INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 2, Number 1, Pages 107–126

OPTIMAL UNIFORM CONVERGENCE ANALYSIS FOR A TWO-DIMENSIONAL PARABOLIC PROBLEM WITH TWO SMALL PARAMETERS

JICHUN LI

(Communicated by Zhimin Zhang)

Abstract. In this paper, we consider a two-dimensional parabolic equation with two small parameters. These small parameters make the underlying problem containing multiple scales over the whole problem domain. By using the maximum principle with carefully chosen barrier functions, we obtain the pointwise derivative estimates of arbitrary order, from which an anisotropic mesh is constructed. This mesh uses very finer mesh inside the small scale regions (where the boundary layers are located) than elsewhere (large scale regions). A fully discrete backward difference Galerkin scheme based on this mesh with arbitrary k-th ($k \geq 1$) order conforming rectangular elements is discussed. Note that the standard finite element analysis technique can not be used directly for such highly nonuniform anisotropic meshes because of the violation of the quasi-uniformity assumption. Then we use the integral identity superconvergence technique to prove the optimal uniform convergence $O(N^{-(k+1)} + M^{-1})$ in the discrete L^2-norm, where N and M are the number of partitions in the spatial (same in both the x- and y-directions) and time directions, respectively.

Key Words. Singular perturbation, anisotropic mesh and uniform convergence.

1. Introduction

Singular perturbation problems (SPPs) appear in many areas, such as in chemical kinetics, heterogeneous flow in porous media, periodic structures, and plate and shell problems, etc. Actually, "Such a situation is relatively common in applications, and this is one of the reasons that perturbation methods are a cornerstone of applied mathematics" [16, Preface]. Those small parameters make the underlying problems contain multiple scales spanning over the whole domain. It is well known that the solutions of singular perturbation problems usually undergo rapid changes within very thin layers near the boundary (boundary layers) or inside the problem domain (interior layers), where the small scales are located.

However, direct numerical simulation by using the standard finite element method to resolve such multiscale problems is very impractical due to the requirement of

Received by the editors June 14, 2004 and, in revised form, July 19, 2004.
2000 *Mathematics Subject Classification.* 65N30, 65N15.

huge computer memory and CPU time. For example, by using a linear finite element on a quasi-uniform mesh to solve the simple model

$$-\varepsilon^2 \triangle u + u = f(x,y) \quad \text{in } \Omega \subseteq \mathcal{R}^2, \quad u|_{\partial\Omega} = 0,$$

where $0 < \varepsilon \ll 1$ is a perturbation parameter, we can obtain the following global error estimate:

$$||u - u_h||_\varepsilon \leq C(\varepsilon + h)h||u||_{H^2(\Omega)},$$

where $||u||_\varepsilon = (\varepsilon^2||\nabla u||^2_{L^2(\Omega)} + ||u||^2_{L^2(\Omega)})^{1/2}$. Noticing the fact that [31, Lemma 2.1]:

$$||u||_{H^2(\Omega)} \leq C\varepsilon^{-2}||f||_{L^2(\Omega)}, \tag{1}$$

we see that, to ensure good approximation, the mesh size h must be in the order of $o(\varepsilon)$. Suppose that $\varepsilon = 10^{-6}$ (which is very common), then $h = o(10^{-6})$. Hence we will end up with 10^{12} unknowns, which is well out of the power of most today's computer resources.

In summary, solving SPPs is a very challenging task because of the fact that ε can be very small leads to notorious computational difficulties [25, pp.310]. Such difficulties have also been emphasized by many researchers [30, 11]. The challenging SPPs serve frequently as test models for new algorithms, e.g., in multigrid methods [14, Ch.10], domain decomposition methods [12], collocation methods [4, Ch.10], and adaptive methods [1, 32].

Recently, the standard finite element methods based on anisotropically refined meshes, which use different scales of mesh size in different subdomains, were proved to give uniform convergence, which is independent of the small perturbation parameters. However, most work was restricted to linear finite element and problems with one perturbation parameter [3, 19, 22, 29, 37]. More details about the unsolved problems in this area can be found in the most recent survey by Roos [28].

In this paper, we will consider the analysis of applying arbitrary order tensor-product finite elements on such highly nonuniform anisotropic mesh to a two-dimensional parabolic equation with two small parameters. The pointwise derivative estimates are essential in the construction of such an anisotropic mesh with optimal uniform convergence. Here we use the maximum principle [26] as our powerful tool to obtain those derivative estimates by carefully choosing all kinds of barrier functions. Then we use the integral identity superconvergence technique [23, 7, 37, 9] originally developed for superconvergence analysis on tensor-product finite elements. We like to remark that uniform convergence can not be obtained directly by the standard finite element analysis for such highly nonuniform anisotropic meshes because of the violation of the quasi-uniformity assumption [8, 5]. Special interpolation estimates have to be obtained on such anisotropic meshes [2]. Also asymptotical expansion or pointwise derivative estimates for the analytical solution has to be investigated in order to obtain such uniform convergence [21, 22].

