On the robust exponential convergence of hp finite element methods for problems with boundary layers

JENS MARKUS MELENK

Seminar für Angewandte Mathematik, ETH Zürich,
CH-8092 Zürich, Switzerland

The hp version of the finite element method for a one-dimensional, singularly perturbed elliptic–elliptic model problem with analytic input data is considered. It is shown that the use of piecewise polynomials of degree p on a mesh consisting of three suitably chosen elements leads to robust exponential convergence, i.e., the exponential rate of convergence depends only on the input data and is independent of the perturbation parameter.

1. Introduction

The approximation of singularly perturbed problems by numerical methods has lately attracted much attention—we mention here only the recent books by Morton (1996), Roos *et al* (1996), Miller *et al* (1996), and the many references therein. The methods for the approximation of singularly perturbed problems discussed in the literature are mostly concerned with robust h versions, that is, they aim at proving that the error (in some suitable norm) is $O(h^p)$ in the mesh width h for some $p > 0$ *uniformly* in the perturbation parameter. Schemes like this lead to *algebraic* rates of convergence only. Often, however, the solution is analytic (or piecewise analytic). Then spectral approximation, i.e., approximation with piecewise polynomials of increasingly higher degree leads to exponential approximation rates. Of course, as the solution depends on the singular perturbation parameter, we expect the rates of the exponential approximability to depend on that parameter as well. The aim of the present paper is to show for a model problem that indeed under the assumption of analyticity of the input data, spectral approximation of the solution leads to exponential convergence and that, with the proper mesh design, *robust, exponentially* converging hp finite element methods (hp FEM) are available.

Robust exponential convergence of hp FEM for an elliptic–elliptic problem with boundary layers was first proved by Schwab & Suri (1996). However, the analysis of Schwab & Suri (1996) is restricted to problems of the form of (1) with polynomial right-hand side and constant coefficients. The main purpose of this paper is to extend these results to general analytic input data; we show that the 'three-element' mesh approach of Schwab & Suri (1996) leads indeed to *robust exponential convergence* for general analytic right-hand sides and variable, analytic coefficients. The novel feature in the proof over the techniques used by Schwab & Suri (1996) is a more careful use of the classical asymptotic expansions available for problems of the type (1). More precisely, the analyticity of the input data gives us complete control over the terms arising in the classical asymptotic expansions and allows us to bound the remainder explicitly *both* in terms of the perturbation parameter d and

the expansion order M; this explicit control over the remainder is the essential ingredient for the proof of robust exponential approximability.

Whereas the books mentioned above are mostly interested in singularly perturbed problems of elliptic–hyperbolic type, we will discuss here a singularly perturbed elliptic–elliptic reaction–diffusion equation. The solution of that equation is the archetype of the boundary layers arising in solid mechanics, for example, in various plate and shell models at small thickness. The results of this paper therefore provide insight in how to design appropriate meshes for two- or three-dimensional problems.

Although we analyze in this paper a simple one-dimensional model problem, the scope is wider. The main tool for our robust exponential approximability result is the ability to use the analyticity of the input data to control the classical asymptotic expansions in terms of the perturbation parameter and the expansion order. Thus, whenever classical asymptotic expansions are available, the techniques employed in the present paper may be used to obtain similar approximation results. For example, the one-dimensional convection–diffusion equation with analytic coefficients falls into that category. Another example is given by the two-dimensional reaction–diffusion equation for which the analysis of the present paper was successfully extended by Melenk & Schwab (1997a) to design an hp FEM converging at a robust exponential rate.

Let us note that the robust exponential approximability obtained in this paper yields automatically exponential rates of convergence of the finite element method for our elliptic–elliptic model problem as the FEM is trivially stable. The situation is more delicate for elliptic–hyperbolic equations, a typical representative of which is the convection–diffusion equation. Whereas the approximability results of this paper hold true for the convection–diffusion equation as well, the stability of finite element methods for that equation is a non-trivial issue. A stable hp method for the convection–diffusion equation featuring robust exponential rates of convergence is presented by Melenk & Schwab (1997b).

In the elliptic–elliptic model problem analyzed in this paper the solution is analytic up to the boundary. Likewise the limiting solution, i.e., the solution of the problem when the perturbation parameter tends to zero, is analytic up to the boundary. In many cases of practical importance, however, neither the solution nor the limiting solution have that much smoothness. Furthermore, the limiting solution may have a substantially different character. For example, in the case of the Reissner–Mindlin plate model with polygonal mid-plane, the solution has corner singularities. The limiting solution, which solves the Kirchhoff equation on a polygonal domain, also has corner singularities, albeit of a different type than those of the solution of the Reissner–Mindlin model. We model this behaviour in our one-dimensional numerical studies by using singular right-hand sides. Of course, such a situation is not covered by the mathematical theory presented in this paper as the classical asymptotic expansions have no meaning. However, we show numerically that a 'union' of the 'three-element' mesh to resolve the boundary layer and a geometrically graded mesh, which is well suited to absorb both the singular behaviour of the solution as well as the limiting solution, leads to very satisfactory schemes.

1.1   *The model problem*

We consider the approximation by the p and hp version of the finite element method of the following singularly perturbed boundary value problem:

$$L_d u_d := -d^2 u_d'' + b(x) u_d = f \qquad \text{on } \Omega := (-1, 1),$$
$$u(\pm 1) = \alpha^\pm \in \mathbf{R}, \tag{1}$$

where f, b are functions *analytic* on the closed interval $\overline{\Omega} = [-1, 1]$, $b(x) \geqslant \underline{b}^2 > 0$ on $\overline{\Omega}$, and $d \in (0, 1]$ is a small parameter which may approach zero. We will make henceforth the assumption that there are C_f, γ_f, C_b, $\gamma_b > 0$ such that

$$\|f^{(n)}\|_{L^\infty(\Omega)} \leqslant C_f \gamma_f^n n! \qquad \forall n \in \mathbf{N}_0, \tag{2}$$

$$\|b^{(n)}\|_{L^\infty(\Omega)} \leqslant C_b \gamma_b^n n! \qquad \forall n \in \mathbf{N}_0, \tag{3}$$

$$b \geqslant \underline{b}^2 > 0 \qquad \text{on } \overline{\Omega}. \tag{4}$$

The analyticity of the input data f, b will allow us to control in Section 2 the asymptotic expansion of the solution u_d explicitly both in terms of d and the expansion order.

The weak formulation of this boundary value problem is:

$$\text{find } u_d \in H_D^1(\Omega) \text{ such that } B_d(u_d, v) = F(v) \qquad \forall v \in H_0^1(\Omega) \tag{5}$$

where we set

$$B_d(u_d, v) = \int_\Omega \left(d^2 u_d' v' + b(x) u_d v \right) \, \mathrm{d}x,$$

$$F(v) = \int_\Omega f v \, \mathrm{d}x,$$

$$H_D^1(\Omega) = \{ u \in H^1(\Omega) \mid u(-1) = \alpha^-, u(1) = \alpha^+ \},$$

$$H_0^1(\Omega) = \{ u \in H^1(\Omega) \mid u(\pm 1) = 0 \},$$

and we denote $H^1(\Omega)$ the usual Sobolev space of all square integrable functions whose (distributional) derivative is also square integrable. Associated with the weak formulation (5) is an 'energy norm'

$$\|u\|_d := (B_d(u, u))^{1/2} . \tag{6}$$

We have the following *a priori* estimate for the solution u_d of (1):

$$\|u_d\|_d \leqslant \underline{b}^{-1} \|f\|_{L^2(\Omega)} + C \left(|\alpha^-| + |\alpha^+| \right), \qquad C \leqslant 2\sqrt{d^2/2 + 2/3 \|b\|_{L^\infty(\Omega)}}. \tag{7}$$

In the finite element method, for a given finite dimensional space $S_0 \subset H_0^1(\Omega)$ of dimension N and a fixed $u_0 \in H_D^1(\Omega)$ (e.g., a linear function), the affine spaces $H_D^1(\Omega)$ and $H_0^1(\Omega)$ are replaced with affine spaces $S_D := u_0 + S_0 \subset H_D^1(\Omega)$, $S_0 \subset H_0^1(\Omega)$ of dimension N. The finite element solution is then given by:

$$\text{find } u_{FE} \in S_D \text{ such that } B_d(u, v) = F(v) \qquad \forall v \in S_0. \tag{8}$$

By the well-known orthogonality relation the finite element solution u_{FE} is the best approximant of the exact solution u in the energy norm, i.e.,

$$\|u - u_{FE}\|_d \leqslant \inf_{v \in S_D} \|u - v\|_d. \tag{9}$$

This paper is therefore only concerned with the approximation properties of the spaces $S_D := S(\kappa, p) \cap H^1_D(\Omega)$ defined in Section 4.

Because the right-hand side f and the coefficient b are analytic, the solution u_d is also analytic and therefore spectral approximation of the exact solution, i.e., approximation by polynomials of increasingly higher degree on the whole domain, is exponential for sufficiently large p. However, this exponential approximation is not robust, that is, the rate of convergence depends on the parameter d. More precisely, the exponential convergence is only visible if $p > d^{-1}$ (cf the discussion following Theorem 1). In the range $p \ll d^{-1}$ which is of practical interest, we will only observe convergence of the type $O(p^{-1}\sqrt{\ln p})$ (see Schwab & Suri (1996)). The deterioration of the performance of the usual spectral method in the range of practical interest, $p \ll d^{-1}$, is due to the presence of boundary layers. In the 'three-element' approach, these boundary layers are resolved by splitting the domain Ω into three elements and approximating by piecewise polynomials of degree p. The two elements adjacent to the boundary points -1 and 1 are of size $O(dp)$ and thus small enough to capture the boundary layer behaviour of the solution near the boundary points ± 1. The introduction of these two additional small elements allows us to obtain exponential convergence for the approximation of solutions of (1) which is robust, i.e., the convergence rate is independent of the small parameter d (Theorem 16).

2. Regularity of the solution

Clearly, in order to find spaces S_D in which the exact solution u_d of (1) can be approximated well, we have to be able to describe the behaviour of u_d precisely. In this section we therefore present the necessary regularity results for the solution u_d.

