) \), if for all \( x \in I \)

\[
|w^{(n)}(x)| \leq C_w \gamma_w^n \nu^{-n} e^{-\text{dist}(x, \partial I)/\nu} \quad \forall n \in \mathbb{N}_0.
\]

Both definitions extend naturally to vector-valued functions by requiring that the above bounds hold componentwise.

With this notation in hand, we can formulate a regularity assertion for the solution \( U \) of (2.1a) and (2.1b).

**Theorem 2.2** (Melenk et al. 2011, Theorem 2.2) Assume (2.3) and (2.7) hold. Then there exist constants \( C, b, \delta, q, \gamma > 0 \), independent of \( 0 < \varepsilon \leq \mu \leq 1 \), such that the following assertions are true for the solution \( U \) of (2.1a) and (2.1b).

(I) \( U \in \mathcal{A}(\varepsilon, C\varepsilon^{-1/2}, \gamma) \).
(II) \( U \) can be written as \( U = W + \tilde{U}_{BL} + \hat{U}_{BL} + R \), where \( W \in \mathcal{A}(1, C, \gamma) \), \( \tilde{U}_{BL} \in \mathcal{B}L^\infty(\delta \mu, C, \gamma) \), \( \hat{U}_{BL} \in \mathcal{B}L^\infty(\delta \varepsilon, C, \gamma) \) and \( ||R||_{L^\infty(\partial I)} + ||R||_{E,I} \leq C[\varepsilon^{-b/\mu} + \varepsilon^{-b/\mu}] \). Furthermore, the second component \( \hat{v} \) of \( \hat{U}_{BL} \) satisfies the stronger assertion \( \hat{v} \in \mathcal{B}L^\infty(\delta \varepsilon, C(\varepsilon/\mu)^2, \gamma) \).

(III) If \( \varepsilon/\mu \leq q \) then \( U \) can be written as \( U = W + \tilde{U}_{BL} + R \), where \( W \in \mathcal{A}(\mu, C, \gamma) \), \( \tilde{U}_{BL} \in \mathcal{B}L^\infty(\delta \varepsilon, C, \gamma) \) and \( ||R||_{L^\infty(\partial I)} + ||R||_{E,I} \leq C\varepsilon^{-b/\mu} \). Furthermore, the second component \( \hat{v} \) of \( \tilde{U}_{BL} \) satisfies the stronger assertion \( \hat{v} \in \mathcal{B}L^\infty(\delta \varepsilon, C(\varepsilon/\mu)^2, \gamma) \).

(IV) \( U \) can be written as \( U = W + \tilde{U}_{BL} + R \), where \( W \in \mathcal{A}(1, C, \gamma) \), \( \tilde{U}_{BL} \in \mathcal{B}L^\infty(\delta \mu, C\sqrt{\mu/\varepsilon}, \gamma \mu/\varepsilon) \), and \( ||R||_{L^\infty(\partial I)} + ||R||_{E,I} \leq C(\mu/\varepsilon)^2 \varepsilon^{-b/\mu} \).

We emphasize that all decompositions of \( U \) stated in (I)–(IV) are valid for the solution \( U \) regardless of the values of \( 0 < \varepsilon \leq \mu \leq 1 \) (caveat: in the form presented here, case III requires \( \varepsilon/\mu \leq q \)); in fact, we will exploit this observation in the proof of Theorem 3.3. Nevertheless, given that the remainder \( R \) is only small in certain parameter regimes (and hence, only then does the corresponding decomposition provide useful information) it may be helpful to think of cases I–IV as being associated with parameter regimes (I)–(IV) mentioned earlier. These ties are further strengthened by how the decompositions of Theorem 2.2 are obtained in Melenk et al. (2011), namely, by asymptotic expansions. In the subsections that follow we will briefly explain what the ansatz is in each case II–IV.

2.1 The three-scale case: case II

Anticipating that boundary layers of length scales \( O(\mu) \) and \( O(\varepsilon) \) will appear at the end points \( x = 0 \) and \( 1 \), we introduce the stretched variables \( \tilde{x} = x/\mu \), \( \tilde{\varepsilon} = x/\varepsilon \) for the expected layers at the left-hand end point \( x = 0 \) and variables \( \tilde{x}^R = (1 - x)/\mu \), \( \tilde{\varepsilon}^R = (1 - x)/\varepsilon \) for the expected behaviour at the right-hand end point \( x = 1 \). We make the following formal ansatz for the solution \( U \):

\[
U \sim \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \left( \frac{\mu}{1} \right)^i \left( \frac{\varepsilon}{\mu} \right)^j [U_{ij}(\tilde{x}) + \tilde{U}_{ij}(\tilde{\varepsilon}) + \hat{U}_{ij}(\tilde{x}) + \hat{U}_{ij}(\tilde{\varepsilon})],
\]

(2.13)

where the functions \( U_{ij}, \tilde{U}_{ij}, \hat{U}_{ij}, \hat{U}_{ij}^R \) are to be determined by inserting the ansatz (2.13) into the boundary value problem (2.1a) and (2.1b), and equating like powers of \( \mu/1 \) and \( \varepsilon/\mu \). The functions \( U_{ij}, \tilde{U}_{ij}, \hat{U}_{ij}, \hat{U}_{ij}^R \) can then be determined recursively as solutions of suitable boundary value problems (see Melenk et al., 2011 for details). The decomposition of Theorem 2.2 is obtained by truncating the asymptotic expansion (2.13) after a finite number of terms:

\[
U(x) = W_M(x) + \tilde{U}^M_{BL}(\tilde{x}) + \hat{U}^M_{BL}(\tilde{\varepsilon}) + \tilde{V}^M_{BL}(\tilde{x}) + \hat{V}^M_{BL}(\tilde{\varepsilon}) + R_M(x),
\]

(2.14)

where

\[
W_M(x) = \sum_{i=0}^{M_1} \sum_{j=0}^{M_2} \mu^i \left( \frac{\varepsilon}{\mu} \right)^j U_{ij}(x)
\]

(2.15)

denotes the outer (smooth) expansion,

\[
\tilde{U}^M_{BL}(\tilde{x}) = \sum_{i=0}^{M_1} \sum_{j=0}^{M_2} \mu^i \left( \frac{\varepsilon}{\mu} \right)^j \tilde{U}_{ij}^M(\tilde{x}), \quad \hat{U}^M_{BL}(\tilde{\varepsilon}) = \sum_{i=0}^{M_1} \sum_{j=0}^{M_2} \mu^i \left( \frac{\varepsilon}{\mu} \right)^j \hat{U}_{ij}^M(\tilde{\varepsilon})
\]

(2.16)
denote the left- and right-hand inner (boundary layer) expansions associated with the variables $\hat{x}$ and $\hat{x}^R$, respectively,

$$
\widehat{U}^M_{\text{BL}}(\hat{x}) = \sum_{i=0}^{M_1} \sum_{j=0}^{M_2} \mu^i \left( \frac{\varepsilon}{\mu} \right)^j \widehat{U}^L_{ij}(\hat{x}), \quad \widehat{V}^M_{\text{BL}}(\hat{x}^R) = \sum_{i=0}^{M_1} \sum_{j=0}^{M_2} \mu^i \left( \frac{\varepsilon}{\mu} \right)^j \widehat{V}^R_{ij}(\hat{x}^R)
$$

(2.17)

denote the left- and right-hand inner (boundary layer) expansions associated with the variables $\hat{x}$ and $\hat{x}^R$, respectively, and the remainder $R_M$ is defined such that (2.14) holds. In establishing Theorem 2.2 the choices $M_1 = O(1/\mu)$, $M_2 = O(\mu/\varepsilon)$ are made. (For full details, see Melenk et al., 2011.)