For simplicity, here we focus on the following parabolic equation

$$L_{\varepsilon\mu}u \equiv \varepsilon\frac{\partial u}{\partial t} - \mu^2 a \triangle u + bu = f(x,y,t,\varepsilon,\mu) \quad \text{in } D \equiv \Omega \times (0,T], \tag{2}$$

$$u|_{\partial\Omega\times(0,T]} = 0, \quad u|_{t=0} = 0, \tag{3}$$

where $\Omega = (0,1)^2$, and the coefficients $a(x,y,t), b(x,y,t)$ and f are sufficiently smooth functions. Here $0 < \varepsilon \ll 1, 0 < \mu \ll 1$ are small parameters. Furthermore we assume

$$b > a\beta_2^2 > 0, \quad b > \beta_1 > 0,$$

where β_1 and β_2 are positive constants.

The asymptotic expansion for (2)-(3) was investigated in [6], which showed where all the boundary layers are located. By the maximum principle, we also find all the boundary layers' exact locations, out of which our anisotropic mesh is built for the standard arbitrary k-th ($k \geq 1$) order conforming tensor-product elements. Our anisotropic mesh separates the boundary layers (small scale regions) totally from other parts of the problem domain (large scale regions). We use very fine mesh inside the boundary layers, and much coarse mesh elsewhere. The mesh ratio can be as high as $1 : \varepsilon|\ln\varepsilon|$. Use of such a mesh [34, 12, 17, 3, 22, 19] is more intuitive than the widely discussed Shishkin type mesh [2, 11, 24, 30, 36, 37]. The fully discrete backward difference [35, pp.748] Galerkin approximation is discussed and optimal uniform convergence rates of $O(N^{-(k+1)} + M^{-1})$ in the discrete L^2-norm are proved for the k-th order conforming tensor-product elements, where N and M are the number of discretization intervals in the spatial and time directions, respectively. Here for simplicity, we use the same number of partitions in both the x- and y-directions.

The rest of the paper is organized as follows. The derivative estimates of arbitrary order for the analytic solution of (2)-(3) are presented in section 2. Our mesh and the Galerkin scheme are constructed in section 3. Section 4 are devoted to the introduction of a special interpolation operator and its interpolation estimates. In the last section, the optimal uniform convergence analysis in the discrete L^2-norm is given.

Throughout the paper, C (or C_i) will denote a generic positive constant, which may be of different value at each occurrence and independent of the mesh size and the perturbation parameter ε. We use the notation $||\cdot||_{k,p,\tau}$ for the standard Sobolev $W^{k,p}(\tau)$ norm defined on the set τ, and v_{ξ^k} for the k-th order derivative of v with respect to the variable ξ. For simplicity, we use $||\cdot||_{k,\tau}$ when $p = 2$.

2. The derivative estimates

Here we use the maximum principle [26] as our powerful tool to obtain derivative estimates for the analytical solution of (2)-(3). Such technique has been proven to be very useful for SPPs [27, 29, 19], the difficult is how to carefully choose all kinds of barrier functions.

Because of the technical difficulty of applying the maximum principle, from now on we assume that a is a positive constant. Such difficulty has been encountered by other researchers ([29, pp.720], [27, pp.50]). Furthermore, we assume that f satisfies the following conditions:

$$|f_{x^i t^k}(x,y,t,\varepsilon,\mu)| \leq C\varepsilon^{-k}(1 + \mu^{-i}e^{-\beta_2 x/\mu} + \mu^{-i}e^{-\beta_2(1-x)/\mu}) \text{ on } \overline{D}, \tag{4}$$

$$|f_{y^i t^k}(x,y,t,\varepsilon,\mu)| \leq C\varepsilon^{-k}(1 + \mu^{-i}e^{-\beta_2 y/\mu} + \mu^{-i}e^{-\beta_2(1-y)/\mu}) \text{ on } \overline{D}, \tag{5}$$

$$|f_{t^i}(x,y,t,\varepsilon,\mu)| \leq C(1 + \varepsilon^{-i}e^{-\beta_1 t/\varepsilon}) \text{ on } \overline{D}, \tag{6}$$

where $\overline{D} = [0,1]^2 \times [0,T],\ i \geq 0,\ k = 0,1.$

Theorem 2.1. *Suppose the solution u of (2)-(3) is sufficiently smooth on $\overline{D}$. Then under the assumptions (4)-(6), we have*

(*i*) $|u_{x^k}(x,y,t)| \le C(1+\mu^{-k}e^{-\beta_2 x/\mu}+\mu^{-k}e^{-\beta_2(1-x)/\mu})$ *on* $\overline{D}$, $\forall\ k \ge 1$,

(*ii*) $|u_{y^k}(x,y,t)| \le C(1+\mu^{-k}e^{-\beta_2 y/\mu}+\mu^{-k}e^{-\beta_2(1-y)/\mu})$ *on* $\overline{D}$, $\forall\ k \ge 1$.