THEOREM 1   Let u_d be the solution of (1). Then there are constants $C, K > 0$ depending only on the right-hand side f, the coefficient b, and the boundary data $\alpha^{\pm}$ such that

$$\|u_d^{(n)}\|_{L^2(\Omega)} \leqslant C K^n \max(n, d^{-1})^n \qquad \forall n \in \mathbf{N}_0. \tag{10}$$

Proof. Choose $K > \max(1, \gamma_f, \gamma_b)$ such that

$$\left[2C_f K^{-2} \left(\frac{\gamma_f}{K} \right)^n + C_b K^{-2} \frac{1}{1 - \gamma_b/K} \right] \leqslant 1 \qquad \forall n \in \mathbf{N}_0.$$

By (7) we may now choose the constant $C \geqslant 1$ such that (10) holds true for $n = 0, 1$. Theorem 1 is now proved by an induction argument. The constants C, K are such that the induction hypothesis holds true for $n = 0, 1$. Let us assume that the induction hypothesis (10) holds for $0 \leqslant \nu \leqslant n+1$ and show that it holds for $n+2$. Differentiating the differential equation n times (note that we already know that u_d is analytic) we get

$$-d^2 u_d^{(n+2)} = f^{(n)} - (b u_d)^{(n)} = f^{(n)} - \sum_{\nu=0}^{n} \binom{n}{\nu} b^{(\nu)} u_d^{(n-\nu)}.$$

Using the induction hypothesis, we get

$$d^2\|u^{(n+2)}\|_{L^2(\Omega)} \leqslant \|f^{(n)}\|_{L^2(\Omega)} + \sum_{\nu=0}^{n}\binom{n}{\nu}C_b\gamma_b^\nu \nu! CK^{n-\nu}\max(n-\nu,d^{-1})^{n-\nu}$$

$$\leqslant 2C_f\gamma_f^n n^n + CC_bK^n\sum_{\nu=0}^{n}\frac{n!}{(n-\nu)!}\left(\frac{\gamma_b}{K}\right)^\nu\max(n-\nu,d^{-1})^{n-\nu}$$

$$\leqslant 2C_f\gamma_f^n\max(n,d^{-1})^n + CC_bK^n\frac{1}{1-\gamma_b/K}\max(n,d^{-1})^n$$

$$\leqslant CK^{n+2}\max(n+2,d^{-1})^n\left[2C_fK^{-2}\left(\frac{\gamma_f}{K}\right)^n + C_bK^{-2}\frac{1}{1-\gamma_b/K}\right].$$

By the choice of K the expression in the brackets is bounded by 1 which concludes the induction argument after dividing both sides by d^2. $\qquad\square$

Note that Theorem 1 yields estimates for the nth derivative of the solution u_d which are independent of d provided that $n \geqslant cd^{-1}$ for some $c > 0$. Roughly speaking, this means that derivatives of order sufficiently large 'don't see' the boundary layer arising for small perturbation parameters d. It is also not too hard to see that from this theorem we can obtain robust exponential convergence of the p version of the finite element method provided that the polynomial degree p is at least $O(d^{-1})$.

Theorem 1 does not reflect the boundary layer behaviour of the solution u_d very well. This boundary layer behaviour can be described in terms of the classical asymptotic expansions for the solution of (1). For $M \in \mathbf{N}_0$ we can decompose u_d in a (smooth) asymptotic part, two boundary layer parts, and a remainder as follows:

$$u_d = w_M + A_M^- u_d^- + A_M^+ u_d^+ + r_M. \tag{11}$$

This decomposition is obtained as follows. Upon inserting the formal *ansatz* $u_d \sim \sum_{j=0}^{\infty}d^j u_j$ in the differential equation (1) and equating like powers of d, we can define the asymptotic (smooth) part w_M by partial sums of this formal series:

$$w_M(x) := \sum_{j=0}^{M}d^{2j}u_{2j}$$

where the terms u_{2j} are defined recursively by

$$u_0(x) = \frac{f(x)}{b(x)}, \qquad u_{j+2} = \frac{1}{b(x)}u_j''(x), \qquad j = 0, 2, 4, \dots$$

A simple calculation shows that $L_d(u_d - w_M) = d^{2M+2}u_{2M}''$ which tends to zero as d tends to zero for each fixed M. Thus the functions w_M satisfy asymptotically, as d tends to zero, the differential equation but they do not satisfy the boundary conditions. This incompatibility can be removed with the aid of two boundary layer functions u_d^-, u_d^+ defined as the solutions of

$$\left.\begin{array}{l}L_d u_d^- = 0 \quad \text{on } \Omega \\ u_d^-(-1) = 1,\ u_d^-(1) = 0\end{array}\right\} \qquad \left.\begin{array}{l}L_d u_d^+ = 0 \quad \text{on } \Omega \\ u_d^+(-1) = 0,\ u_d^+(1) = 1.\end{array}\right\} \tag{12}$$

Upon setting

$$A_M^- := \alpha^- - w_M(-1), \qquad A_M^+ := \alpha^+ - w_M(1)$$

the function $w_M + A^- u_d^- + A^+ u_d^+$ satisfies the correct boundary and still the same differential equation as w_M. Finally, let us define the remainder r_M in such a way that the decomposition (11) holds true, i.e., define r_M as the solution of

$$L_d r_M = d^{2M+2} u_{2M}'' \quad \text{on } \Omega, \qquad r_M(\pm 1) = 0. \tag{13}$$

Our aim is now to analyze the behaviour of each of four terms in the decomposition (11).

LEMMA 2   Let $G \subset \mathbb{C}$ be a complex neighbourhood of $\overline{\Omega} = [-1, 1]$. Let $B, u_0 : G \to \mathbb{C}$ be two functions holomorphic and bounded on G. Define functions u_{2j} recursively via

$$u_{j+2}(x) := B(x) u_j''(x), \qquad j = 0, 2, 4, \dots$$

Then there are C, K_1', $K_2' > 0$ depending only on G and $\|B\|_{L^\infty(G)}$ such that

$$\|u_j^{(n)}\|_{L^\infty(\Omega)} \leqslant C j! n! K_1'^j K_2'^n \|u_0\|_{L^\infty(G)}, \qquad j = 0, 2, 4, \dots, \quad \forall n \in \mathbb{N}_0.$$

Proof. For $0 < \delta \leqslant 1$, define the sets G_δ by

$$G_\delta := \{ z \in G \mid \operatorname{dist}(z, \partial G) \geqslant \delta \}.$$

The claim of the lemma follows immediately from Cauchy's integral theorem for derivatives, if we can show the following, stronger assertion: There is $K > 0$ such that for $0 < \delta \leqslant 1$

$$\|u_j\|_{L^\infty(G_\delta)} \leqslant \delta^{-j} K^j j! \|u_0\|_{L^\infty(G)}, \qquad j = 0, 2, 4, \dots \tag{14}$$

Fix $K^2 > 2e \|B\|_{L^\infty(G)}$. Clearly (14) is true for $j = 0$. We proceed now by induction on j. Assume that (14) holds true for a given $j \geqslant 0$. For any $0 < \kappa < 1, x \in G_\delta$, we may apply Cauchy's integral formula for derivatives to u_j where the integration path is chosen as the circle of radius $\kappa\delta$ about x. Hence, we obtain

$$\|u_{j+2}\|_{L^\infty(G_\delta)} \leqslant \|B\|_{L^\infty(G)} \|u_j''\|_{L^\infty(G_\delta)} \leqslant \|B\|_{L^\infty(G)} \frac{2}{2\pi} \frac{2\pi\kappa\delta}{(\kappa\delta)^3} \|u_j\|_{L^\infty(G_{(1-\kappa)\delta})}$$

$$\leqslant 2 \|B\|_{L^\infty(G)} (\kappa\delta)^{-2} j! K^j ((1-\kappa)\delta)^{-j} \|u_0\|_{L^\infty(G)}$$

$$\leqslant \delta^{-(j+2)} K^{j+2} (j+2)! \|u_0\|_{L^\infty(G)} \left[\frac{2 \|B\|_{L^\infty(G)}}{K^2 \kappa^2 (1-\kappa)^j (j+1)(j+2)} \right].$$

Choosing $\kappa = 1/(j+2)$ and observing that this choice implies

$$\frac{1}{\kappa^2 (1-\kappa)^j (j+1)(j+2)} = \frac{(j+2)^{j+1}}{(j+1)^{j+1}} = \left(1 + \frac{1}{j+1} \right)^{j+1} \leqslant e \qquad \forall j \in \mathbb{N}_0$$

allows us to infer that the expression in brackets is bounded by 1 by the choice of K. $\quad\square$

THEOREM 3   There are constants C, K_1, $K_2 > 0$ depending only on the input data f, b, and $\alpha^\pm$ such that the functions w_M of (11) satisfy the following estimate: Under the assumption $0 < 2Md K_1 \leqslant 1$

$$\|w_M^{(n)}\|_{L^\infty(\Omega)} \leqslant C K_2^n n! \qquad \forall n \in \mathbb{N}_0.$$

Proof. As the function $B(x) := 1/b(x)$ is analytic on $\overline{\Omega}$, there is a complex neighbourhood G of $\overline{\Omega}$ on which B is holomorphic and bounded. Because f is analytic on $\overline{\Omega}$, we may assume without loss of generality that f is holomorphic on G as well. Hence Lemma 2 is applicable to the terms u_{2j} appearing in the definition of w_M and yields:

$$\|u_{2j}^{(n)}\|_{L^\infty(\Omega)} \leqslant C K_2'^n n! K_1'^{2j}(2j)! \qquad \forall n, j \in \mathbf{N}_0.$$

Hence we obtain

$$\|w_M^{(n)}\|_{L^\infty(\Omega)} \leqslant C K_2'^n n! \sum_{j=0}^{M} d^{2j} K_1'^{2j}(2j)! \qquad \forall n \in \mathbf{N}_0.$$

Estimating $d^{2j} K_1'^{2j}(2j)! \leqslant (dK_1'2M)^{2j}$ we see that the sum can be majorized by a converging geometric series provided that $dK_2'2M \leqslant q$ for some fixed $q < 1$. The claim of the theorem follows by setting $K_1 = K_1'/q$, $K_2 = K_2'$. □

This theorem allows us to control the growth of the derivatives of the asymptotic part w_M. An immediate corollary is that we can control the coefficients A_M^-, A_M^+.

COROLLARY 4   With K_1 as in Theorem 3 there is $C > 0$ depending only on the input data f, b, and $\alpha^\pm$ such that for any d, M satisfying $2MdK_1 \leqslant 1$

$$|A_M^-|, |A_M^+| \leqslant C.$$

Proof. Noting that $A_M^- = \alpha^- - w_M(-1)$, $A_M^+ = \alpha^+ - w_M(1)$, the proofs follows from Theorem 3.   □

Let us now consider the two boundary layers.