### 2.2 The first two-scale case: case III

We employ again the notation of stretched variables $\hat{x} = x/\varepsilon$ and $\hat{x}^R = (1 - x)/\varepsilon$. Since $\mu/1$ is not assumed to be small, only the scales $O(1)$ and $O(\varepsilon)$ are expected to be present in the problem. Inserting the formal ansatz

$$
U(x) \sim \sum_{i=0}^{\infty} \left( \frac{\varepsilon}{\mu} \right)^i \left[ U_i(x) + \hat{U}^L_i(\hat{x}) + \hat{U}^R_i(\hat{x}^R) \right]
$$

(2.18)

into the boundary value problem (2.1a) and (2.1b) and equating like powers of $\varepsilon/\mu$ yields recursions for the functions $U_i$, $\hat{U}^L_i$ and $\hat{U}^R_i$. Truncating (2.18) after $M$ terms leads to the representation

$$
U(x) = W_M(x) + \hat{U}^M_{\text{BL}}(\hat{x}) + \hat{V}^M_{\text{BL}}(\hat{x}^R) + R_M(x),
$$

(2.19)

with outer (smooth) expansion $W_M$, left- and right-hand inner (boundary layer) expansions $\hat{U}^M_{\text{BL}}$ and $\hat{V}^M_{\text{BL}}$, and a remainder $R_M$ such that (2.19) is valid. In establishing Theorem 2.2 the choice $M = O(\mu/\varepsilon)$ is made (see Melenk et al., 2011 for more details).

### 2.3 The second two-scale case: case IV

In this case, $\mu/1$ is assumed to be small and $\varepsilon/\mu$ is not deemed small, which leads us to the ansatz

$$
U \sim \sum_{i=0}^{\infty} \mu^i \left[ U_i(x) + \hat{U}^L_i(\hat{x}) + \hat{U}^R_i(\hat{x}^R) \right],
$$

(2.20)

with the stretched variables $\hat{x} = x/\mu$ and $\hat{x}^R = (1 - x)/\mu$. Inserting this ansatz into the boundary value problem (2.1a) and (2.1b) and equating like powers of $\mu$ yields recursions for the functions $U_i$, $\hat{U}^L_i$ and $\hat{U}^R_i$. The truncated series $W_M(x) := \sum_{i=0}^{M_0} \mu^i U_i(x)$, $\hat{U}^M_{\text{BL}} := \sum_{i=0}^{M_0} \mu^i \hat{U}^L_i(\hat{x})$ and $\hat{V}^M_{\text{BL}} := \sum_{i=0}^{M_0} \mu^i \hat{U}^R_i(\hat{x}^R)$ yield the decomposition stated in Theorem 2.2 if $M = O(1/\mu)$. (The details are given in Melenk et al., 2011.)
3. Approximation results

3.1 Main results

In this section we will describe the finite-dimensional subspace \([SN]^2\), which appears in (2.11), in order to construct an \(hp\) scheme for the approximation of the solution to (2.4). To this end, let \(\Delta = \{0 = x_0 < x_1 < \cdots < x_M = 1\}\) be an arbitrary partition of \(I = (0, 1)\) and set

\[I_j = (x_{j-1}, x_j), \quad h_j = x_j - x_{j-1}, \quad j = 1, \ldots, M.\]

Also, define the master (or standard) element \(I_{ST} = (-1, 1)\), and note that it can be mapped onto the \(j\)th element \(I_j\) by the linear mapping

\[x = Q_j(t) = \frac{1}{2}(1 - t)x_{j-1} + \frac{1}{2}(1 + t)x_j.\]

With \(\Pi_p(I_{ST})\) the space of polynomials of degree \(\leq p\) on \(I_{ST}\), we define our finite-dimensional subspace as following (with \(\circ\) denoting composition of functions):

\[S_N \equiv S^p(\Delta) = \{V \in [H^1_0(I)]^2 : V \circ Q_j^{-1} \in (\Pi_p(I_{ST}))^2, \quad j = 1, \ldots, M\},
\]

\[S^p_0(\Delta) := S^p(\Delta) \cap [H^1_0(I)]^2.\]  \hfill (3.1)

We restrict our attention here to constant polynomial degree \(p\) for all elements, i.e., \(p_j = p, j = 1, \ldots, M\); however, clearly, more general settings with variable polynomial degree are possible.

The following spectral boundary layer mesh is essentially the minimal mesh that yields robust exponential convergence. Loosely speaking, one inserts nodes to resolve the boundary layers, i.e., upon setting \(x_\varepsilon := \kappa p \varepsilon\) and \(x_\mu = \kappa p \mu\), one inserts the nodes \(x_\varepsilon\) and \(1 - x_\varepsilon\) if \(\kappa p \varepsilon < 1/2\); the nodes \(x_\mu\) and \(1 - x_\mu\) are inserted if \(\kappa p \mu < 1/2\). Here, \(\kappa > 0\) is a user-specified parameter. The optimal choice of \(\kappa\) is problem specific as it depends on the length scales of the boundary layers, and we refer the reader to Schwab & Suri (1996) for a more detailed discussion of this issue.

**Definition 3.1** (Spectral boundary layer mesh) For \(\kappa > 0, p \in \mathbb{N}\) and \(0 < \varepsilon \leq \mu \leq 1\), define the spaces \(S(\kappa, p)\) of piecewise polynomials by

\[S(\kappa, p) := \begin{cases} S^p_0(\Delta) ; \Delta = \{0, \kappa p \varepsilon, \kappa p \mu, 1 - \kappa p \mu, 1 - \kappa p \varepsilon, 1\} & \text{if } \kappa p \mu < \frac{1}{2}, \\ S^p_0(\Delta) ; \Delta = \{0, \kappa p \varepsilon, 1 - \kappa p \varepsilon, 1\} & \text{if } \kappa p \varepsilon < \frac{1}{2} \leq \kappa p \mu, \\ S^p_0(\Delta) ; \Delta = \{0, 1\} & \text{if } \kappa p \varepsilon \geq \frac{1}{2}. \end{cases}\]

**Remark 3.2** Strictly speaking, the method considered here is not a true \(hp\) version, since the location (and not the number) of mesh points changes as \(p\) increases. A better characterization would be a \(p\)-version FEM on a moving mesh. Obviously, additional refinement and/or using a true \(hp\) version would yield better results but at the cost of using more degrees of freedom—see Xenophontos & Oberbroeckling (2007) for a numerical comparison.

We now present the main result of this paper.

**Theorem 3.3** Let \(f, g\) and \(A\) be analytic on \(I\) and satisfy the conditions in (2.3) and (2.7). Let \(U = (u, v)^T\) be the solution to (2.1a) and (2.1b). Then there exist constants \(\kappa_0, C, \beta > 0\) depending only on \(f, g\) and...
such that for each \( \kappa \in (0, \kappa_0] \) there is an \( \tilde{\mathcal{J}}_p \mathbf{U} = [\mathcal{J}_p u, \mathcal{J}_p v]^T \in S(\kappa, p) \) with \( \tilde{\mathcal{J}}_p \mathbf{U} = \mathbf{U} \) on \( \partial I \) and
\[
\| \mathbf{U} - \tilde{\mathcal{J}}_p \mathbf{U} \|_{E,I} \leq C e^{-\beta \kappa p}.
\]

**Proof.** The proof is given at the end of Section 3.2. \( \square \)

Using the above theorem and the quasi-optimality result (2.12) we have the following corollary.