Proof. We only present the proofs of (i) for $k = 1,2,3,4$ in the following Lemmas 2.2-2.7. From the proofs given below, it is not difficult to see that other higher order derivatives can be obtained by the inductive method. (ii) can be proved by symmetry. □

Let $\partial D = \partial\Omega \times (0,T] \bigcup \Omega \times \{t = 0\}$. Here we will make repeated use of the following weak maximum principle [26]:

Lemma 2.1. *For any functions $w(x,y,t) \in C^2(D) \cap C^0(\overline{D})$, if $w \ge 0$ on ∂D and $L_{\varepsilon\mu} w \ge 0$ on D, then $w \ge 0$ on $\overline{D}$.*

Lemma 2.2.

$$u_x|_{y=0,1} = u_x|_{t=0} = 0, \tag{7}$$

$$|u_x|_{x=0,1} \le C\mu^{-1}, \tag{8}$$

$$|u| \le C(1-e^{-\beta_2 x/\mu}) \le C, \quad (x,y,t) \in \overline{D}. \tag{9}$$

Proof. By the boundary condition (3), i.e.,

$$u|_{y=0,1} = u|_{t=0} = 0,$$

the proof of (7) follows directly.

Consider the barrier function $\phi = C(1-e^{-\beta_2 x/\mu})$, we have

$$\begin{aligned} L_{\varepsilon\mu}(\phi \pm u) &= aC\beta_2^2 e^{-\beta_2 x/\mu} + bC(1-e^{-\beta_2 x/\mu}) \pm f \\ &= aC\beta_2^2 + C(1-e^{-\beta_2 x/\mu})(b - a\beta_2^2) \pm f \\ &\ge 0, \quad \text{for sufficiently large } C, \end{aligned}$$

which along with Lemma 2.1 and the fact that $(\phi \pm u)|_{\partial D} \ge 0$ gives us

$$|u| \le \phi = C(1-e^{-\beta_2 x/\mu}), \quad \text{on } \overline{D},$$

which completes the proof of (9).

From (9) and the fact that $u|_{x=0} = 0$, we have

$$\begin{aligned} |u_x(0,y,t)| &= |\lim_{x\to 0^+} \frac{u(x,y,t)-u(0,y,t)}{x}| \le \lim_{x\to 0^+} |\frac{u(x,y,t)-u(0,y,t)}{x}| \\ &\le \lim_{x\to 0^+} \frac{C(1-e^{-\beta_2 x/\mu})}{x} = C\frac{\beta_2}{\mu} \le C\mu^{-1}. \end{aligned} \tag{10}$$

By the same technique, it is easy to prove that

$$|u| \le C(1-e^{-\beta_2(1-x)/\mu}), \quad \text{on } \overline{D},$$

from which and $u|_{x=1} = 0$, we obtain

$$\begin{aligned} |u_x(1,y,t)| &= |\lim_{x\to 1^-} \frac{u(x,y,t)-u(1,y,t)}{x-1}| \le \lim_{x\to 1^-} |\frac{u(x,y,t)-u(1,y,t)}{x}| \\ &\le \lim_{x\to 1^-} \frac{C(1-e^{-\beta_2(1-x)/\mu})}{1-x} = C\frac{\beta_2}{\mu} \le C\mu^{-1}, \end{aligned}$$

which together with (10) completes the proof of (8). □

Lemma 2.3.

$$|u_x| \leq C(1+\mu^{-1}e^{-\beta_2 x/\mu}+\mu^{-1}e^{-\beta_2(1-x)/\mu}), \quad (x,y,t)\in \overline{D}.$$

Proof. Consider the barrier function $\phi = C(1+\mu^{-1}e^{-\beta_2 x/\mu}+\mu^{-1}e^{-\beta_2(1-x)/\mu})$, we have

$$\begin{aligned} L_{\varepsilon\mu}(\phi \pm u_x) &= -aC\mu^{-1}\beta_2^2(e^{-\beta_2 x/\mu}+e^{-\beta_2(1-x)/\mu}) \\ &\quad +bC(1+\mu^{-1}e^{-\beta_2 x/\mu}+\mu^{-1}e^{-\beta_2(1-x)/\mu}) \pm (f_x - b_x u) \\ &= bC+(b-a\beta_2^2)C\mu^{-1}(e^{-\beta_2 x/\mu}+e^{-\beta_2(1-x)/\mu}) \pm (f_x-b_x u) \\ &\geq 0, \quad \text{for sufficiently large } C, \end{aligned} \tag{11}$$

where in the last step we used the assumption (4) and the estimate (9).

From Lemma 2.2, we have

$$(\phi \pm u_x)|_{\partial D} \geq 0, \quad \text{for sufficiently large } C,$$

which along with (11) and Lemma 2.1 completes the proof. ☐

Lemma 2.4.

$$|u_{x^2}| \leq C(1+\mu^{-2}e^{-\beta_2 x/\mu}+\mu^{-2}e^{-\beta_2(1-x)/\mu}), \quad (x,y,t)\in \overline{D}.$$

Proof. From (2)-(3), we have

$$u_{x^2}|_{x=0,1} = -\mu^{-2}a^{-1}f|_{x=0,1}, \tag{12}$$

$$u_{x^2}|_{y=0,1} = u_{x^2}|_{t=0} = 0, \tag{13}$$

from which we obtain

$$|\ u_{x^2}\ |_{\partial D} \leq C\mu^{-2}. \tag{14}$$

Consider the barrier function $\phi = C(1+\mu^{-2}e^{-\beta_2 x/\mu}+\mu^{-2}e^{-\beta_2(1-x)/\mu})$, we have

$$\begin{aligned} & L_{\varepsilon\mu}(\phi \pm u_{x^2}) \\ =\ & bC+(b-a\beta_2^2)C\mu^{-2}(e^{-\beta_2 x/\mu}+e^{-\beta_2(1-x)/\mu}) \pm (f_{x^2}-b_{x^2}u-2b_x u_x) \\ \geq\ & 0, \quad \text{for sufficiently large } C, \end{aligned} \tag{15}$$

where in the last step we used the assumption (4) and the obtained estimates for u and u_x.