THEOREM 5   Let u_d^-, u_d^+ be the solutions of (12). Then there are C, $K_3 > 0$ depending only on the function b such that

$$|(u_d^-)^{(n)}(x)| \leqslant e^{-\underline{b}(x+1)/d} C K_3^n \max(n, d^{-1})^n \qquad \forall x \in \overline{\Omega}, \quad n \in \mathbf{N}_0, \qquad (15)$$

$$|(u_d^+)^{(n)}(x)| \leqslant e^{-\underline{b}(1-x)/d} C K_3^n \max(n, d^{-1})^n \qquad \forall x \in \overline{\Omega}, \quad n \in \mathbf{N}_0. \qquad (16)$$

Proof. We observe that an induction argument similar to that of the proof of Theorem 1 leads to the desired estimates provided that we can show the induction hypothesis for $n = 0$ and $n = 1$. These are slight variations of estimates based on the maximum principle first proved by O'Riordan & Stynes (1986) (see also the book by Miller *et al* (1996) for a nice exposition).   □

Let us finally consider the remainder r_M.

THEOREM 6   There are constants C, $K_4 > 0$ depending only on the functions f and b such that the remainders r_M defined in (11) satisfy

$$\|r_M^{(n)}\|_{L^2(\Omega)} \leqslant C d^{2-n} (2MdK_4)^{2M}, \qquad n = 0, 1, 2.$$

Proof. The functions r_M satisfy

$$L_d r_M = d^{2M+2} u_{2M}'' \quad \text{on } \Omega, \qquad r_M(\pm 1) = 0.$$

We saw in the proof of Theorem 3 that

$$\|u_{2M}^{(n)}\|_{L^\infty(\Omega)} \leqslant C K_1'^{2M}(2M)! K_2'^n n! \qquad \forall n \in \mathbf{N}_0.$$

The *a priori* estimate (7) therefore gives

$$\|r_M\|_d \leqslant C d^{2M+2} K_1'^{2M}(2M)!$$

which gives the desired estimates for $n = 0$ and $n = 1$. Using the differential equation satisfied by r_M gives the result for $n = 2$. $\qquad \Box$

REMARK 7   We see that an induction argument would allow us to control also all derivatives of r_M explicitly in terms of M, d, and the order n of the differentiation.

REMARK 8   Theorem 6 asserts that the remainder r_M is indeed small provided that $2MdK_4 < 1$, i.e., if $2Md$ is small. In the complementary case, $2Md$ large, the asymptotic expansion loses its meaning.

3. Polynomial approximation results

The aim of this section is to show that, for the H^1 conforming approximation with piecewise polynomials, it is enough to control the growth of the derivatives on each element.
  Let $I = [-1, 1]$.

LEMMA 9   Let $u \in C^\infty(I)$ satisfy

$$\|D^p u\|_{L^2(I)} \leqslant C_u p! \gamma^p \tag{17}$$

for some C_u, $\gamma > 0$. Then there is a sequence of polynomials $(P_p)_{p=0}^\infty$ of degree p such that

$$\|u - P_p\|_{L^\infty(I)} + \|(u - P_p)'\|_{L^\infty(I)} \leqslant C C_u e^{-\sigma p}$$

where the constants C, $\sigma > 0$ depend only on γ.

Proof. From Sobolev's embedding theorem, we have that $\|D^p u\|_{L^\infty(I)} \leqslant C_u C_1 p! \gamma'^p$ for some C_1, γ' depending only on γ. Therefore, u is analytic on the closed set I and can be extended analytically to a complex neighbourhood of I. The result follows from standard theory: for example, the polynomial P_p may be taken by interpolating u in the Chebyshev points (see Chapter 4 of Davis (1974) for the details). $\qquad \Box$

  Define on the space $C(I)$ for $p \geqslant 1$ the operator i_p by interpolation in the $p + 1$ Gauss–Lobatto points. As shown by Sündermann (1980), there is $C_G > 0$ such that the following stability estimate holds:

$$\|i_p u\|_{L^\infty(I)} \leqslant C_G(1 + \ln p)\|u\|_{L^\infty(I)} \qquad \forall u \in C(I). \tag{18}$$

LEMMA 10   Let $u \in C^1(I)$. Then

$$\|u - i_p u\|_{L^\infty(I)} \leqslant (1 + C_G)(1 + \ln p)\|u\|_{L^\infty(I)},$$

$$\|(u - i_p u)'\|_{L^\infty(I)} \leqslant \|u'\|_{L^\infty(I)} + C_G(1 + \ln p)p^2 \|u\|_{L^\infty(I)}.$$

Proof. The proof of the first estimate follows immediately from the stability estimate (18). For the second one, we use Markov's inequality, $\|v_p'\|_{L^\infty(I)} \leqslant p^2 \|v_p\|_{L^\infty(I)}$, valid for all polynomials v_p of degree p to get

$$\|(u - i_p u)'\|_{L^\infty(I)} \leqslant \|u'\|_{L^\infty(I)} + \|(i_p u)'\|_{L^\infty(I)} \leqslant \|u'\|_{L^\infty(I)} + p^2 \|i_p u\|_{L^\infty(I)}$$

and then use (18). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

For the interpolation error in the Gauss–Lobatto points, we have

LEMMA 11   Let u satisfy the assumptions of Lemma 9. Then there are $C, \sigma > 0$ depending only on γ of Lemma 9 and C_G such that

$$\|u - i_p u\|_{L^\infty(I)} + \|(u - i_p u)'\|_{L^\infty(I)} \leqslant C C_u e^{-\sigma p}.$$

Proof. Let P_p be the approximant constructed in Lemma 9. As the interpolation operator i_p reproduces polynomials of degree p, we have $u - i_p u = (u - P_p) - i_p(u - P_p)$. Applying Lemma 10 to the function $u - P_p$ and estimating $\|u - P_p\|_{L^\infty(I)}$, $\|(u - P_p)'\|_{L^\infty(I)}$ by means of Lemma 9 allows us to conclude the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

REMARK 12   As the endpoints ± 1 are sampling points of the Gauss–Lobatto interpolation operator, $u(\pm 1) = i_p u(\pm 1)$.

4. Main theorem

DEFINITION 13   Denote $\Pi_p(J)$ the set of all polynomials of degree p on the interval J. Let $\Delta = \{x_i \mid 0 \leqslant i \leqslant n\}$ be a collection of $n + 1 \geqslant 2$ nodal points satisfying

$$-1 = x_0 < x_1 < \cdots < x_n = 1.$$

These $n + 1$ nodal points define a mesh on Ω with elements $I_i = (x_{i-1}, x_i)$, $i = 1, ..., n$. On this mesh, define spaces of piecewise polynomials of degree p by

$$S^p(\Delta) := \{u \in H^1(\Omega) \mid u|_{I_i} \in \Pi_p(I_i), i = 1, ..., n\},$$
$$S_0^p(\Delta) := S^p(\Delta) \cap H_0^1(\Omega).$$

DEFINITION 14   *(Three-element mesh)* For $\kappa > 0$ and $p \in \mathbf{N}$ introduce spaces $S(\kappa, p) \subset H^1(\Omega)$ of piecewise polynomials of degree p by

$$S(\kappa, p) := \begin{cases} S^p(\Delta), & \Delta = \{-1, 1\} & \text{for } \kappa p d \geqslant 1/2 \\ S^p(\Delta), & \Delta = \{-1, -1 + \kappa p d, 1 - \kappa p d, 1\} & \text{for } \kappa p d < 1/2. \end{cases}$$

REMARK 15   We see that the spaces $S(\kappa, p)$ are based on three elements for the range of practical interest $p \ll d^{-1}$. In this case the elements at the boundary are of size $O(dp)$. Since we expect exponential rates of convergence for p sufficiently large, we switch to one element at $p = O(d^{-1})$; cf also the discussion following Theorem 1.

THEOREM 16   Let f, b be analytic on $\overline{\Omega}$, $b \geqslant \underline{b}^2$ on $\overline{\Omega}$. Let u_d be the solution of (1). Then there are $C, \tau, \kappa_0 > 0$ depending only on $f, b, \underline{b}, \alpha^\pm$ such that the following holds. For each $0 < \kappa \leqslant \kappa_0$ and $p \in \mathbf{N}$ there is $v_p \in S(\kappa, p)$ with $v_p(\pm 1) = u_d(\pm 1)$ and

$$\|u_d - v_p\|_{L^\infty(\Omega)} + \kappa p d \|(u_d - v_p)'\|_{L^\infty(\Omega)} \leqslant C(1 + \ln p) p^2 e^{-\tau \kappa p}.$$

REMARK 17   Let us stress that the constants κ_0, τ, and C are independent of the perturbation parameter d, the spectral order p, and the parameter κ. In particular, the dependence on the parameter κ is made explicit; we will make use of this fact in Section 5. Furthermore, the value of κ_0 is in principle accessible from the proof of Theorem 16. It depends on f and b. It can be shown that for the case of $b \equiv 1$, κ may be chosen arbitrarily in the open interval $(0, 4/e)$ as suggested by the analysis of Schwab & Suri (1996).

Proof. We will construct the approximant v_p explicitly by means of piecewise Gauss–Lobatto interpolants. Note that piecewise Gauss–Lobatto interpolants are automatically continuous and our construction will ensure $v_p(\pm 1) = u_d(\pm 1)$ (cf Remark 12).

Let us first consider the asymptotic case, i.e., $\kappa pd \geqslant 1/2$. By Theorem 1 we have

$$\|u_d^{(n)}\|_{L^2(\Omega)} \leqslant C K^n \max(n, d^{-1})^n \qquad \forall n \in \mathbf{N}_0.$$

We have furthermore

$$\max(d^{-1}, n)^n \leqslant \max(n! d^{-n}/n!, n^n) \leqslant \max(n! e^{1/d}, n^n) \leqslant \left(C e^{1/d}\right) e^n n! \qquad (19)$$

where we used Stirling's formula in the last estimate. Lemma 11 allows us to conclude that

$$\|u_d - i_p u_d\|_{L^\infty(\Omega)} + \|(u_d - i_p u_d)'\|_{L^\infty(\Omega)} \leqslant C e^{1/d} e^{-\sigma p}$$

for some $C, \sigma > 0$ independent of p and d. The assumption $\kappa pd \geqslant 1/2$ implies $e^{1/d} \leqslant e^{2\kappa p}$. Thus, assuming that κ_0 is chosen so small that $2\kappa \leqslant 2\kappa_0 \leqslant \sigma/2$, we obtain with $\kappa pd \geqslant 1/2$ and $d \leqslant 1$

$$\|u_d - i_p u_d\|_{L^\infty(\Omega)} + \kappa pd\|(u_d - i_p u_d)'\|_{L^\infty(\Omega)} \leqslant C \kappa pd e^{-(\sigma/2)p} \leqslant C \kappa_0 p e^{-2\kappa p}.$$

Let us now turn to the pre-asymptotic case $\kappa pd < 1/2$. The mesh consists of three elements $I_1 = (-1, -1 + \kappa pd)$, $I_2 = (-1 + \kappa pd, 1 - \kappa pd)$, and $I_3 = (1 - \kappa pd, 1)$. We decompose u_d as in (11) by writing $u_d = w_M + A_M^- u_d^- + A_M^+ u_d^+ + r_M$. The expansion order M is chosen as

$$2M = \mu \kappa p \qquad (20)$$

with $\mu > 0$ being a fixed parameter satisfying

$$\mu K_1 \frac{1}{2} \leqslant 1, \qquad \mu \frac{1}{2} K_4 =: q < 1 \qquad (21)$$

where the constants K_1 and K_4 are the constants of Theorems 3 and 6, respectively. Strictly speaking, we should take M as the integer part of $\mu \kappa p/2$—for notational convenience, however, we will ignore this point henceforth. The choice of μ guarantees that, as $\kappa pd \leqslant 1/2$,

$$2M d K_1 \leqslant \mu \kappa pd K_1 \leqslant \mu \frac{1}{2} K_1 \leqslant 1, \qquad 2M d K_4 \leqslant \mu \kappa pd K_4 \leqslant \mu \frac{1}{2} K_4 = q < 1,$$

and thus the assumptions of Theorem 3 are satisfied and the remainder r_M is indeed small by Theorem 6.