**Corollary 3.4** Let \( \mathbf{U} \) be the solution to (2.4) and let \( \mathbf{U}_N \in S_0^p(\Delta) \) be the solution to (2.11) based on the spectral boundary layer mesh of Definition 3.1. Then there exist constants \( \kappa_0, C, \beta > 0 \) depending only on the input data \( f, g \) and \( A \), such that for any \( 0 < \kappa \leq \kappa_0 \),
\[
\| \mathbf{U} - \mathbf{U}_N \|_{E,I} \leq C e^{-\beta \kappa p}.
\]

### 3.2 Proof of Theorem 3.3

#### 3.2.1 An approximation operator

**Lemma 3.5** There exists a bounded linear operator \( \mathcal{J}_p : H^1(I_{ST}) \to \Pi_p(I_{ST}) \) satisfying \( u(\pm 1) = \mathcal{J}_p u(\pm 1) \) with the following approximation properties: if \( u \in C^\infty(I_{ST}) \), then
\[
\| u - \mathcal{J}_p u \|_{0,I_{ST}}^2 \leq \frac{1}{p^2} \frac{(p-s)!}{(p+s)!} \| u^{(s+1)} \|_{0,I_{ST}}^2 \quad \forall s = 0, 1, \ldots, p, \tag{3.2}
\]
\[
\| (u - \mathcal{J}_p u') \|_{0,I_{ST}}^2 \leq \frac{(p-s)!}{(p+s)!} \| u^{(s+1)} \|_{0,I_{ST}}^2 \quad \forall s = 0, 1, \ldots, p. \tag{3.3}
\]

**Proof.** The result is taken from Schwab (1998, Corollary 3.15). \( \square \)

**Lemma 3.6** Let \( p \in \mathbb{N}, \lambda \in (0, 1] \) such that \( \lambda p \in \mathbb{N} \). Then
\[
\frac{(p-\lambda p)!}{(p+\lambda p)!} \leq \left[ \frac{(1-\lambda)^{(1-\lambda)\gamma} + \gamma}{(1+\lambda)^{(1+\lambda)\gamma} + \gamma} \right] p^{-2\lambda p} e^{2\lambda p+1}.
\]

**Proof.** This follows from Stirling’s formula—see Xenophontos & Oberbroeckling (2007, Lemma 3.1) for the details. \( \square \)

On the reference element \( I_{ST} \), we have the following additional stability results.

**Lemma 3.7** Let \( \mathcal{J}_p : H^1(I_{ST}) \to \Pi_p(I_{ST}) \) be as in Lemma 3.5. Then, on the reference element \( I_{ST} \), we have
\[
| \mathcal{J}_p u |_{1,I_{ST}} \leq C | u |_{1,I_{ST}}, \quad \| \mathcal{J}_p u \|_{0,I_{ST}} \leq \| u \|_{0,I_{ST}} + C \frac{1}{p} | u |_{1,I_{ST}},
\]
\[
\| \mathcal{J}_1 u \|_{L^\infty(I_{ST})} \leq \| u \|_{L^\infty(I_{ST})}, \quad \| (\mathcal{J}_1 u') \|_{L^\infty(I_{ST})} \leq C \| u \|_{L^\infty(I_{ST})}.
\]

**Proof.** The estimates for the linear interpolant \( \mathcal{J}_1 \) are standard. The stability in the \( H^1 \) seminorm is a direct consequence of (3.3) with \( s = 0 \) and the triangle inequality. The \( L^2 \) stability follows from the triangle inequality and an approximation result:
\[
\| \mathcal{J}_p u \|_{0,I_{ST}} \leq \| u \|_{0,I_{ST}} + \| u - \mathcal{J}_p u \|_{0,I_{ST}} \leq \| u \|_{0,I_{ST}} + C p^{-1} | u |_{1,I_{ST}}.
\]
\( \square \)
The approximation operator \( \mathcal{I}_p \) on the reference element can be used to define an approximation operator in an elementwise fashion: for a mesh \( \Delta \) with elements \( I_j \), \( j = 1, \ldots, \mathcal{M} \), element maps \( Q_j \) and given degree vector \( \overrightarrow{p} = (p_1, \ldots, p_{\mathcal{M}}) \), with \( 1 \leq p_j \leq p \), we define the operator \( \mathcal{I}_{\overrightarrow{p}, \Delta} : [H^1(I)]^2 \rightarrow S^p(\Delta) \) elementwise in the standard way with the operators \( \mathcal{I}_{p_j} \), by requiring
\[
(\mathcal{I}_{\overrightarrow{p}, \Delta} \mathbf{V})|_{I_j} \circ Q_j = \mathcal{I}_{p_j}(\mathbf{V}|_{I_j} \circ Q_j), \quad j = 1, \ldots, \mathcal{M}.
\] (3.4)

Since the operators \( \mathcal{I}_{p_j} \) interpolate at the end points of \( I_{\mathcal{ST}} \), this operator is indeed well defined. We will write \( \mathcal{I}_{p, \Delta} \) if \( p_j = p \) for \( j = 1, \ldots, \mathcal{M} \). We point out that \( p_j = 1 \) corresponds to the linear interpolant. Finally, since the operators \( \mathcal{I}_{\overrightarrow{p}, \Delta} \) are defined elementwise, we will work with the abbreviation
\[
\mathcal{I}_{p_j}(\mathbf{V}) := (\mathcal{I}_{\overrightarrow{p}, \Delta} \mathbf{V})|_{I_j} = \mathcal{I}_{p_j}(\mathbf{V}|_{I_j} \circ Q_j)) \circ Q_j^{-1}.
\] (3.5)

The following approximation result on the reference element will be one of our main tools for the proof of Theorem 3.3.

**Lemma 3.8** Let \( u \in C^\infty(I_{\mathcal{ST}}) \) satisfy, for some \( C_u, \gamma_u > 0 \), \( K \geq 1 \), \( h \in (0, 1] \),
\[
\|u^{(n)}\|_{0,I_{\mathcal{ST}}} \leq C_u(\gamma_u h)^n \max\{n,K\}^n \quad \forall n \in \mathbb{N}.
\] (3.6)

Then there exist \( \eta, \beta, C > 0 \) depending solely on \( \gamma_u \), such that under the condition
\[
\frac{hK}{p} \leq \eta,
\] (3.7)
the approximation \( \mathcal{I}_p u \in \Pi_p(I_{\mathcal{ST}}) \) given by Lemma 3.5 satisfies \( u(\pm 1) = (\mathcal{I}_p u)(\pm 1) \) and
\[
\|u - \mathcal{I}_p u\|_{0,I_{\mathcal{ST}}} + \|(u - \mathcal{I}_p u)'\|_{0,I_{\mathcal{ST}}} \leq CC_u \frac{hK}{p} e^{-\beta p}.
\]