The proof follows from Lemma 2.1, (14) and (15). ☐

Denote

$$g(x,y,t,\varepsilon,\mu) = a^{-1}\mu^{-2}[\frac{1-e^{-\beta_2(1-x)/\mu}}{1-e^{-\beta_2/\mu}}f(0,y,t,\varepsilon,\mu)+\frac{1-e^{-\beta_2 x/\mu}}{1-e^{-\beta_2/\mu}}f(1,y,t,\varepsilon,\mu)],$$

and

$$\overline{u} = u_{x^2} + g(x,y,t,\varepsilon,\mu). \tag{16}$$

Differentiating (2) twice with respect to x, it is not difficult to see that $\overline{u}$ satisfies

$$L_{\varepsilon\mu}\overline{u} \equiv \varepsilon\frac{\partial \overline{u}}{\partial t} - \mu^2 a \triangle \overline{u} + b\overline{u} = \overline{f}(x,y,t,\varepsilon,\mu) \quad \text{in } D, \tag{17}$$

$$\overline{u} = 0 \quad \text{on } \partial D, \tag{18}$$

where $\overline{f} = f_{x^2} - 2b_x u_x - b_{x^2} u + L_{\varepsilon\mu} g$. To obtain (18), we used the following compatibility conditions:

$$f(0,0,t,\varepsilon,\mu) = f(0,1,t,\varepsilon,\mu) = f(1,0,t,\varepsilon,\mu) = f(1,1,t,\varepsilon,\mu) = 0, \tag{19}$$

$$f(0,y,0,\varepsilon,\mu) = f(1,y,0,\varepsilon,\mu) = 0, \tag{20}$$

$$f(x,0,0,\varepsilon,\mu) = f(x,1,0,\varepsilon,\mu) = 0, \tag{21}$$

Similar compatibility conditions of (19) were obtained for a steady problem in Section 3 of [15]. Let us show how (20) can be obtained. By letting $t = 0$ in (2) and using the boundary condition (3), we have

$$\varepsilon \frac{\partial u}{\partial t}|_{t=0} = f(x,y,0,\varepsilon,\mu). \tag{22}$$

On the other hand, from the boundary conditions $u|_{x=0,1} = 0$, we have

$$u_t|_{x=0,1} = 0. \tag{23}$$

Substituting (23) into (22) with $x = 0, 1$ gives us (20). (19) and (21) can be obtained similarly.

Lemma 2.5.

$$\overline{u}_x|_{y=0,1} = \overline{u}_x|_{t=0} = 0, \tag{24}$$

$$|\overline{u}_x|_{x=0,1} \leq C\mu^{-3}. \tag{25}$$

Proof. The proof of (24) follows directly from the boundary conditions (18).
The proof of (25) can be obtained by the same technique used in Lemma 2.2.
Consider the barrier function $\phi = C\mu^{-2}(1 - e^{-\beta_2 x/\mu})$, we have

$$\begin{aligned} L_{\varepsilon\mu}(\phi \pm \overline{u}) &= aC\beta_2^2\mu^{-2} + C\mu^{-2}(1 - e^{-\beta_2 x/\mu})(b - a\beta_2^2) \pm \overline{f} \\ &\geq 0, \quad \text{for sufficiently large } C, \end{aligned}$$

which along with Lemma 2.1 and the fact that $(\phi \pm \overline{u})|_{\partial D} \geq 0$ gives us

$$|\overline{u}| \leq \phi = C\mu^{-2}(1 - e^{-\beta_2 x/\mu}), \quad \text{on } \overline{D}. \tag{26}$$

From (26) and (18), we have

$$\begin{aligned} |\overline{u}_x(0,y,t)| &\leq \lim_{x\to 0^+} \left| \frac{\overline{u}(x,y,t) - \overline{u}(0,y,t)}{x} \right| \\ &\leq \lim_{x\to 0^+} \frac{C\mu^{-2}(1 - e^{-\beta_2 x/\mu})}{x} = C\mu^{-2}\frac{\beta_2}{\mu} \leq C\mu^{-3}. \end{aligned} \tag{27}$$

Similarly, it is not difficult to obtain

$$|\overline{u}_x(1,y,t)| \leq C\mu^{-3},$$

which along with (27) finishes the proof of (25). □

Lemma 2.6.

$$|u_{x^3}| \leq C(1 + \mu^{-3}e^{-\beta_2 x/\mu} + \mu^{-3}e^{-\beta_2(1-x)/\mu}), \quad (x,y,t) \in \overline{D}.$$

Proof. Similar to Lemma 2.3, by considering the barrier function $\phi = C(1 + \mu^{-3}e^{-\beta_2 x/\mu} + \mu^{-3}e^{-\beta_2(1-x)/\mu})$, we have

$$\begin{aligned} L_{\varepsilon\mu}(\phi \pm \overline{u}_x) &= bC + (b - a\beta_2^2)C\mu^{-3}(e^{-\beta_2 x/\mu} + e^{-\beta_2(1-x)/\mu}) \pm (\overline{f}_x - b_x\overline{u}) \\ &\geq 0, \quad \text{for sufficiently large } C, \end{aligned} \tag{28}$$

where in the last step we used the assumption (4) and the obtained estimates for $u_{x^k}, k = 0, 1, 2$.