We will now analyze the approximation of each of the four terms in the decomposition

(11) by appropriate piecewise Gauss–Lobatto interpolants. Let us first consider w_M. By Theorem 3 and Lemma 11, the Gauss–Lobatto interpolant $i_p w_M$ on the whole domain Ω satisfies

$$\|w_M - i_p w_M\|_{L^\infty(\Omega)} + \|(w_M - i_p w_M)'\|_{L^\infty(\Omega)} \leqslant C e^{-\sigma p}$$

where the constants $C, \sigma > 0$ are independent of d, p, and κ. As $\kappa p d \leqslant 1/2$, the approximation of w_M by its Gauss–Lobatto interpolant satisfies the desired bounds if we make sure that τ and κ_0 are so small that $\tau\kappa \leqslant \tau\kappa_0 \leqslant \sigma$.

Let us now consider the approximation of the boundary layer parts. We will only provide the arguments for $A_M^- u_d^-$ as the arguments for the other boundary layer, $A_M^+ u_d^+$, are completely analogous. From Corollary 4 and our particular choice of μ, we have that the constant $A_M^- \leqslant C$ for some $C > 0$ independent of d. We may therefore concentrate on the approximation of u_d^-. We approximate u_d^- by its Gauss–Lobatto interpolants on the intervals $\overline{I}_1 = [-1, -1 + \kappa p d]$ and $\widetilde{I}_2 := \overline{I}_2 \cup \overline{I}_3 = [-1 + \kappa p d, 1]$. For the approximation of u_d^- on $\overline{I}_1$, denote l_1 the linear map from the reference element $I = [-1, 1]$ onto $\overline{I}_1$ and observe that Theorem 5 yields together with $l_1' = 2/(\kappa p d)$:

$$
\begin{aligned}
\left\|\left(u_d^- \circ l_1\right)^{(n)}\right\|_{L^\infty(I)} &\leqslant C K_3^n (\kappa p d/2)^n \max{(n, d^{-1})^n} \\
&\leqslant C(K_3/4)^n \max{(2\kappa p d n, 2\kappa p)^n} \\
&\leqslant C(K_3/4)^n \max{(n^n, n!(2\kappa p)^n/n!)} \leqslant C(e K_3/4)^n n! e^{2\kappa p}
\end{aligned}
$$

where we made use of the assumption that $\kappa p d \leqslant 1/2$ and argued as in estimate (19). Hence, Lemma 11 allows us to conclude that

$$\|u_d^- \circ l_1 - i_p\left(u_d^- \circ l_1\right)\|_{L^\infty(I)} + \|\left(u_d^- \circ l_1 - i_p\left(u_d^- \circ l_1\right)\right)'\|_{L^\infty(I)} \leqslant C e^{2\kappa p} e^{-\sigma p}$$

for some $C, \sigma > 0$ independent of d, p, and κ. Note that under the assumption $2\kappa \leqslant 2\kappa_0 \leqslant \sigma/2$, this is exponentially small. Scaling back, we obtain on I_1

$$
\begin{aligned}
\|u_d^- - i_p(u_d^- \circ l_1) \circ l_1^{-1}\|_{L^\infty(I_1)} &\leqslant C e^{-(\sigma/2)p} \leqslant C e^{-2\kappa p}, \\
\kappa p d \|\left(u_d^- - i_p(u_d^- \circ l_1) \circ l_1^{-1}\right)'\|_{L^\infty(I_1)} &\leqslant C e^{-(\sigma/2)p} \leqslant C e^{-2\kappa p}.
\end{aligned}
\tag{22}
$$

Let us now control u_d^- on $\widetilde{I}_2$. Denoting l_2 the linear map from the reference element I onto $\widetilde{I}_2$, Theorem 5 gives together with $l_2' = 2/(2 - \kappa p d) \leqslant 4/3$

$$\|u_d^- \circ l_2\|_{L^\infty(I)} \leqslant C e^{-\underline{b}\kappa p}, \qquad \|(u_d^- \circ l_2)'\|_{L^\infty(I)} \leqslant C d^{-1} e^{-\underline{b}\kappa p}.$$

These estimates together with Lemma 10 give for the approximation properties of the Gauss–Lobatto interpolant of $u_d^- \circ l_2$

$$
\begin{aligned}
\|u_d^- \circ l_2 - i_p\left(u_d^- \circ l_2\right)\|_{L^\infty(I)} &\leqslant C(1 + \ln p) e^{-\underline{b}\kappa p}, \\
\|\left(u_d^- \circ l_2 - i_p(u_d^- \circ l_2)\right)'\|_{L^\infty(I)} &\leqslant C(d^{-1} + (1 + \ln p)p^2) e^{-\underline{b}\kappa p}.
\end{aligned}
\tag{23}
$$

As $\kappa \leqslant \kappa_0$, $d \leqslant 1$, $l_2' \geqslant 1$, (23) allows us to conclude that on $\widetilde{I}_2$

$$\|u_d^- - i_p(u_d^- \circ l_2) \circ l_2^{-1}\|_{L^\infty(\widetilde{I}_2)} \leqslant C(1 + \ln p) e^{-\underline{b}\kappa p},$$

$$\kappa p d \|\left(u_d^- - i_p(u_d^- \circ l_2) \circ l_2^{-1}\right)'\|_{L^\infty(\widetilde{I}_2)} \leqslant C(p + (1 + \ln p)p^2) e^{-\underline{b}\kappa p}$$

for some C independent of d, p, and κ. Combining this estimate with (22) yields the desired estimate for the term u_d^- if $\tau \leqslant \min(2, \underline{b})$.

Let us finally turn our attention to r_M. We have by Theorem 6 and the embedding theorem

$$\|r_M\|_{L^\infty(\Omega)} \leqslant C \|r_M\|_{H^1(\Omega)} \leqslant Cd(2MdK_4)^{2M},$$
$$\|r_M'\|_{L^\infty(\Omega)} \leqslant C \|r_M\|_{H^2(I)} \leqslant C(2MdK_4)^{2M}.$$

As $2MdK_4 = \mu p dK_4 = q < 1$ by the choice of M and μ, we obtain by Lemma 10 for the Gauss–Lobatto interpolant $i_p r_M$ on the whole domain Ω

$$\|r_M - i_p r_M\|_{L^\infty(\Omega)} \leqslant Cq^{2M}d(1 + \ln p)$$
$$\|(r_M - i_p r_M)'\|_{L^\infty(\Omega)} \leqslant C \left[q^{2M} + C_G(1 + \ln p)p^2 q^{2M} d\right].$$

As $2M$ is proportional to κp, $q < 1$, and $\kappa p d \leqslant 1/2$, the approximation of the remainder r_M by its Gauss–Lobatto interpolant also satisfies the desired bounds if $\tau \leqslant \mu |\ln q|$. □

Theorem 16 leads to the following corollary for the FEM discretization (8).

COROLLARY 18   In (8) set $S_0 := S(\kappa, p) \cap H_0^1(\Omega)$ and $S_D := l(x) + S_0$ with $l(x) = \alpha^-(1 - x)/2 + \alpha^+(x + 1)/2$. Denote u_{FE} the finite element solution of (8). Then there is $\kappa_0 > 0$ depending only on the input data f, b, and $\alpha^\pm$ such that for all $0 < \kappa \leqslant \kappa_0$ there are $C, \sigma > 0$ independent of d, p such that

$$\|u_d - u_{FE}\|_d \leqslant Ce^{-\sigma p}.$$

Proof. This follows immediately from the quasi-optimality result (9) and Theorem 16. As we observed in Remark 17, κ may be chosen arbitrarily in the interval $(0, 4/e)$ in the special case $b \equiv 1$. □

5. Implementational aspects

5.1 *Remarks on meshes*

The 'three-element' meshes considered in the preceding section are the 'minimal' meshes which lead to robust exponential convergence. However, in a practical implementation, the following two points may arise. Firstly, reasonable lower bounds on κ_0 are needed in order to choose κ in the definition of the 'three-element' mesh appropriately. Secondly, the 'three-element' meshes depend on the perturbation parameter d as well as the spectral order p. In practice, it is more convenient to fix a mesh and increase the polynomial degree rather than change the mesh each time the polynomial degree is increased.

The following proposition addresses the first point. We show that linking the parameter κ to the polynomial degree p avoids the necessity of explicit knowledge of κ_0 at the expense of marginally relaxing the error bounds.