**Proof.** First note that \( s \geq 1, (j + 1)/j \leq 2 \). Then, in view of \( hK/p \leq \eta, h \leq 1 \) and \( \lambda \leq 1 \), we compute for \( s = \lambda p \) the following:
\[
\begin{align*}
    h^{2(s+1)} (\max\{s+1,K\})^{2(s+1)} &\leq e^2 h^2 \max\{s+1,K\}^2 (\max\{hs,hK\})^{2s} \\
    &\leq e^2 h^2 \max\{K,p+1\}^2 (\max\{h\lambda p,\eta p\})^{2s} \\
    &\leq e^2 h^2 (2Kp)^2 (p \max\{\lambda,\eta\})^{2s}.
\end{align*}
\]

This and Lemma 3.6 give
\[
\begin{align*}
    \frac{(p-s)!}{(p+s)!} \|u^{(s+1)}\|_{0,I_{\mathcal{ST}}}^2 &\leq 4 e^2 (hKp)^2 \gamma_u^2 C_u^2 \gamma_u^{2\lambda p} (p \max\{\lambda,\eta\})^{2\lambda p} (1 - \lambda)^{1-\lambda} (1 + \lambda)^{1+\lambda} \left[ \frac{(1 - \lambda)^{1-\lambda}}{(1 + \lambda)^{1+\lambda}} \right]^p p^{-2\lambda p} e^{2\lambda p+1} \\
    &\leq 4 C_u^3 e^3 (hKp)^2 \left[ \frac{(1 - \lambda)^{1-\lambda}}{(1 + \lambda)^{1+\lambda}} \right]^p [\gamma_u e \max\{\lambda,\eta\}]^{2\lambda p}.
\end{align*}
\]
Select now \( \lambda \in (0, 1) \) and \( \eta > 0 \) such that \( \gamma_u e \max\{\lambda, \eta\} \leq 1. \) Since for this choice of \( \lambda \) we have \( (1 - \lambda)^{1 - \lambda}/(1 + \lambda)^{1 + \lambda} =: q < 1, \) we conclude

\[
\frac{(p - s)!}{(p + s)!} \|u^{(s+1)}\|^2_{0,I_J} \leq 4e^3 \gamma_u^2 (hKp)^2 e q^p = 4e^3 \gamma_u^2 \left( \frac{hK}{p} \right)^2 C_u^2 p^4 q^p,
\]

which is the desired bound, since the algebraic factor \( p^4 \) may be absorbed in the exponentially decaying one by suitably adjusting the constants. □

We reformulate the approximation result of Lemma 3.8 in a form that will be convenient for the approximation of the smooth and the boundary layer parts of the expansion (the latter within the layer).

**Corollary 3.9** Let \( I_j \) be an interval of length \( h_j \), and let \( V \in C^\infty(I_j) \) satisfy for some \( C_u, \gamma_u > 0, K \geq 1, \)

\[
\|V^{(n)}\|_{L^\infty(I_j)} \leq C_u \gamma_u^n \max\{n, K\}^n \quad \forall n \in \mathbb{N}.
\]

Then there exist constants \( C, \eta, \beta > 0 \) depending only on \( \gamma_u \), such that under the scale resolution condition

\[
h_j K \leq \eta, \tag{3.8}
\]

the polynomial approximation \( \mathcal{I}_{p_j, I_j} V \) satisfies

\[
h_j^{-1} \|V - \mathcal{I}_{p_j, I_j} V\|_{0,I_J} + \|V - \mathcal{I}_{p_j, I_j} V\|_{1,I_J} \leq CC_u \frac{h_j^{1/2} K}{p_j} e^{-\beta p_j}. \tag{3.9}
\]

**Proof.** Let \( \hat{V} := V \circ Q_j, \) where \( Q_j : I_{ST} \to I_j \) is an affine bijection. Then \( \hat{V} \) satisfies

\[
\|\hat{V}^{(n)}\|_{L^2(I_{ST})} \leq CC_u (\gamma_u h_j/2)^n \max\{n, K\}^n \quad \forall n \in \mathbb{N}.
\]

Therefore, Lemma 3.8 gives the existence of \( C, \beta \) and \( \eta \) such that under assumption (3.8) we have

\[
\|\hat{V} - \mathcal{I}_{p_j, I} \hat{V}\|_{0,I_{ST}} + \|(\hat{V} - \mathcal{I}_{p_j, I} \hat{V})\|_{0,I_{ST}} \leq CC_u \frac{h_j K}{p_j} e^{-\beta p_j}.
\]

Transforming back to \( I_j \) gives the result. □

The following result will be useful for the approximation of the remainder \( R \) and the boundary layer contributions.

**Lemma 3.10** Let \( I_j \) be an interval of length \( h_j \) and \( p_j \in \mathbb{N}. \) Then, for scalar functions (and analogously for vector-valued ones):

\[
\|u - \mathcal{I}_{p_j, I_j} u\|_{0,I_J} \leq C \frac{h_j}{p_j} \|u\|_{1,I_J}, \tag{3.10}
\]

\[
\|u - \mathcal{I}_{p_j, I_j} u\|_{1,I_J} \leq C \|u\|_{1,I_J}, \tag{3.11}
\]

\[
\|u - \mathcal{I}_{1,I_j} u\|_{0,I_j} \leq \|u\|_{0,I_j} + C h_j^{1/2} \|u\|_{L^\infty(I_j)}. \tag{3.12}
\]

**Proof.** The estimates follow from Lemmas 3.5, 3.7 and standard scaling arguments. □
Finally, we formulate a (repetition of approximation) result for the approximation of functions of boundary-layer type outside the layer.

**Lemma 3.11** Let $\nu > 0$ and let $u$ satisfy

\[ |u(x)| + \nu |u'(x)| \leq C_u e^{-\text{dist}(x, \partial I)/\nu} \quad \forall x \in I. \]

Let $\Delta$ be an arbitrary mesh on $I$ with mesh points $\xi$ and $1 - \xi$, where $\xi \in (0, 1/2)$. Then, for some $C > 0$ independent of $\nu$, on $(\xi, 1 - \xi)$ the piecewise linear interpolant $I_{1,\Delta} u$ satisfies

\[ \nu |u - I_{1,\Delta} u|_{1,\xi,1-\xi} + \|u - I_{1,\Delta} u\|_{0,\xi,1-\xi} \leq CC_u e^{-\xi/\nu}. \]

**Proof.** Let $I_j$ be an element in the interval $(\xi, 1 - \xi)$ of length $h_j$. We distinguish between the cases $h_j \leq \nu$ and $h_j > \nu$. In the case $h_j \leq \nu$, we note that (3.10) and (3.11) yield

\[ \|u - I_{1,I_j} u\|_{0,I_j} + \nu\|u - I_{1,I_j} u\|_{0,1-I_j} \leq C(h_j + \nu)|u_{1,I_j}| \leq Cv|u_{1,I_j}|. \]

In the converse case $h_j \geq \nu$, we use (3.11) and (3.12) to get

\[ \|u - I_{1,I_j} u\|_{0,I_j} + \nu\|u - I_{1,I_j} u\|_{0,1-I_j} \leq C[h_j^{1/2}\|u\|_{L^\infty(I_j)} + \nu|u_{1,I_j}|]. \]

Summation over all elements in $(\xi, 1 - \xi)$ then gives the result. \(\square\)

### 3.2.2 Details of the proof of Theorem 3.3.

The approximation of the exact solution $U$ is constructed element by element with the aid of the operator $I_p$. The basic ingredient of the proof is that the four regularity assertions of Theorem 2.2 permit us to show that, on each element, the features on all length scales can either be resolved or are sufficiently small to be safely ignored. The parameter $\kappa_0$ appearing in the statement of Theorem 3.3 will be determined in the course of the proof.