From Lemma 2.5, we have

$$(\phi \pm \overline{u}_x)|_{\partial D} \geq 0, \quad \text{for sufficiently large } C,$$

which along with (28) and Lemma 2.1 gives us

$$|\overline{u}_x| \leq C(1 + \mu^{-3}e^{-\beta_2 x/\mu} + \mu^{-3}e^{-\beta_2(1-x)/\mu}),$$

which together with the definition $\overline{u}$ of (16) gives us

$$|u_{x^3}| \leq |\overline{u}_x - g_x| \leq |\overline{u}_x| + |g_x| \leq C(1 + \mu^{-3}e^{-\beta_2 x/\mu} + \mu^{-3}e^{-\beta_2(1-x)/\mu}),$$

which completes the proof. □

Lemma 2.7.

$$|u_{x^4}| \leq C(1 + \mu^{-4}e^{-\beta_2 x/\mu} + \mu^{-4}e^{-\beta_2(1-x)/\mu}), \quad (x, y, t) \in \overline{D}.$$

Proof. The proof is very similar to the one given for Lemma 2.4.

From (17)-(18), we have

$$\overline{u}_{x^2}|_{x=0,1} = -\mu^{-2}a^{-1}\overline{f}|_{x=0,1}, \tag{29}$$

$$\overline{u}_{x^2}|_{y=0,1} = \overline{u}_{x^2}|_{t=0} = 0,, \tag{30}$$

from which and the estimate of $\overline{f}$ we obtain

$$|\ \overline{u}_{x^2}\ |_{\partial D} \leq C\mu^{-4}. \tag{31}$$

Consider the barrier function $\phi = C(1+\mu^{-4}e^{-\beta_2 x/\mu}+\mu^{-4}e^{-\beta_2(1-x)/\mu})$. By simple calculations, we obtain

$$\begin{aligned} & L_{\varepsilon\mu}(\phi \pm \overline{u}_{x^2}) \\ = \;& bC + (b - a\beta_2^2)C\mu^{-4}(e^{-\beta_2 x/\mu} + e^{-\beta_2(1-x)/\mu}) \pm (\overline{f}_{x^2} - b_{x^2}\overline{u} - 2b_x\overline{u}_x) \\ \geq \;& 0, \quad \text{for sufficiently large } C, \end{aligned} \tag{32}$$

where in the last step we used the assumption (4), the definitions of $\overline{u}$ and $\overline{f}$, and the obtained estimates for $u_{x^k}, 0 \leq k \leq 3$.

Lemma 2.1, (31) and (32) complete the proof. □

It is not difficult to see that the above proofs for Lemmas 2.5-2.7 can be carried out repeatedly to obtain higher order derivative estimates of u_{x^k} $(k > 4)$.

Lemma 2.8.

$$|u_t| \leq C(1 + \varepsilon^{-1}e^{-\beta_1 t/\varepsilon}), \quad (x, y, t) \in \overline{D}.$$

Proof. Consider the barrier function $\phi = C(1 + \varepsilon^{-1}e^{-\beta_1 t/\varepsilon})$. By simple calculations, we obtain

$$\begin{aligned} L_{\varepsilon\mu}(\phi \pm u_t) &= -C\beta_1\varepsilon^{-1}e^{-\beta_1 t/\varepsilon} + bC(1 + \varepsilon^{-1}e^{-\beta_1 t/\varepsilon}) \pm (f_t - b_t u) \\ &= bC + C\varepsilon^{-1}e^{-\beta_1 t/\varepsilon}(b - \beta_1) \pm (f_t - b_t u) \\ &\geq 0, \quad \text{for sufficiently large } C, \end{aligned} \tag{33}$$

where we used the assumption (6) and the estimate (9).

On the other hand, letting $t = 0$ in (2), we have

$$u_t(x, y, 0) = \varepsilon^{-1}f(x, y, 0, \varepsilon, \mu) \leq C\varepsilon^{-1}. \tag{34}$$

From the boundary conditions (3), we have

$$u_t|_{x=0,1} = u_t|_{y=0,1} = 0, \tag{35}$$

which along with (34) give us

$$(\phi \pm u_t)|_{\partial D} \geq 0,$$

which together with (33) and Lemma 2.1 finishes the proof. □

Denote

$$\tilde{g}(x, y, t, \varepsilon, \mu) = \varepsilon^{-1} f(x, y, 0, \varepsilon, \mu) e^{-\beta_1 t/\varepsilon}, \tag{36}$$

$$\tilde{u} = u_t(x, y, t) - \tilde{g}(x, y, t, \varepsilon, \mu), \tag{37}$$

Differentiating (2) once with respect to t, it is not difficult to see that $\tilde{u}$ satisfies

$$L_{\varepsilon\mu}\tilde{u} \equiv \varepsilon \frac{\partial \tilde{u}}{\partial t} - \mu^2 a \triangle \tilde{u} + b\tilde{u} = \tilde{f}(x, y, t, \varepsilon, \mu) \quad \text{in } D, \tag{38}$$

$$\tilde{u} = 0 \quad \text{on } \partial D, \tag{39}$$

where $\tilde{f} = f_t - b_t u - L_{\varepsilon\mu}\tilde{g}$. Note that the compability conditions (20)-(21), (34), and (35) were used to obtain (39).