PROPOSITION 19   Let u_d be the solution of (1) and assume that hypotheses (2)–(4) hold. For $p \geqslant 2$ choose $\kappa = 1/\ln p$ in the definition of $S(\kappa, p)$. Then there are C, σ depending

only on f, b, $\underline{b}$, $\alpha^{\pm}$ such that for all $p \geqslant 2$ there is $v_p \in S(\kappa, p) = S(1/\ln p, p)$ with $v_p(\pm 1) = u_d(\pm 1)$ and

$$\|u_d - v_p\|_{L^\infty(\Omega)} + d \left\|\left(u_d - v_p\right)'\right\|_{L^\infty(\Omega)} \leqslant C e^{-\sigma p/\ln p}.$$

Proof. Let κ_0 be as in the statement of Theorem 16. There is $p_0 \in \mathbf{N}$ such that $\kappa = 1/\ln p < \kappa_0$ if $p \geqslant p_0$. Hence Theorem 16 is applicable for $p \geqslant p_0$, and the assertion of the proposition follows from the error estimate given in Theorem 16. $\square$

Let us now address the second point. We will show that fixed meshes which are refined geometrically towards the endpoints $x = \pm 1$ share some of the approximation properties of the 'minimal', 'three-element' meshes. Let us introduce a *geometric mesh* with *grading factor* $0 < \sigma < 1$ and $L \in \mathbf{N}$ *layers* which is refined towards the left endpoint $x = -1$ by the $L + 2$ nodes $\Delta = \{x_i \mid i = 0,..., L + 1\}$ given by

$$x_0 = -1, \quad x_i = -1 + \sigma^{L-i+1}, \quad i = 1,..., L, \quad x_{L+1} = 1. \qquad (24)$$

Completely analogously we define meshes that are refined geometrically towards the right endpoint or both endpoints. For meshes that are refined geometrically towards both endpoints we have the following approximation result:

PROPOSITION 20   Let u_d be the solution of (1) and assume hypotheses (2)–(4). Let $C' > 0, 0 < \sigma < 1$ be fixed. Furthermore, let a mesh, geometrically refined towards the endpoints $x = \pm 1$, be given by the nodes $\Delta = \{-1, -1 + \sigma^i, 1 - \sigma^i, 1 \mid i = 1,..., L\}$ where the number of layers L satisfies $\sigma^L \leqslant C'd$. Then there are $C, \tau > 0$ depending only on f, b, $\underline{b}$, $\alpha^{\pm}$, and σ, C' such that for each $p \in \mathbf{N}$ there is $v_p \in S^p(\Delta)$ with $v_p(\pm 1) = u_d(\pm 1)$ and

$$\|u_d - v_p\|_{L^\infty(\Omega)} + d \left\|\left(u_d - v_p\right)'\right\|_{L^\infty(\Omega)} \leqslant C e^{-\tau p}.$$

Proof. Let κ_0 be as in Theorem 16. If $\kappa_0 p d \geqslant 1/2$ then $S(\kappa_0, p) = \Pi_p(\Omega) \subset S^p(\Delta)$, and hence the claim of the proposition follows immediately from Theorem 16. If $\kappa_0 p d < 1/2$, we may assume without loss of generality that the polynomial degree p satisfies $p \geqslant C'/(\sigma \kappa_0)$. It is easy to see that there are then $i \in \{1,..., L\}$ and $\kappa \in [\sigma \kappa_0, \kappa_0]$ such that $\kappa p d = \sigma^i$. Thus, for this κ we have $S(\kappa, p) \subset S^p(\Delta)$, and the error bounds of Theorem 16 together with $\sigma \kappa_0 \leqslant \kappa \leqslant \kappa_0$ imply the desired result. $\square$

REMARK 21   It should be noted that the number of elements in the geometric meshes of Proposition 20 does depend on the perturbation parameter d, albeit in a very weak way: The condition $\sigma^L \leqslant C'd$ stipulates that $L \sim |\ln d|$. Hence, in terms of the number of degrees of freedom $N = \dim S^p(\Delta)$, the error estimate of Proposition 20 reads

$$\|u_d - v_p\|_{L^\infty(\Omega)} + d \left\|\left(u_d - v_p\right)'\right\|_{L^\infty(\Omega)} \leqslant C e^{-\tau' N/|\ln d|}$$

for some $C, \tau' > 0$.

5.2   *Impact of numerical quadrature*

If the coefficient b and the right-hand side f are general (analytic) functions, any implementation of (8) has to resort to numerical quadrature for the evaluation of the stiffness

matrix and the load vector. We show that the use of composite Gaussian quadrature rules of order $p + q$ with q proportional to p is sufficient to retain the robust exponential convergence of the finite element method (8) based on the spaces $S(\kappa, p)$.

For ease of exposition, let us assume the case of homogeneous boundary conditions, $\alpha^{\pm} = 0$. For any interval J denote $G_J^q(w)$ the Gaussian quadrature rule with q points in the interval J applied to the function w. Let the mesh with n elements I_i, $i = 1,...,n$, be determined by a collection of nodes Δ and consider the finite element method (8) based on the space $S_0^p(\Delta)$. If the integrals are replaced with composite Gaussian quadrature rules with $p + q$ points in each element, we obtain the following method:

$$\text{find } u_G \in S_0^p(\Delta) \text{ such that } B_d^{p+q}(u_G, v) = F^{p+q}(v) \quad \forall v \in S_0^p(\Delta) \qquad (25)$$

where the bilinear form B_d^{p+q} and the right-hand side F^{p+q} are given by

$$B_d^{p+q}(u, v) = \sum_{i=1}^n \left(d^2 G_{I_i}^{p+q}(u'v') + G_{I_i}^{p+q}(b\,uv) \right), \quad F^{p+q}(v) = \sum_{i=1}^n G_{I_i}^{p+q}(f v).$$

THEOREM 22   Let f, b be analytic on $\overline{\Omega}$, u_d be the solution of (1) with $\alpha^{\pm} = 0$. Let u_G be defined by (25). Then there are constants $C, \sigma > 0$ depending only on the input data f, b such that for all $p, q \in \mathbb{N}$

$$\|u_d - u_G\|_d \leqslant C \left(\inf_{v \in S_0^p(\Delta)} \|u_d - v\|_d + e^{-\sigma q} \right).$$

The proof of Theorem 22 is postponed until the end of this section as a few technical lemmas are needed. Theorem 22 allows us to assess the impact of numerical quadrature. The first term, the infimum, may be estimated by $Ce^{-\sigma p}$ if the mesh is chosen properly as was shown in Section 4. Choosing $q = \tau p$, $\tau > 0$, also makes the second term exponentially small in the polynomial degree p. Hence, overintegration (measured by q) proportional to the polynomial degree p retains the exponential convergence of the finite element method.

The analysis of the quadrature error will be done in the framework of the first Strang lemma (see, e.g., Ciarlet (1978)):

LEMMA 23   (*First Strang Lemma*)   Assume that there is $\beta > 0$ such that

$$\beta \|u\|_d^2 \leqslant B_d^{p+q}(u, u) \qquad \forall u \in S_0^p(\Delta). \qquad (26)$$

Then (25) has a unique solution $u_G \in S_0^p(\Delta)$ satisfying

$$\|u_d - u_G\|_d \leqslant (1 + \beta^{-1}) \left\{ \inf_{v \in S_0^p(\Delta)} \left(\|u_d - v\|_d + \sup_{w \in S_0^p(\Delta)} \frac{|B_d(v, w) - B_d^{p+q}(v, w)|}{\|w\|_d} \right) \right.$$

$$\left. + \sup_{w \in S_0^p(\Delta)} \frac{|F(w) - F^{p+q}(w)|}{\|w\|_d} \right\}.$$

Furthermore, we need the following lemma.

LEMMA 24   Let $p, q \in \mathbf{N}$, $w \in C^0(\overline{J})$. Then for all polynomials u, v of degree p

$$\left| \int_J uwv \, dx - G_J^{p+q}(uwv) \right| \leqslant 2 \, \|u\|_{L^2(J)} \|v\|_{L^2(J)} \inf_{\pi_{2q-1}} \|w - \pi_{2q-1}\|_{L^\infty(J)}$$

where the infimum is taken over all polynomials π_{2q-1} of degree $2q - 1$.

Proof. Let π_{2q-1} be a polynomial of degree $2q - 1$. The function $uv \, \pi_{2q-1}$ is a polynomial of degree $2(p + q) - 1$ which is integrated exactly by the Gaussian quadrature rule with $p + q$ points. Hence

$$\left| \int_J uwv \, dx - G_J^{p+q}(uwv) \right| \leqslant \left| \int_J uv(w - \pi_{2q-1})dx - G_J^{p+q}\big(uv(w - \pi_{2q-1})\big) \right|.$$

The integral may be estimated using Schwarz's inequality by

$$\left| \int_J uv(w - \pi_{2q-1})dx \right| \leqslant \|u\|_{L^2(J)} \|v\|_{L^2(J)} \|w - \pi_{2q-1}\|_{L^\infty(J)}.$$

For the second term, we observe that Schwarz's inequality for sums together with the positivity of the weights of the Gaussian quadrature rule leads to

$$\left| G_J^{p+q}\big(uv(w - \pi_{2q-1})\big) \right| \leqslant \sqrt{G_J^{p+q}(u^2)} \sqrt{G_J^{p+q}(v^2)} \|w - \pi_{2q-1}\|_{L^\infty(J)}.$$

As $q \geqslant 1$, the Gaussian quadrature formulae integrate the polynomials u^2, v^2 exactly, and we obtain the desired bound.   □

Proof of Theorem 22. Let $u \in S_0^p(\Delta)$. As $q \geqslant 1$, the Gaussian quadrature formulae $G_{I_i}^{p+q}$ integrate the polynomials $|u'|^2$, u^2 exactly. Hence,

$$B_d^{p+q}(u, u) = \sum_{i=1}^n d^2 G_{I_i}^{p+q}(|u'|^2) + G_{I_i}^{p+q}(b|u|^2)$$

$$\geqslant \sum_{i=1}^n d^2 \int_{I_i} |u'|^2 \, dx + \underline{b}^2 G_{I_i}^{p+q}(|u|^2) \geqslant \sum_{i=1}^n d^2 \int_{I_i} |u'|^2 + \underline{b}^2 |u|^2 \, dx$$

$$\geqslant \min\left(1, \frac{\underline{b}^2}{\|b\|_{L^\infty(\Omega)}}\right) \|u\|_d^2 =: \beta \|u\|_d^2.$$

Thus, assumption (26) of Lemma 23 is satisfied with a constant $\beta > 0$ independent of p, q. Let us now control the two consistency errors appearing in Lemma 23. We have for u, $w \in S_0^p(\Delta)$

$$B_d(u, w) - B_d^{p+q}(u, w)$$

$$= \sum_{i=1}^n d^2 \int_{I_i} u'w' \, dx - d^2 G_{I_i}^{p+q}(u'w') + \int_{I_i} buv \, dx - G_{I_i}^{p+q}(buw)$$

$$= \sum_{i=1}^n \int_{I_i} buv \, dx - G_{I_i}^{p+q}(buw)$$

because polynomials of degree $2(p + q) - 1 \geqslant 2p$ are integrated exactly by the Gaussian quadrature rules. The analyticity of b, Lemma 9, and the scaling invariance of the L^∞ norm imply that there are constants $C, \sigma > 0$ depending only on b such that for each element I_i there is a polynomial P_i of degree $2q - 1$ satisfying

$$\|b - P_i\|_{L^\infty(I_i)} \leqslant Ce^{-\sigma q}.$$

Appealing to Lemma 24 allows us to conclude

$$\left| B_d(u, w) - B_d^{p+q}(u, w) \right| \leqslant \sum_{i=1}^n 2\|u\|_{L^2(I_i)}\|w\|_{L^2(I_i)}\|b - P_i\|_{L^\infty(I_i)}$$

$$\leqslant Ce^{-\sigma q}\|u\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)} \leqslant Ce^{-\sigma q}\|u\|_d\|v\|_d$$

by the Schwarz inequality for sums. This allows us to estimate the first consistency term by $Ce^{-\sigma q}$. The second consistency error is controlled in a similar manner. □

REMARK 25   Inspection of the proof of Theorem 22 shows that the effect of the numerical quadrature is controlled by the approximability of f, b by piecewise (discontinuous) polynomials of degree $2q - 1$ on the mesh given by the nodes Δ. Therefore, the impact of numerical quadrature can be estimated as stated in Theorem 22 if the data f, b are merely piecewise analytic, provided that the points of non-analyticity of f and b are nodal points of the mesh.