We remark here that we make the simplifying assumption that

\[ \frac{\varepsilon}{\mu} \leq q \quad (3.13) \]

with $q$ as in Theorem 2.2. In the converse case, $\mu$ and $\varepsilon$ are comparable, and the regularity assertion (III) follows from (IV) by suitably adjusting constants.

**Case 1:** $\kappa p \varepsilon \geq 1/2$, which implies in particular that $\kappa p \mu \geq 1/2$ and $\kappa p \varepsilon / \mu \geq 1/2$. This is the no-scale separation case (from the point of view of regularity) or, equivalently, the ‘asymptotic case of $p$’ (from the point of view of approximability). The mesh is given by $\Delta = \{0, 1\}$.

We employ the regularity assertion (I) of Theorem 2.2 for the solution $U$, i.e., $U \in A(\varepsilon, C\varepsilon^{-1/2}, \gamma)$. The mesh $\Delta$ consists of the single element $I$ with length 1. Corollary 3.9 then implies the existence of $\eta > 0$ (depending solely on $\gamma$) such that the condition

\[ \frac{1}{p \varepsilon} \leq \eta \quad (3.14) \]
implies \(\|U - \mathcal{I}_{p,J} U\|_{H^1(I)} \leq C e^{-\beta p}\), which is even stronger than what is required. The crucial condition (3.14) is easily satisfied by making sure that \(\kappa_0 < \eta/2\), since then the assumption \(\kappa p \varepsilon > 1/2\) produces

\[
\frac{1}{p \varepsilon} = \frac{\kappa}{\kappa p \varepsilon} \leq 2 \kappa \leq 2 \kappa_0.
\]

**case 2:** \(\kappa p \mu < 1/2\) and \(\kappa p (\varepsilon/\mu) < 1/2\). This is the three-scale case (regularitywise) or the 'pre-asymptotic case of \(\varepsilon\)' (approximabilitywise). The mesh is given by \(\Delta = \{0, \kappa p \varepsilon, \kappa p \mu, 1 - \kappa p \mu, 1 - \kappa p \varepsilon, 1\}\), i.e., the mesh has five elements \(I_1, \ldots, I_5\). Statement (II) of Theorem 2.2 gives the decomposition

\[
U = W + \tilde{U}_{BL} + \hat{U}_{BL} + R,
\]

where the smooth part satisfies \(W \in \mathscr{A}(1, C_w, \gamma_w)\), the boundary layers satisfy \(\tilde{U}_{BL} \in \mathscr{B} \mathcal{L}^\infty(\delta \mu, C_\tilde{u}, \gamma_{\tilde{u}})\), \(\hat{U}_{BL} \in \mathscr{B} \mathcal{L}^\infty(\delta \varepsilon, C_\hat{v}, \gamma_{\hat{v}})\), and the remainder satisfies

\[
\|R\|_{L^\infty(\partial I)} + \|R\|_{E,I} \leq C[e^{-b/\mu} + e^{-b \mu/\varepsilon}] \leq C e^{-2b \kappa p}.
\]

For all \(n \in \mathbb{N}_0\) these give

\[
\|W^{(n)}\|_{L^\infty(I)} \leq C_w \gamma_w^n n!,
\]

\[
|\tilde{U}^{(n)}_{BL}(x)| \leq C_{\tilde{u}} \gamma_{\tilde{u}} n (\delta \mu)^{-n} e^{-\text{dist}(x, \partial I)/(\delta \mu)},
\]

\[
|\hat{U}^{(n)}_{BL}(x)| \leq C_{\hat{v}} \gamma_{\hat{v}} n (\delta \varepsilon)^{-n} e^{-\text{dist}(x, \partial I)/(\delta \varepsilon)},
\]

\[
|\hat{v}_{BL}^{(n)}(x)| \leq C_{\hat{v}} \left(\frac{\varepsilon}{\mu}\right)^2 \gamma_{\hat{v}} n (\delta \varepsilon)^{-n} e^{-\text{dist}(x, \partial I)/(\delta \varepsilon)}.
\]

We approximate \(W\) by \(\mathcal{I}_{p,\Delta} W\), the first boundary layer function \(\tilde{U}_{BL}\) by \(\mathcal{I}_{(p,1,p,\Delta)} \tilde{U}_{BL}\), the second boundary layer function \(\hat{U}_{BL}\) by \(\mathcal{I}_{(p,1,1,\rho,\Delta)} \hat{U}_{BL}\), and the remainder \(R\) by its global linear interpolant \(\mathcal{I}_{p,[0,1]} R\).

With the aid of Corollary 3.9, it is easy to see that \(\|W - \mathcal{I}_{p,\Delta} W\|_{1,I} \leq C e^{-\beta p}\) for some \(C, \beta > 0\), independent of \(\varepsilon\) and \(\mu\). For the remainder \(R\), in view of (3.16) which gives control of \(R\) at the end points of \(I\), we get that

\[
\|R - \mathcal{I}_{1,[0,1]} R\|_{E,I} \leq \|R\|_{E,I} + \|\mathcal{I}_{1,[0,1]} R\|_{E,I} \leq C e^{-2b \kappa p}.
\]

We now turn to the approximation of the boundary layer contribution \(\tilde{U}_{BL}\). Lemma 3.11 (with \(v = \delta \mu\)) immediately produces, for the element \(I_3 = (\kappa p \mu, 1 - \kappa p \mu)\), the estimate \(\|\tilde{U}_{BL} - \mathcal{I}_{1,J_3} \tilde{U}_{BL}\|_{E,I_3} \leq C e^{-\kappa \rho/\delta}\), which is exponentially small. For the small elements \(I_1, I_2, I_4\) and \(I_5\), we note that their length is smaller than \(\kappa p \mu\). Corollary 3.9 is applicable with \(K = 1/\mu\), which, for these elements \(I_j, j \in \{1, 2, 4, 5\}\) produces

\[
h_j^{-1} \|\tilde{U}_{BL} - \mathcal{I}_{p,J} \tilde{U}_{BL}\|_{0,I} + \|\hat{U}_{BL} - \mathcal{I}_{p,J} \hat{U}_{BL}\|_{1,I} \leq C \frac{h_j^{1/2}}{p \mu} e^{-\beta p}
\]

if the scale resolution condition

\[
\frac{\kappa p \mu}{p \mu} \leq \eta
\]
is satisfied, where the parameter \( \eta \) depends only on \( \gamma_0 \). Taking \( \kappa_0 \) sufficiently small, this condition is satisfied. Recalling that \( h_j \leq \kappa p \mu \), we see that (3.21) is a much stronger result than required.