Lemma 2.9.

$$|u_{t^k}| \leq C(1 + \varepsilon^{-k} e^{-\beta_1 t/\varepsilon}), \quad (x, y, t) \in \overline{D}, \quad k \geq 2.$$

Proof. The proof is similar to the one given for Lemma 2.8. Consider the barrier function $\phi = C(1 + \varepsilon^{-2} e^{-\beta_1 t/\varepsilon})$. By simple calculations, we obtain

$$\begin{aligned} L_{\varepsilon\mu}(\phi \pm \tilde{u}_t) &= bC + C\varepsilon^{-2} e^{-\beta_1 t/\varepsilon}(b - \beta_1) \pm (\tilde{f}_t - b_t \tilde{u}) \\ &\geq 0, \quad \text{for sufficiently large } C, \end{aligned} \tag{40}$$

where we used the assumption (6), and the definitions of $\tilde{f}$ and $\tilde{u}$.

On the other hand, letting $t = 0$ in (38), we have

$$\tilde{u}_t(x, y, 0) = \varepsilon^{-1} \tilde{f}(x, y, 0, \varepsilon, \mu) \leq C\varepsilon^{-2}. \tag{41}$$

From the boundary conditions (39), we have

$$\tilde{u}_t|_{x=0,1} = \tilde{u}_t|_{y=0,1} = 0, \tag{42}$$

which along with (41) give us

$$(\phi \pm \tilde{u}_t)|_{\partial D} \geq 0,$$

which together with (40) and Lemma 2.1 gives us

$$|\tilde{u}_t| \leq C(1 + \varepsilon^{-2} e^{-\beta_1 t/\varepsilon}). \tag{43}$$

By the definition of $\tilde{u}$ and the fact (43), we have

$$|u_{tt} \leq |\tilde{u}_t + \tilde{g}_t| \leq |\tilde{u}_t| + |\tilde{g}_t| \leq C(1 + \varepsilon^{-2} e^{-\beta_1 t/\varepsilon}),$$

which concludes the proof for $k = 2$.

It is not difficult to see that the above procedures can be used repeatedly for higher order estimates of u_{t^k} for $k \geq 3$. □

Lemma 2.10. *Suppose the solution u of (2)-(3) is sufficiently smooth on $\overline{D}$. Then under the assumptions (4)-(6), we have*

$$(i) \qquad |u_{x^k t}(x, y)| \leq C\varepsilon^{-1}(1 + \mu^{-k} e^{-\beta_2 x/\mu} + \mu^{-k} e^{-\beta_2 (1-x)/\mu}) \quad on\ \overline{D},\ \forall\ k \geq 1,$$

$$(ii) \qquad |u_{y^k t}(x, y)| \leq C\varepsilon^{-1}(1 + \mu^{-k} e^{-\beta_2 y/\mu} + \mu^{-k} e^{-\beta_2 (1-y)/\mu}) \quad on\ \overline{D},\ \forall\ k \geq 1.$$

Proof. Note that (38)-(39) are in the same form as our original problem (2)-(3), except that the right hand side $\tilde{f}$ satisfies

$$|\tilde{f}_{x^i}(x,y,t,\varepsilon,\mu)| \le C\varepsilon^{-1}(1+\mu^{-i}e^{-\beta_2 x/\mu}+\mu^{-i}e^{-\beta_2(1-x)/\mu}) \text{ on } \overline{\Omega},\ \forall\, i \ge 0,$$
$$|\tilde{f}_{y^i}(x,y,t,\varepsilon,\mu)| \le C\varepsilon^{-1}(1+\mu^{-i}e^{-\beta_2 y/\mu}+\mu^{-i}e^{-\beta_2(1-y)/\mu}) \text{ on } \overline{\Omega},\ \forall\, i \ge 0,$$
$$|\tilde{f}_{t^i}(x,y,t,\varepsilon,\mu)| \le C\varepsilon^{-1}(1+\varepsilon^{-i}e^{-\beta_1 t/\varepsilon}) \text{ on } \overline{\Omega},\ \forall\, i \ge 0,$$

which differ from (4)-(6) only in a constant ε^{-1}.

Hence by carrying out the same procedures used in Lemmas 2.2-2.7, it is easy to obtain that

$$|\tilde{u}_{x^k}(x,y,t)| \le C\varepsilon^{-1}(1+\mu^{-k}e^{-\beta_2 x/\mu}+\mu^{-k}e^{-\beta_2(1-x)/\mu}) \quad \text{on } \overline{D},\ \forall\, k \ge 1,$$

from which and the definitions of $\tilde{u}$, we have

$$|u_{x^k t}| \le |\tilde{u}_{x^k}(x,y,t)+\tilde{g}_{x^k}| \le C\varepsilon^{-1}(1+\mu^{-k}e^{-\beta_2 x/\mu}+\mu^{-k}e^{-\beta_2(1-x)/\mu}),$$

where in the last step we used the definition of $\tilde{g}$ and the assumption (4).