Similarly, if composite Gaussian quadrature rules are used on each element, the impact of numerical quadrature can be assessed by the approximability of the data by piecewise polynomials on the submesh. For example, if the right-hand side f is (nearly) singular as in the numerical examples of the next section, properly selecting submeshes for the numerical quadrature guarantees that the quadrature error does not contribute significantly to the overall error.

6. Numerical examples

In this section we want to present a few numerical examples to illustrate Theorem 16. We consider the problem

$$-d^2u'' + u = f_a(x) := (a + x)^{-0.45} \qquad \text{on } \Omega = (-1, 1),$$
$$u(\pm 1) = 0, \tag{27}$$

where $a \geqslant 1$ is a parameter. Note that in the case $a > 1$, the right-hand side f_a is analytic on $[-1, 1]$, and thus Theorem 16 and Corollary 18 apply. For the case $a = 1$, (27) is still a well-posed problem for all $d > 0$ as $f_1 \in L^2(-1, 1)$. However, as the right-hand side is not analytic on the closed set $[-1, 1]$, the mathematical theory developed in this paper does not cover this case. Nonetheless, this case is interesting as it is a one-dimensional model for two-dimensional problems with corner singularities.

For (27), a particular solution is given by

$$u_{part}(x) = -\frac{1}{d}\int_{-1}^x \sinh\frac{x - t}{d} f_a(t)\, dt,$$

and hence the solution is seen to be

$$u(x) = u_{part}(x) - \frac{\sinh(x+1)/d}{\sinh(2/d)} u_{part}(1). \tag{28}$$

6.1   *Description of the numerical set-up*

Figures 1–8 show the results of the numerical experiments for various choices of a and d. The parameter a is chosen to range from the relatively smooth case of $a = 1 \cdot 1$ down to the limiting case $a = 1 \cdot 0$. The perturbation parameter d varies from $d = 10^{-2}$ to $d = 10^{-8}$. In all the graphs, we report the relative error in energy versus the number of degrees of freedom; here the energy is the square of the energy norm defined in (6). All calculations are performed using MATLAB, i.e., with double precision (16 digits) accuracy. Quadrature rules of sufficient accuracy are used in order to be able to concentrate on the finite element error (cf Section 5.2 and Remark 25).

We consider in our calculations the p version of the FEM with three types of mesh, that is, the finite element formulation (8) is used with three different choices of finite element spaces $S_D = S_I$, S_{II}, and S_{III}, each one consisting of continuous, piecewise polynomials of degree p. The first type of mesh is the 'three-element' mesh, i.e., the mesh consists of two small elements of size κpd near the endpoints $x = \pm 1$ and one large element in the middle. In the notation of Definitions 13, 14, the finite element spaces are the spaces

$$S_I(p) := S_0^p(\Delta_I) = S(\kappa, p) \cap H_0^1(-1, 1)$$
$$\Delta_I := \{-1, -1 + \kappa pd, 1 - \kappa pd, 1\}, \qquad \kappa = 0 \cdot 71.$$

The choice $\kappa = 0 \cdot 71$ was previously made by Schwab & Suri (1996) for their calculations, and we refer to the discussion there for the optimal choice of κ. Let us point out that the numerical studies of Schwab & Suri (1996) indicate that for the present problem (27) the 'three-element' approach is fairly insensitive to the precise choice of κ as long as it stays away from 0 and $4/e$.

For the other two types of mesh, we use 'unions' of boundary layer meshes and geometric meshes. In our particular examples, all geometric mesh refinements are towards the left endpoint $x = -1$ as in (24). In all our calculations, we use the grading factor $\sigma = 0 \cdot 15$ (cf Babuška & Gui (1986) for a justification of this choice). The 'union' of such a geometric mesh and the 'three-element' mesh is a mesh with $L + 3$ elements whose nodes are given by the nodes of (24) and the two additional nodes $-1 + \kappa pd$, $1 - \kappa pd$ (again with $\kappa = 0 \cdot 71$). The finite element spaces are then given by

$$\Delta_{II}(L) := \{-1, -1 + \sigma^i, -1 + \kappa pd, 1 - \kappa pd, 1 \mid i = 1,..., L\}, \sigma = 0 \cdot 15, \kappa = 0 \cdot 71$$
$$S_{II}(L, p) := S_0^p(\Delta_{II}(L)).$$

Finally, the third type of mesh considered in our numerical studies is a 'union' of the boundary layer mesh at the right endpoint and the geometric mesh (24) at the left endpoint. The mesh is determined by the nodes of (24) and one additional node at $1 - \kappa pd$. Hence it is a mesh with $L + 2$ elements, and we get for the finite element spaces

$$\Delta_{III}(L) := \{-1, -1 + \sigma^i, 1 - \kappa pd, 1 \mid i = 1,..., L\}, \qquad \sigma = 0 \cdot 15, \kappa = 0 \cdot 71$$
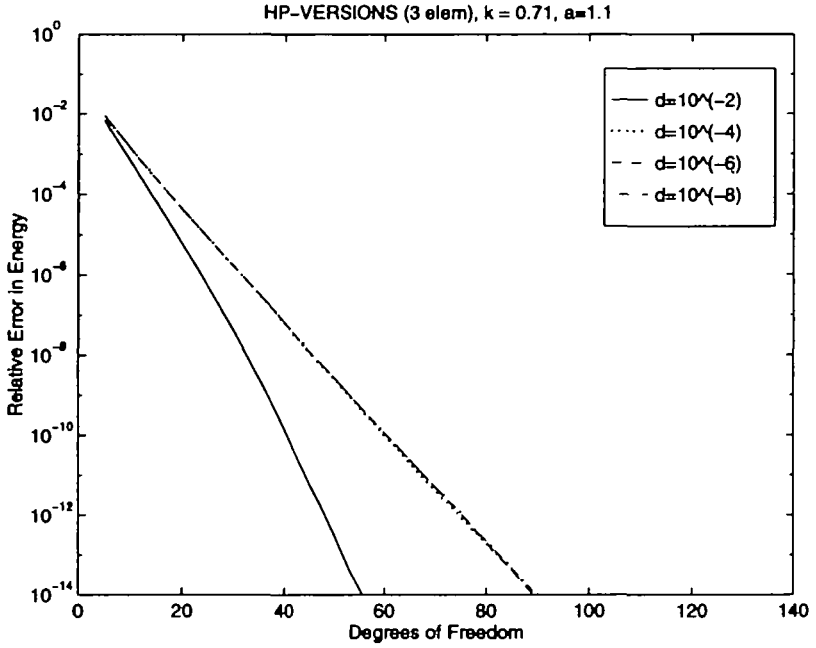$$S_{III}(L, p) := S_0^p(\Delta_{III}(L)).$$

HP-VERSIONS (3 elem), k = 0.71, a=1.1

FIG. 1. Three-element mesh for various d; $a = 1\cdot1$

HP-VERSIONS (3 elem), k = 0.71, a=1.01

FIG. 2. Three-element mesh for various d; $a = 1\cdot01$
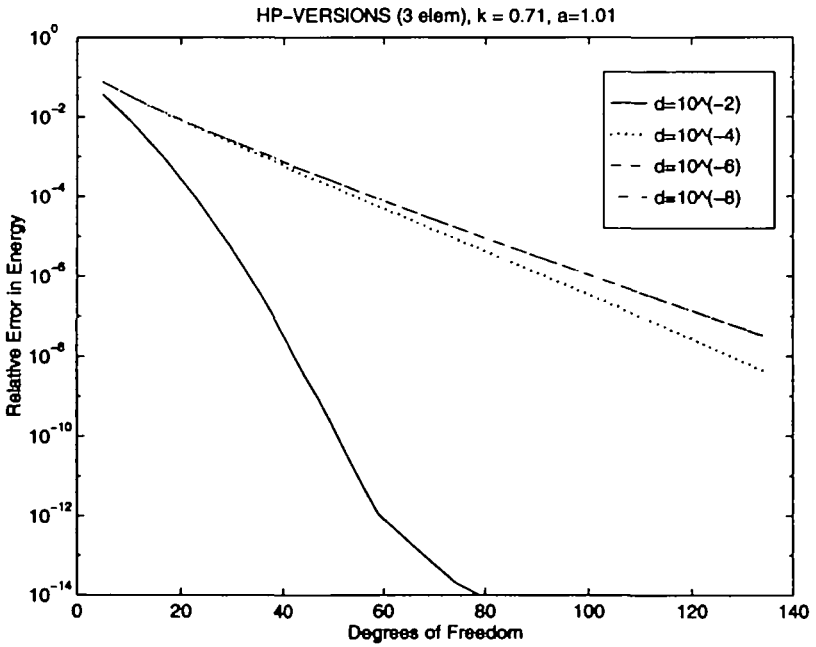
HP-VERSIONS (3 elem), k = 0.71, a=1.001
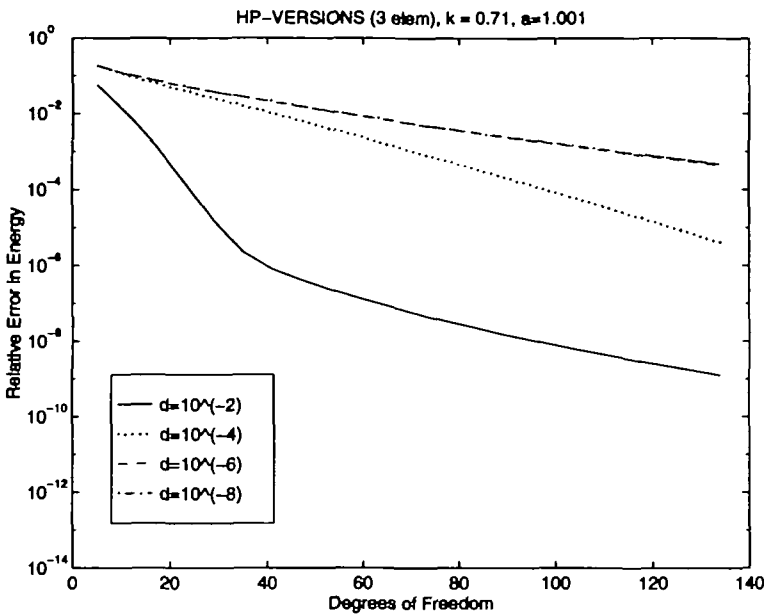
FIG. 3. Three-element mesh for various d; $a = 1.001$

6.2   Discussion of the numerical results

6.2.1   *The case $a > 1$.*

Let us first consider the case $a > 1$. In Figs. 1–3 we see the performance of the 'three-element' approach for $a = 1.1$, $a = 1.01$, and $a = 1.001$ and various choices of d. We indeed see that the 'three-element' approach, i.e., the finite element spaces $S_l(p)$, yields robust exponential convergence: The error curves are practically straight lines in the semi-logarithmic plot which indicates exponential convergence, and for fixed a, the error curves tend to a limiting curve as d approaches zero (the curves for $d = 10^{-6}$ and $d = 10^{-8}$ are practically on top of each other in Figs. 1–3), which is in agreement with our theoretical results on robustness.