We next study the approximation of \( \hat{U}_{BL} \). We first consider the approximation on the elements \( I_2 \), \( I_3 \) and \( I_4 \), which are all in the interval \((\kappa p \varepsilon, 1 - \kappa p \varepsilon)\). For the \( \hat{u} \) component of \( \hat{U}_{BL} \), Lemma 3.11 (with \( \nu = \delta \varepsilon \), \( C_u = C_{\hat{u}} \)) yields the desired exponential approximation result. For the \( \hat{v} \) component, we also apply Lemma 3.11 (with \( \nu = \delta \varepsilon \), \( C_u = C_{\hat{v}} (\varepsilon / \mu)^2 \)) to get

\[
\| \hat{v} - \mathcal{I}_{1,\Delta} \hat{v} \|_{0,(\kappa p \varepsilon, 1 - \kappa p \varepsilon)} + \varepsilon | \hat{v} - \mathcal{I}_{1,\Delta} \hat{v} |_{1,(\kappa p \varepsilon, 1 - \kappa p \varepsilon)} \leq CC \frac{\varepsilon}{\mu} e^{-\kappa p / \delta}.
\]

This implies

\[
\| \hat{v} - \mathcal{I}_{1,\Delta} \hat{v} \|_{0,(\kappa p \varepsilon, 1 - \kappa p \varepsilon)} + \mu | \hat{v} - \mathcal{I}_{1,\Delta} \hat{v} |_{1,(\kappa p \varepsilon, 1 - \kappa p \varepsilon)} \leq CC \frac{\varepsilon}{\mu} e^{-\kappa p / \delta} \leq CC \frac{\varepsilon}{\mu} e^{-\kappa p / \delta},
\]

which is the desired bound.

The approximation of \( \hat{U}_{BL} \) on the remaining elements \( I_1 \) and \( I_5 \) is achieved with the aid of Corollary 3.9 in exactly the same way as \( \hat{U}_{BL} \) was approximated on \( I_1 \), \( I_2 \), \( I_4 \) and \( I_5 \). Again, for the \( \hat{v} \) component, we may exploit the fact that the bounds (3.20) feature an additional factor \((\varepsilon / \mu)^2\).

**Case 3:** \( \kappa p \mu \geq 1/2 \) and \( \kappa p \varepsilon < 1/2 \). This is the first two-scale case (regularitywise) or the ‘semi-asymptotic case of \( p' \) (approximabilitywise). The mesh is given by \( \Delta = \{0, \kappa p \varepsilon, 1 - \kappa p \varepsilon, 1\} \).

Statement (III) of Theorem 2.2 gives the decomposition

\[
U = W + \hat{U}_{BL} + R,
\]

where \( W \in \mathcal{A}(\mu, C_w, \gamma_w) \), \( \hat{U}_{BL} \in \mathcal{B}L^\infty(\delta \varepsilon, C_u, \gamma_u) \), and \( \| R \|_{L^\infty(\partial I)} + \| R \|_{E,I} \leq C e^{-b / \varepsilon} \). Furthermore, the second component \( \hat{v}_{BL} \) of \( \hat{U}_{BL} \) satisfies the stronger assertion \( \hat{v}_{BL} \in \mathcal{B}L^\infty(\delta \varepsilon, C_{\hat{v}} (\varepsilon / \mu)^2, \gamma_{\hat{v}}) \). Recall that the mesh consists of three elements,

\[
I_1 = (0, \kappa p \varepsilon), \quad I_2 = (\kappa p \varepsilon, 1 - \kappa p \varepsilon), \quad I_3 = (1 - \kappa p \varepsilon, 1),
\]

and we approximate each component as follows: \( W \) is approximated by \( \mathcal{I}_{p,\Delta} W \), the boundary layer \( \hat{U}_{BL} \) is approximated by \( \mathcal{I}_{p,1,\Delta} \hat{U}_{BL} \), and \( R \) is approximated by \( \mathcal{I}_{1,\Delta} R \). For the approximation of \( W \), we use Corollary 3.9, which yields exponential convergence, since the scale resolution condition

\[
\eta \geq \frac{h_j}{p \mu} = \frac{h_j \kappa}{\kappa p \mu} \geq 2 \kappa h_j
\]

can be satisfied for all elements by taking \( \kappa_0 \) sufficiently small.

The approximation of \( \hat{U}_{BL} \) follows by the same arguments as in case 2.

Finally, for the remainder \( R \), we again use stability of the piecewise linear approximation and the fact that the algebraic factor \( \varepsilon^{-1} \) can be absorbed into the exponentially small factor \( e^{-b / \varepsilon} \) at the expense of slightly reducing \( b \):

\[
\| R - \mathcal{I}_{1,\Delta} R \|_{E,I} \leq \| R - \mathcal{I}_{1,\Delta} R \|_{1,I} \leq C \| R \|_{1,I} \leq C \varepsilon^{-1} \| R \|_{E,I} \leq C \varepsilon^{-1} e^{-b / \varepsilon} \leq C e^{-b' / \varepsilon} \leq C e^{-2b' \kappa \varepsilon}.
\]
case 4: $\kappa \varepsilon \leq \kappa \rho \mu < 1/2$ and $\kappa \rho (\varepsilon / \mu) \geq 1/2$. This is the second two-scale case or the second 
'semi-asymptotic case of $p$. The mesh is given by $\Delta = \{0, \kappa \varepsilon, \kappa \rho \mu, 1 - \kappa \rho \mu, 1 - \kappa \rho \varepsilon, 1\}$.

Statements (III) and (IV) of Theorem 2.2 give the decompositions

$$U = W_3 + U_{BL} + R_3,$$
$$U = W_4 + U_{BL} + R_4,$$

where $W_4 \in H(1, C, \gamma), U_{BL} \in B.L^\infty (\delta \mu, C \sqrt{\mu / \varepsilon}, \gamma \mu / \varepsilon)$ and $\| R_4 \|_{L^\infty (\partial I)} + \| R_4 \|_{E, I} \leq C (\mu / \varepsilon)^2 e^{-b / \mu}$; furthermore, $W_3 \in H(\mu, C, \gamma), U_{BL} \in B.L^\infty (\delta \varepsilon, C, \gamma)$ and $\| R_3 \|_{L^\infty (\partial I)} + \| R_3 \|_{E, I} \leq C e^{-b / \varepsilon}$ and additionally, the $\hat{v}$ component of $U_{BL}$ satisfies the stronger estimate

$$\hat{v} \in B.L^\infty (\delta \varepsilon, C (\varepsilon / \mu)^2, \gamma).$$

The mesh consists of five elements $I_1, \ldots, I_5$. On $I_1 = (0, \kappa \rho \varepsilon)$ (and analogously on $I_5 = (1 - \kappa \rho \varepsilon, 1)$), we approximate $U$ by

$$\mathcal{J}_{p, I_1} W_4 + \mathcal{J}_{p, I_1} U_{BL} + \mathcal{J}_{I_1} R_4.$$

On $I_2 = (\kappa \rho \varepsilon, \kappa \rho \mu)$ (and analogously on $I_4 = (1 - \kappa \rho \mu, 1 - \kappa \rho \varepsilon)$), we approximate $U$ by

$$\mathcal{J}_{p, I_2} W_3 + \mathcal{J}_{I_2} U_{BL} + \mathcal{J}_{I_2} R_3.$$

For the middle element $I_3$, we use

$$\mathcal{J}_{p, I_3} W_4 + \mathcal{J}_{p, I_3} U_{BL} + \mathcal{J}_{I_3} R_4.$$

The approximation of the functions $W_4$ and $W_3$ is done with the aid of Corollary 3.9. The interesting case is the approximation of $W_3$ on $I_2$ and $I_4$, for which we note that the element size satisfies $h_j = \kappa \rho (\mu - \varepsilon) \leq \kappa \rho \mu$.