By symmetry, (ii) can be proved directly. □

3. The mesh and the scheme

From Theorem 2.1 and Lemma 2.9, we see that the solution u of (2)-(3) has sharp boundary layers at faces $x=0,1, y=0,1$, and $t=0$. Hence we need finer mesh inside the boundary layers than elsewhere. First we divide Ω into nine subdomains Ω_i, $1\le i\le 9$, i.e., $\Omega=\cup_{i=1}^{9}\Omega_i$, where

$$\begin{aligned}
\Omega_1 &\equiv (0,\sigma_x)\times(0,\sigma_y),\ \Omega_2\equiv(\sigma_x,1-\sigma_x)\times(0,\sigma_y),\ \Omega_3\equiv(1-\sigma_x,1)\times(0,\sigma_y),\\
\Omega_4 &\equiv (0,\sigma_x)\times(\sigma_y,1-\sigma_y),\quad \Omega_5\equiv(\sigma_x,1-\sigma_x)\times(\sigma_y,1-\sigma_y),\\
\Omega_6 &\equiv (1-\sigma_x,1)\times(\sigma_y,1-\sigma_y),\quad \Omega_7\equiv(0,\sigma_x)\times(1-\sigma_y,1),\\
\Omega_8 &\equiv (\sigma_x,1-\sigma_x)\times(1-\sigma_y,1),\quad \Omega_9\equiv(1-\sigma_x,1)\times(1-\sigma_y,1).
\end{aligned}$$

Here $\sigma_x=\sigma_y=(k+1)\beta_2^{-1}\mu|\ln\mu|$ for the standard k-th order conforming tensor-product finite elements. Then each subdomain Ω_i is divided quasi-uniformly in both the x- and y-directions. We assume that the meshes are matching globally, hence we obtain an *a prior* anisotropically refined mesh (see Figure 1), which is refined only in the directions of the boundary layers. To simplify the notation, we assume equal total number of partitions, denoted as N, in both the x- and y-directions. The number of divisions in each subdomain is some fraction of N.

Similarly in the time space, we use smaller time step inside the boundary layer than elsewhere. We divide $[0,T]$ into two subdomains, i.e.,

$$[0,T]=[0,\sigma_t]\bigcup[\sigma_t,T],$$

where $\sigma_t=\beta_1^{-1}\varepsilon|\ln\varepsilon|$. Then each subdomain is partitioned uniformly into $M/2$ intervals. Note that to build our mesh, the conditions of $\sigma_x<1/2$ and $\sigma_t<T/2$ are implied to be true. Otherwise, the underlying problem is not considered to be singularly perturbed.

The weak formulation of (2)-(3) is given by finding $u:(0,T]\to H_0^1(\Omega)$ such that

$$\varepsilon(\frac{\partial u}{\partial t},v)+\mu^2(a\nabla u,\nabla v)+(bu,v)=(f,v),\quad v\in H_0^1(\Omega). \tag{44}$$

Here and throughout the paper, $(\cdot,\cdot)$ denotes the inner product in $L^2(\Omega)$.

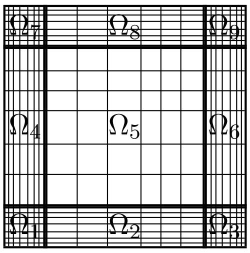

FIGURE 1. An exemplary anisotropic mesh

Consider the following fully discrete time Galerkin method for approximating the solution of (2)-(3): find $u_{h,k}^{n+1} \in S_h^k$ such that

$$\varepsilon\left(\frac{u_{h,k}^{n+1}-u_{h,k}^{n}}{\delta t_n}, v_h\right) + \mu^2(a\nabla u_{h,k}^{n+1}, \nabla v_h) + (bu_{h,k}^{n+1}, v_h) = (f, v_h), \quad v_h \in S_h^k, \tag{45}$$

where $S_h^k \subseteq H_0^1(\Omega)$ is the k-th order conforming tensor-product finite elements on the above special rectangular partition of Ω. Here and below we denote

$$\delta t_n = t_{n+1} - t_n, \quad n = 0, \cdots, M-1,$$

where $t_0 = 0, t_{M/2} = \sigma_t, t_M = T$. Note that (45) is the so-called backward scheme [35, pp.748]. More sophisticated discrete time schemes [10, 33] can be discussed similarly.