As a approaches 1, the overall approximation rate deteriorates. This is to be expected from the proof of Theorem 16. Essentially, the solution is split into an asymptotic part and a boundary layer part. The boundary layer part can be approximated well with the aid of the two small elements at the endpoints. However, the approximation of the asymptotic part is poor on the large middle element of size $O(1)$ if a is close to 1. In the limiting case $a = 1$, the exponential rate of convergence breaks down with the 'three-element' mesh (cf Fig. 4).

For the case a close to 1, we are thus led to considering meshes designed such that the asymptotic part as well as the boundary layer can be approximated well. As the 'three-element' mesh can approximate the boundary layer well, let us turn our attention to the approximation of the asymptotic part. The asymptotic part is analytic on $[-1, 1]$ but has a singularity at $x = -a < -1$. Piecewise polynomials on meshes which are graded *geometrically* towards the point closest to the singularity (here: the left endpoint $x = -1$)
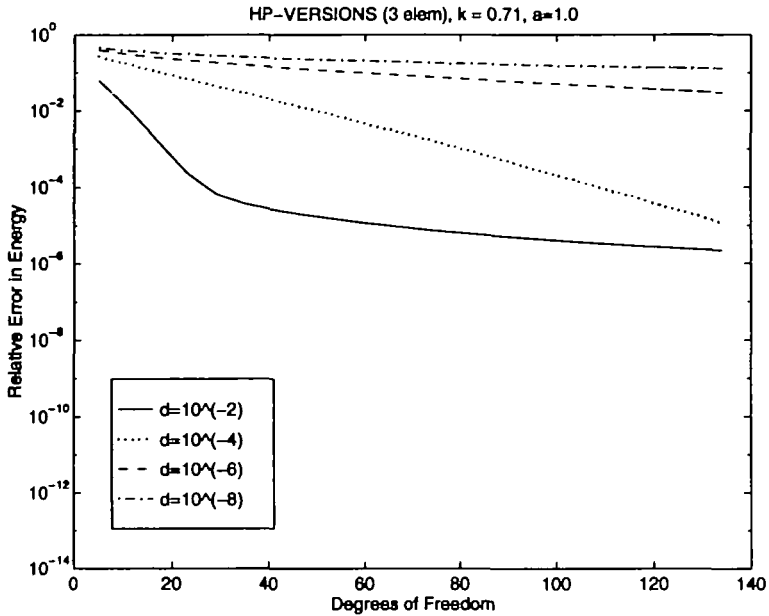
FIG. 4. Three-element mesh for various d; $a = 1 \cdot 0$

deal very successfully with this kind of singularity. A simple scaling argument suggests that the mesh should be chosen such that the ratio of the length of the elements to the distance (of the elements) to the singularity is bounded from below. This can be achieved with geometric meshes of the type (24) where the number of layers L is such that the smallest element (x_0, x_1) has length proportional to dist(a, Ω), i.e., $\sigma^L \sim |a - 1|$. For $\sigma = 0 \cdot 15$ and $a = 1 \cdot 01$, we can choose $L = 2$ and for $a = 1 \cdot 001$ we may choose $L = 3$ to get $\sigma^L \sim |a - 1|$. As the spaces $S_{II}(2, p)$ and $S_{III}(3, p)$ are based on the 'union' of these geometric meshes and the 'three-element' meshes, we expect these *ansatz* spaces to perform well for the nearly singular cases $a = 1 \cdot 01$ and $a = 1 \cdot 001$. We report the numerical results of the FEM with these *ansatz* spaces in Figs. 5, 6. We see that indeed this 'union' of the geometric mesh and the 'three-element' mesh leads to a very robust method both with respect to the perturbation parameter d and the parameter a.

6.2.2 *The case $a = 1$.* Let us now turn our attention to the limiting case $a = 1$. This case is of course no longer covered by the mathematical theory of this paper as the right-hand side is no longer differentiable at $x = -1$. We study this case numerically because it is a simple one-dimensional model for various singularly perturbed elliptic–elliptic problems. For example, as alluded to in the introduction, the solution of the Reissner–Mindlin plate model with polygonal mid-plane has corner singularities. The limiting solution (as the thickness of the plate tends to zero) is the solution of the Kirchhoff plate equation. The solution of that equation also has corner singularities which are, however, of a different character than those of the Reissner–Mindlin equation.
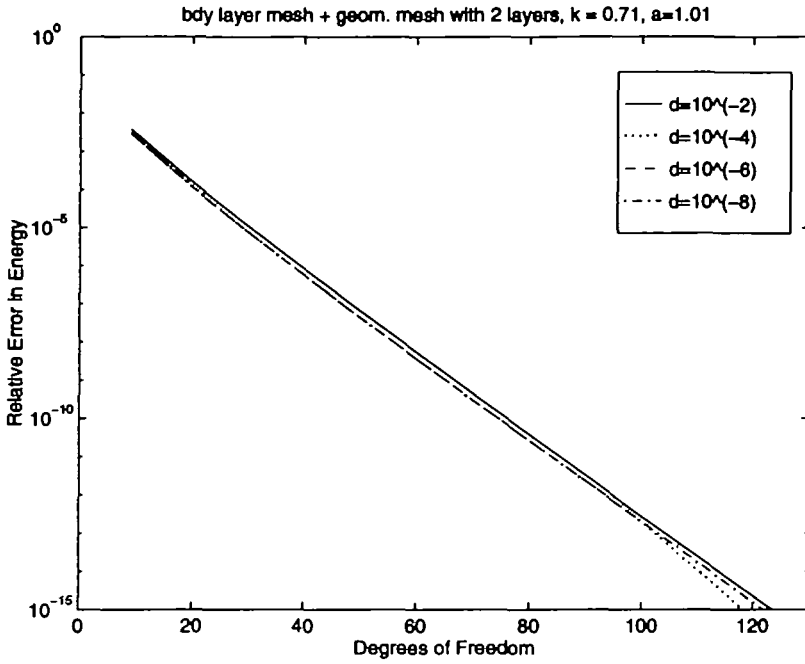
FIG. 5. Union of geometric mesh and three-element mesh; $a = 1·01$, 2 layers

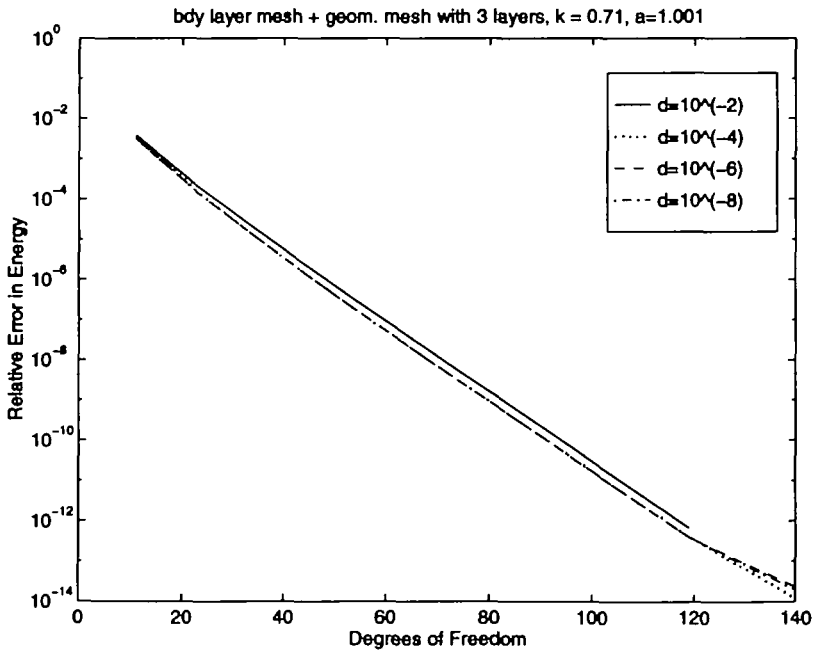

FIG. 6. Union of geometric mesh and three-element mesh; $a = 1·001$, 3 layers

For $a = 1$, the solution u of (27) has a (weak) singularity at the left endpoint $x = -1$: From (28) we see that the leading *singular* part of the solution u is

$$u \sim -\frac{1}{(1 - 0 \cdot 45)(2 - 0 \cdot 45)d^2}(x + 1)^{2 - 0 \cdot 45} \tag{29}$$

in an $O(d)$ neighbourhood of $x = -1$. On the other hand, as d tends to zero, the limiting solution is the singular function f_1 which has a different, stronger singularity at $x = -1$. It is reasonable to expect that the proper mesh should be such that it can resolve boundary layers, if present, that it can deal with the (weak) singularity at $x = -1$ for $d > 0$, and that it can also resolve the stronger singular behaviour of the limiting solution f_1. Functions of the type (29) and the limiting solution f_1 are analytic on Ω but have a singularity at the left endpoint. The approximation of such functions by the hp version of the FEM was analyzed by Babuška & Gui (1986). The mesh proposed there is a geometric mesh where the number of layers is proportional to p, the polynomial degree. For our problem, we therefore advocate a union of the 'three-element' mesh and such a geometrically refined mesh.

Let us remark that the hp FEM analyzed by Babuška & Gui (1986), which is based on geometric meshes with a number of layers proportional to p, has several noteworthy features. Firstly, exponential rates of convergence can be achieved in this way for the approximation of singular functions of the type considered here. Secondly, the *same* geometric mesh achieves this exponential convergence for singular functions of the type (29) and f_1 *simultaneously*. This is a very convenient feature of p or hp extensions. If singular functions are approximated using the h version of the FEM, the optimal, 'radical' meshes depend strongly on the type of singularity (cf Babuška & Gui (1986)). Hence, an h-version approach with optimal mesh design would be much more complicated in this situation as the type of mesh has to change as d tends to zero.