Next, we study the approximation of $U_{BL}$ on elements $I_1$ and $I_5$. In order to be able to employ Corollary 3.9, we rewrite the regularity assertion for $U_{BL} \in B.L^\infty (\delta \mu, C \sqrt{\mu / \varepsilon}, \gamma \mu / \varepsilon)$ as follows:

$$\| (U_{BL})^{(\alpha)} \|_{L^\infty (I_1)} \leq C \sqrt{\mu / \varepsilon} \left( \frac{\mu}{\varepsilon} \right)^n (\delta \mu)^{-n} \leq C \sqrt{\mu / \varepsilon} \left( \frac{\mu}{\varepsilon} \right)^n \varepsilon^{-n} \forall n \in \mathbb{N}.$$

Hence, Corollary 3.9 implies

$$\| U_{BL} - \mathcal{J}_{p, I_1} U_{BL} \|_{1, I_1} \leq C \sqrt{\mu / \varepsilon} e^{-\beta \rho}$$

if the scale resolution condition

$$\frac{h_1}{p_1 \varepsilon} \leq \eta$$

is satisfied, where $\eta$ depends solely on $\gamma / \delta$. In view of $h_1 = \kappa \rho \varepsilon$ and $p_1 = p$, this condition is satisfied if $\kappa_0$ is sufficiently small. Finally, the factor $\sqrt{\mu / \varepsilon}$ can be controlled since $\mu / \varepsilon \leq 2 \kappa \rho$, and this algebraic factor can be absorbed into the exponentially decaying one.

The approximation of $U_{BL}$ on $I_2$ and $I_3$ is achieved with Lemma 3.11 (as in case 2, the $\hat{u}$ component and the $\hat{v}$ component have to be studied separately). Finally, the approximation $U_{BL}$ on the middle element $I_3$ is covered by Lemma 3.11.
We next turn to the remainders. The stability properties of the linear interpolant stated in (3.10) and (3.11) yield for each element $I_j$ and arbitrary $\mathbf{R} \in H^1(I_j)$,

$$
\|\mathbf{R} - \mathcal{J}_{1,I_j} \mathbf{R}\|_{1,I_j} \leq C\|\mathbf{R}\|_{1,I_j}.
$$

Hence, we get for arbitrary $\mathbf{R} \in H^1(I)$ that $\|\mathbf{R} - \mathcal{J}_{1,\Omega} \mathbf{R}\|_{1,I} \leq C\|\mathbf{R}\|_{1,I} \leq C\epsilon^{-1}\|\mathbf{R}\|_{E,I}$ and thus

$$
\begin{align*}
\epsilon^{-1}\|\mathbf{R}^{IV}\|_{E,I} & \leq C\epsilon^{-1}(\mu/\epsilon)^2 e^{-b/\mu} \leq C(\mu/\epsilon)^3 \mu^{-1} e^{-b/\mu} \leq C(2\kappa p)^3 e^{-b/\mu}, \\
\epsilon^{-1}\|\mathbf{R}^{III}\|_{E,I} & \leq C\epsilon^{-1} e^{-b/\epsilon} \leq C e^{-b/\epsilon}.
\end{align*}
$$

Using $\epsilon \leq \mu$ and $2\kappa p \geq 1/\mu$ shows that these terms are exponentially small in $\kappa p$ as desired.

This completes the proof of Theorem 3.3.

**Remark 3.12** Theorem 3.3 asserts the existence of values $\kappa$ such that the spectral boundary layer $S(\kappa, p)$ has good approximation properties. Some indication on how to select $\kappa$ can be inferred from a more detailed analysis of the layer structure of the solution in conjunction with the results of Schwab & Suri (1996). Specifically, we recall from Schwab & Suri (1996) that on the interval $I = (0, 1)$ the function $x \mapsto e^{-x/d}$ (with $d > 0$) can be approximated well by piecewise polynomials of degree $p$ on the mesh with nodes $\{0, kp, 1\}$, where $\kappa \in (0, 4/\epsilon)$; in particular, $\kappa = 1$ is a good choice in this situation.

In order to apply this result here, let us assume that $\mu$ and $\epsilon/\mu$ are both sufficiently small that the leading-order terms of the asymptotic expansion provide an accurate description of the solution. By Melenk et al. (2011, Section 4) we then have

$$
\mathbf{U} \approx \mathbf{W} + C_1 e^{-\lambda_1 x/\mu} + C_2 e^{-\lambda_2 x/\epsilon} + C_3 e^{-\lambda_3 (1-\epsilon)/\mu} + C_4 e^{-\lambda_4 (1-\epsilon)/\epsilon}
$$

(3.23)

with vectors $C_i \in \mathbb{R}^2$ of size $O(1)$ (additionally, the second components of $C_2$ and $C_4$ are of size $(\epsilon/\mu)^2$). The coefficients $\lambda_i > 0$ are given by

$$
\lambda_1^2 = \frac{\det A(0)}{A_{11}(0)}, \quad \lambda_2^2 = A_{11}(0), \quad \lambda_3^2 = \frac{\det A(1)}{A_{11}(1)}, \quad \lambda_4^2 = A_{11}(1).
$$

(3.24)

This indicates that a good choice of mesh should include points close to $p\mu/\lambda_1$, $1-p\mu/\lambda_3$ to resolve the $O(\mu)$ layers and close to $p\epsilon/\lambda_2$, $1-p\epsilon/\lambda_4$ to resolve the $O(\epsilon)$ layers.

### 4. Numerical results

In this section, we present the results of numerical computations for the model problem considered in Linß & Madden (2004a), Madden & Stynes (2003) and Xenophontos & Oberbroeckling (2007). The data are chosen as follows:

$$
\mathbf{A} = \begin{bmatrix}
2(x+1)^2 & -(1+x^3) \\
-2\cos(\pi x/4) & 2.2e^{(1-x)}
\end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix}
2e^x \\
10x + 1
\end{bmatrix},
$$

with $\kappa$, appearing in the definition of the mesh, taken as 1. An exact solution is not available, hence for the computations we use a reference solution $\mathbf{U}_{\text{REF}}$, obtained with a high number of degrees of freedom. Since various numerical results for this problem have already been presented in Xenophontos & Oberbroeckling (2007), we focus on issues not studied there. We will present the estimated percentage...
relative error in the energy norm given by

\[ 100 \times \frac{\|U_{\text{REF}} - U_N\|_{E,I}}{\|U_{\text{REF}}\|_{E,I}}, \tag{4.1} \]

versus the number of degrees of freedom \( N \). In the last subsection, we will address the issue of convergence and robustness when the maximum norm \( \| \cdot \|_{\infty,I} \) is used as an error measure.

4.1 The case of a single boundary layer: case III

In this case \( \mu \) is not small but \( \epsilon/\mu \) is. It may therefore be sufficient to use a three-element mesh \( \Delta_\epsilon = \{0, \epsilon, 1 - \epsilon, 1\} \). For this choice of mesh, the left graph in Fig. 1 depicts the cases \( \mu = 0.1, \epsilon = 10^{-j}, j = 2, \ldots, 6 \) and shows exponential convergence of the error (4.1). We note that for small polynomial degree, the convergence is somewhat reduced since \( \mu = 0.1 \) is not very large.

4.2 Five elements versus three elements

We focus on the case \( \epsilon = 10^{-3} \) and \( \mu = 10^{-2} \). The numerical results are shown in the right-hand graph in Fig. 1, where the error is plotted versus the polynomial degree. We compare the performance of the spectral boundary layer mesh (with five elements) with two meshes with merely three elements, namely, \( \Delta_\epsilon = \{0, \epsilon, 1 - \epsilon, 1\} \) (which is able to resolve the features on the \( \mathcal{O}(\epsilon) \) scale) and \( \Delta_\mu = \{0, \mu, 1 - \mu, 1\} \) (which is adequate to capture the features on the \( \mathcal{O}(\mu) \) scale). We recall that the problem size \( N = 2(M - 1)p \), where \( M \) is the number of elements (here \( M = 3 \) or 5). We note that the spectral boundary layer mesh with five elements yields exponential convergence. The other two meshes show a significantly worse performance. To qualitatively understand this behaviour, we assume that \( \mu \) and \( \epsilon/\mu \) are sufficiently small for asymptotic expansions to be meaningful, and we study the approximation of the leading-order layers, which take the form (3.23) with parameters \( \lambda_i \) given by (3.24), i.e.,
1. Polynomial degree $p$

Estimated percentage relative error in energy norm $\varepsilon = 10^{-3}, \mu = 10^{-2}$, three-element mesh resolving $\mu$ scale $\varepsilon$-scale polynomial degree $p$

error $O(1/p)$

2. $\lambda_1 \approx 2.23, \lambda_2 \approx 1.41, \lambda_3 \approx 2.82, \lambda_4 \approx 1.36$. The detailed analysis of Schwab & Suri (1996) leads us to expect, for the case of small $p$ under consideration here, the following convergence behaviours: for the three-element mesh capable of resolving features on the $\varepsilon$ scale, the approximation of the contributions $C_1 e^{-\lambda_1 x/\mu}$ and $C_3 e^{-\lambda_3 (1-x)/\mu}$ is achieved at the rate $O(e^{-p\lambda_{\min}\varepsilon/\mu})$, where $\lambda_{\min} = \min\{\lambda_1, \lambda_3\}$. For the three-element mesh that is capable of resolving features on the $\mu$ scale, the approximation of the contributions $C_2 e^{-\lambda_2 x/\varepsilon}$ and $C_4 e^{-\lambda_4 (1-x)/\varepsilon}$ can be done with $O(1/p)$. Figure 2 shows the performance of these meshes on doubly-logarithmic and semilogarithmic scales.

4.3 Convergence in the maximum norm

Even though the natural energy norm associated with (2.4) was used in our analysis (and numerical experiments), we wish to comment on the performance of the proposed method when the maximum norm $\|\cdot\|_{\infty, I}$ is used as an error measure. First, we note that from the (Sobolev embedding) inequality $\|v\|_{\infty, I} \leq C \|v\|_{0, I}^{1/2} \|v'\|_{0, I}^{1/2}$, we obtain bounds in the maximum norm from the estimates given in Corollary 3.4, namely,

$$\|U - U_N\|_{\infty, I} \leq Ce^{-1/2} e^{-\beta kp}.$$ 

While this estimate is obviously $\varepsilon$ dependent, this dependence is rather weak since it implies exponential convergence with $\varepsilon$-independent constants under the weak side constraint $p \geq c \log \varepsilon^{-1}$ for some $c > 0$ which depends on the problem. Figure 3 shows the (estimated) percentage relative error in the maximum norm, for the three-scale case, i.e., when $\varepsilon \ll \mu \ll 1$. As the figure shows, the method yields exponential convergence, even in the maximum norm, with the values of the singular perturbation parameters not affecting the method’s performance. See also Xenophontos & Oberbroeckling (2007) for more numerical results (in the maximum norm) and comparisons with other methods.

5. Conclusions and extensions

We considered a coupled system of two reaction–diffusion equations with two singular perturbation parameters $0 < \varepsilon \leq \mu \leq 1$. We have proved that the $hp$ FEM proposed in Xenophontos &
Fig. 3. Maximum norm convergence for the *hp* version on the spectral boundary layer mesh.

Oberbroeckling (2007), indeed exhibits exponential rates of convergence, independently of the singular perturbation parameters $\epsilon$ and $\mu$, as the degree $p$ of the approximating polynomial is increased. The mesh design principle is simple in that for each boundary layer a small element that is capable of resolving it is inserted in the mesh. This mesh design principle is applicable beyond the problem class here, for example, in one dimension to systems with more than two components (and thus more than two layers) and convection–diffusion problems. We also believe that an extension to systems of reaction–diffusion equations in two dimensions is possible. The key ingredient of the present analysis is the detailed regularity theory of Melenk et al. (2011). A rigorous analysis of the above extensions would require analogous regularity results. The straightforward extension of these regularity results seems feasible but technically involved; for example, while the present system of two equations required a regularity theory distinguishing between four different cases of relationships between the scales, the extension to the problem

$$-\mathbf{E} U'' + A(x) U = f \quad \text{on } I = (0, 1), \quad U(0) = U(1) = 0 \in \mathbb{R}^d$$

with a positive definite diagonal matrix $\mathbf{E} = \text{diag}(\epsilon_1, \ldots, \epsilon_d)$ would require distinguishing between $2^d$ cases.

Inspection of the proof of the approximation result Theorem 3.3 shows that some refinements are possible. We highlight three of them.

1. Our boundary layer approximation relies on Lemma 3.11, which provides estimates for the piecewise linear approximation of boundary layer functions outside the layer. This approximation can be improved using the technique employed in Schwab & Suri (1996, Theorem 5.1) of adding an appropriate additional piecewise linear function. The end result is then that one can construct a piecewise polynomial approximation $\pi$ to a function $u \in \mathcal{B}_c \mathcal{D}^\infty(v, C, \gamma)$ on the three-element mesh $\Delta = \{0, \kappa \nu p, 1 - \kappa \nu p, 1\}$ satisfying

$$\|u - \pi\|_{0, I} + (\kappa \nu \gamma) \|u - \pi\|_{1, I} \leq C \sqrt{\kappa \nu \gamma} e^{-\beta \kappa p}.$$
These boundary layer approximation results can, for example, lead to improved error bounds if the model problem (2.1a) and (2.1b) is considered with $f = g = 0$ and inhomogeneous Dirichlet boundary data. More generally, improved error bounds (with respect to $\mu$ and $\varepsilon$) can be expected if the leading-order term of the outer expansion (of an asymptotic expansion) is contained in the approximation space. This happens, for example, for a constant $A$ and polynomial right-hand sides $f$ and $g$. These observations parallel the case of a scalar singularly perturbed problem analysed in Schwab & Suri (1996) and are visible in the numerics of Xenophontos & Oberbroeckling (2007).

2. The sublayer on the $O(\varepsilon)$ scale is particularly weak in the $\hat{v}$ component. This could be exploited further. For example, in certain parameter ranges, one could remove the mesh points $k\varepsilon$ and $1 - k\varepsilon$ in the mesh for the second component without compromising, up to certain tolerances, the accuracy of the FEM. On the other hand, this may be awkward to program, since it would require different meshes for each component of $U$.

3. The proof of Theorem 3.3 relies on the regularity assertions of Theorem 2.2. There, $L^\infty$-based estimates are given for the solution $U$. The proof of Theorem 3.3, however, mostly uses $L^2$-based regularity estimates.

REFERENCES