Let us return to the use of geometric meshes for our problem. Although the geometric meshes of Babuška & Gui (1986) are optimal in some sense and lead to exponential rates of convergence, it is more convenient in practical applications to fix a mesh and increase the polynomial degree p. This approach is taken in the numerical experiments of Figs. 7, 8. In those computations, the perturbation parameter d is chosen as $d = 10^{-4}$ or $d = 10^{-6}$, and a 'union' of a geometric mesh at the left endpoint and a boundary layer mesh for the right endpoint is used, i.e., the finite element spaces are $S_{III}(L, p)$ for various numbers of layers L. This 'union' thus takes care of the boundary layer at the right endpoint $x = 1$. As the mesh design at the left endpoint $x = -1$ is independent of the perturbation parameter d, we cannot expect the finite element spaces $S_{III}(L, p)$ to perform completely robustly with respect to d. However, as soon as the number of layers L is $O(|\ln d|)$, the smallest element of the geometric mesh is of size $O(d)$ (cf Proposition 20, Remark 21); the dependence of the method on the perturbation parameter d is therefore rather weak as geometric meshes with few layers create elements of size $O(d)$. For $d = 10^{-4}$ or $d = 10^{-6}$ and $\sigma = 0 \cdot 15$ already $L = 2$, resp. $L = 4$, leads to meshes whose smallest element at the left endpoint is $O(d)$.

Because the mesh is fixed and the solution has a singularity at $x = -1$, the asymptotic rate of convergence is algebraic and given by that of the p version of the FEM. For singular functions of the type (29), the analysis of Babuška & Gui (1986) shows that the H^1 error of the smallest element behaves like $p^{-2(1 - 0 \cdot 45) - 1}$, i.e., a rate of $DOF^{-4 \cdot 2}$ for the asymptotic
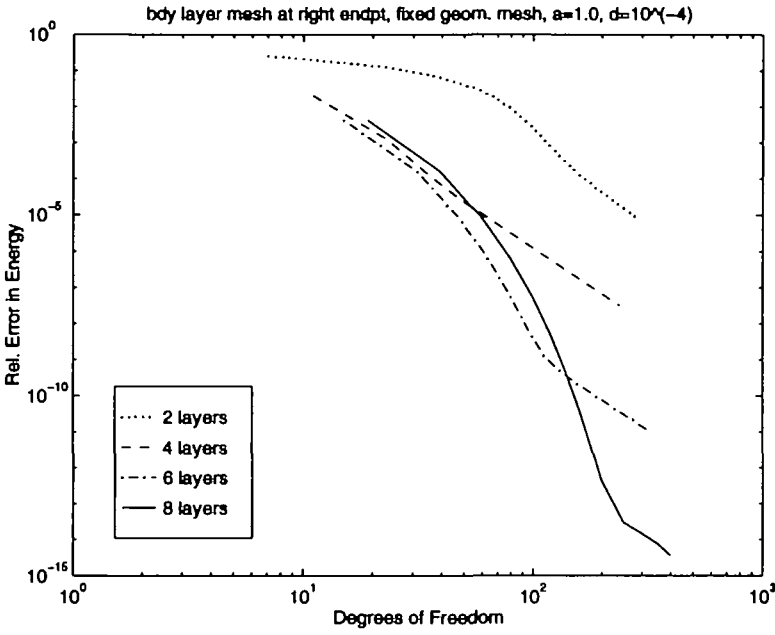
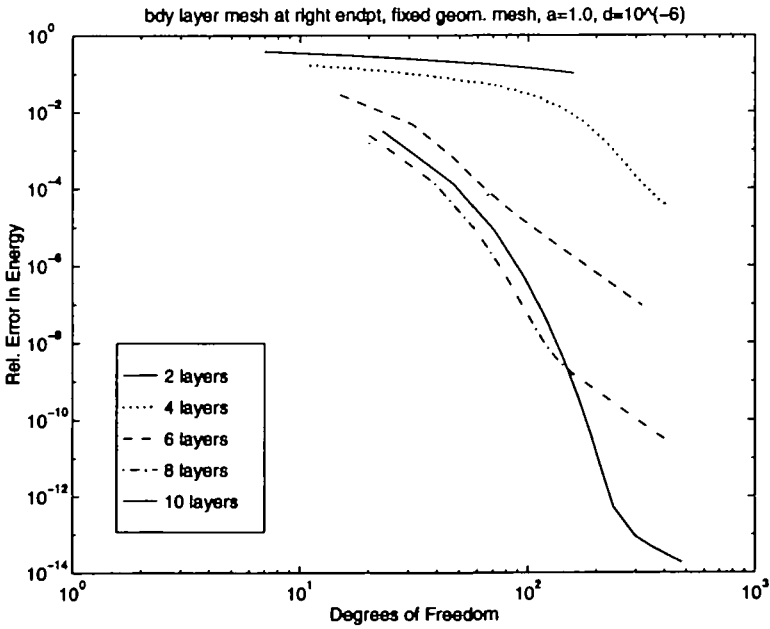FIG. 7. Geometric mesh at $x = -1$, boundary layer mesh at $x = 1$; $a = 1.0$, $d = 10^{-4}$

FIG. 8. Geometric mesh at $x = -1$, boundary layer mesh at $x = 1$; $a = 1.0$, $d = 10^{-6}$

error behaviour in the energy. Indeed, asymptotically, the error curves in Figs. 7, 8 are practically straight lines with slopes slightly over 4. However, depending on the number of layers, the pre-asymptotic range can be quite large: for example, in Fig. 7 with 6 layers, the asymptotic behaviour is not visible until the global error in energy is $\simeq 10^{-9}$. For a mesh with 8 layers, the asymptotic behaviour does not start until the energy error is 10^{-14}. In this pre-asymptotic range the global error is not dominated by the error in the first element abutting on the singularity. Rather, the error reduction is determined by the error reduction possible in the elements away from the singularity. There, exponential rates of convergence (in p) are possible and this exponential rate of convergence is visible in the pre-asymptotic regime.

6.2.3 *The final conjecture.*
We saw that in the case of an unsmooth right-hand side the use of a 'union' of a geometric mesh with a 'three-element' mesh is very successful. This is due to two facts. Firstly, 'three-element' meshes are designed such that the two small elements can resolve the boundary layers well. Secondly, geometric meshes can absorb both the singular behaviour of the solution for positive d as well as the singular behaviour of the limiting solution for $d = 0$. This 'union' of meshes is therefore very versatile.

For practical purposes, the introduction of small boundary layer elements is unnecessary at those boundary points towards which a strong geometric refinement is done because the geometric mesh leads to small elements of size $O(d)$ with fairly few layers. For problems whose solution is not analytic up to the boundary (as in the case $a = 1$), the use of a fixed geometric mesh can, asymptotically, only lead to algebraic rates of convergence. However, the use of a sufficient number of layers ensures that the asymptotic behaviour of the p version is pushed beyond the practical ranges of polynomial degrees p, and we have the pre-asymptotic exponential convergence.

7. Concluding remarks

In the present paper, we analyzed the hp FEM for a one-dimensional singularly perturbed problem of elliptic–elliptic type. We showed that for analytic input data, the introduction of two small elements of size $O(pd)$ near the boundary leads to robust exponential convergence of the hp FEM.

Although we considered a simple model problem, the techniques used here apply to more general situations. The essential tool for the proof of the approximation result Theorem 16 are classical asymptotic expansions for which the asymptotic part as well as the remainder can be controlled explicitly in terms of the perturbation parameter d and the expansion order M (Theorems 3, 6). Similar asymptotic expansions hold true for the convection–diffusion equation with analytic coefficients. The *approximation result* Theorem 16 holds therefore for the convection–diffusion equations with analytic coefficients as well. Of course, as the solutions of the convection–diffusion equation have a boundary layer at the outflow boundary only, it would be enough to use two elements where the small element is located at the outflow boundary. Let us conclude our remarks on the convection–diffusion equation by stressing that stability of finite element methods for the convection–diffusion equation is, as opposed to the reaction–diffusion equation considered in this paper, a non-trivial issue; a stable hp FEM for the convection–diffusion equation

able to make use of robust exponential approximability results of the type proved in this paper is presented by Melenk & Schwab (1997b).

Finally, the ideas developed in this paper are not restricted to one-dimensional problems; this is demonstrated by Melenk & Schwab (1997a) who successfully employ the ideas of the present paper for the construction of a robust, exponentially converging *hp* FEM for a reaction–diffusion equation in two dimensions.

We confirmed our theoretical result of robust exponential convergence by numerical experiments. Additionally, we studied numerically the case when the solution and the limiting solution (as the perturbation parameter tends to zero) are singular. There, we advocated the use of a 'union' of the proper boundary layer mesh with a geometrically graded mesh which is able to absorb the singular behaviour of the solution and the limiting solution. We showed numerically that this approach leads to very satisfactory results.

Acknowledgement

REFERENCES

BABUŠKA, I., & GUI, W. 1986 The *p* and *hp* versions of the finite element method in one dimension. Part I: the error analysis of the *p* version. Part II: the error analysis of the *hp* version. *Numer. Math.* **49**, 577–612 and 613–57.

CIARLET, P. G. 1978 *The Finite Element Method for Elliptic Problems*. Amsterdam: North-Holland.

DAVIS, P. J. 1974 *Interpolation and Approximation*. New York: Dover.

MELENK, J. M., & SCHWAB, C. 1997a *hp* FEM for reaction-diffusion equations. Part I: robust exponential convergence. Part II: regularity. *Research Reports 97-03 and 97-04*, Seminar für Angewandte Mathematik, ETH Zürich, CH-8092 Zürich, Switzerland.

MELENK, J. M., & SCHWAB, C. 1997b An *hp* finite element method for convection-diffusion problems. *Research Report 97-05*, Seminar für Angewandte Mathematik, ETH Zürich, CH-8092 Zürich, Switzerland.

MILLER, J. J. H., O'RIORDAN, E., & SHISHKIN, G. I. 1996 *Fitted Numerical Methods for Singular Perturbation Problems*. Singapore: World Scientific.

MORTON, K. W. 1996 *Numerical Solution of Convection-diffusion Problems*. London: Chapman & Hall.

O'RIORDAN, E., & STYNES, M. 1986 A uniformly accurate finite-element method for a singularly perturbed one-dimensional reaction-diffusion problem. *Math. Comput.* **47**, 555–70.

ROOS, H.-G., STYNES, M., & TOBISKA, L. 1996 *Numerical Methods for Singularly Perturbed Differential Equations*. Berlin: Springer.

SCHWAB, C., & SURI, M. 1996 The *p* and *hp* version of the finite element method for a problem with boundary layers. *Math. Comput.* **65**, 1403–29.

SÜNDERMANN, B. 1980 Lebesgue constants in Lagrangian interpolation at the Fekete points. *Ergebnisberichte der Lehrstühle Mathematik III & VIII (Angewandte Mathematik) 44*, Universität Dortmund.