Remark. The scheme (45) is unconditionally stable, hence there is no restriction on the time step. This can be seen by taking $v_h = u_{h,k}^{n+1}$ in (45) and using the Hölder inequality, we obtain

$$\begin{aligned} &\frac{\varepsilon}{\delta t_n}||u_{h,k}^{n+1}||_{0,\Omega}^2 + \mu^2 C||\nabla u_{h,k}^{n+1}||_{0,\Omega}^2 + C||u_{h,k}^{n+1}||_{0,\Omega}^2 \\ \le\ & ||f||_{0,\Omega}||u_{h,k}^{n+1}||_{0,\Omega} + \frac{\varepsilon}{\delta t_n}||u_{h,k}^{n}||_{0,\Omega}||u_{h,k}^{n+1}||_{0,\Omega}, \end{aligned} \tag{46}$$

where $C = \min_{(x,y,t)\in\overline{\Omega}\times[0,T]}(a(x,y,t), b(x,y,t))$. Using

$$||f||_{0,\Omega}||u_{h,k}^{n+1}||_{0,\Omega} \le \frac{C}{2}||u_{h,k}^{n+1}||_{0,\Omega}^2 + \frac{1}{2C}||f||_{0,\Omega}^2,$$

and

$$||u_{h,k}^{n}||_{0,\Omega}||u_{h,k}^{n+1}||_{0,\Omega} \le \frac{1}{2}||u_{h,k}^{n}||_{0,\Omega}^2 + \frac{1}{2}||u_{h,k}^{n+1}||_{0,\Omega}^2,$$

we can rewrite (46) as

$$\begin{aligned} &\frac{\varepsilon}{2\delta t_n}||u_{h,k}^{n+1}||_{0,\Omega}^2 + \mu^2 C||\nabla u_{h,k}^{n+1}||_{0,\Omega}^2 + \frac{C}{2}||u_{h,k}^{n+1}||_{0,\Omega}^2 \\ \le\ & \frac{1}{2C}||f||_{0,\Omega}^2 + \frac{\varepsilon}{2\delta t_n}||u_{h,k}^{n}||_{0,\Omega}^2, \end{aligned} \tag{47}$$

from which we have

$$||u_{h,k}^{n+1}||_{0,\Omega}^2 \le \frac{\delta t_n}{\varepsilon C}||f||_{0,\Omega}^2 + ||u_{h,k}^{n}||_{0,\Omega}^2. \tag{48}$$

Summing up (48) from $n = 0$ to $M-1$ and using the fact $\sum_{n=0}^{M-1}\delta t_n = T$, we obtain

$$||u_{h,k}^{M}||_{0,\Omega}^2 \le \frac{T}{\varepsilon C}||f||_{0,\Omega}^2 + ||u_{h,k}^{0}||_{0,\Omega}^2,$$

which shows that the scheme is unconditionally stable.

4. The special interpolation operator and its interpolation estimates

Consider a special interpolation operator $\Pi_h^k w : w \in C^0(\bar{\Omega}) \to Q_k$ defined on each rectangular element τ of Ω by the following conditions [13, pp.108]:

$$\Pi_h^k w(a_i) = w(a_i),\ i = 1,2,3,4, \tag{49}$$
$$\int_{l_j} (\Pi_h^k w - w)v = 0,\ \forall v|_{l_j} \in P_{k-2}(l_j),\ j = 1,2,3,4,$$
$$\int_\tau (\Pi_h^k w - w)v = 0,\ \forall v|_\tau \in Q_{k-2}(\tau).$$

for $k \geq 2$, where a_i and l_j denote the vertices and edges of τ, which are illustrated in Figure 1 of [18]. Here P_k is the k-th order polynomial in one dimension, and

$$Q_k(\tau) = \text{span}\{x^i y^j : 0 \leq i, j \leq k, (x,y) \in \tau\}.$$

Note that when $k = 1$, Π_h^k is defined by (49) only, i.e., Π_h^1 is the standard bilinear interpolation.

By [13, pp.108], Π_h^k is well defined. Furthermore, using standard scaling technique [8, 5], we can easily prove the following interpolation estimates [20, Lemma 3.1]:

Lemma 4.1. *Let integer $k \geq 1$, and real p with $1 \leq p \leq \infty$. Then for all $v \in W^{k+1,p}(\tau)$, we have*

$$||v - \Pi_h^k v||_{0,p,\tau} \leq C(h_{x,\tau}^{k+1} ||v_{x^{k+1}}||_{0,p,\tau} + h_{y,\tau}^{k+1} ||v_{y^{k+1}}||_{0,p,\tau}),$$

where τ is an arbitrary rectangular element with width $h_{x,\tau}$ and length $h_{y,\tau}$.

Using the integral identity technique [23], we have [18, Lemmas 1-2]:

Lemma 4.2. *For $\forall v \in Q_k(\tau), k \geq 1$ and $\forall w \in H^{k+2}(\tau)$, we have*

$$
\begin{aligned}
(i) \quad \int_\tau (\Pi_h^k w - w)_x v_x &= \frac{(-2)^k}{(2k)!} \int_\tau F^k(y) w_{xy^{k+1}}(x,y) v_{xy^{k-1}}(x, y_\tau) \\
&\quad + \frac{(-2)^k}{(2k+2)!} \int_\tau (F^{k+1}(y))_y w_{xy^{k+1}}(x,y) v_{xy^k}(x, y_\tau) \\
&= O(h_{y,\tau}^{k+1}) ||w_{xy^{k+1}}||_{0,\tau} ||v_x||_{0,\tau} \\
(ii) \quad \int_\tau (\Pi_h^k w - w)_y v_y &= \frac{(-2)^k}{(2k)!} \int_\tau E^k(x) w_{yx^{k+1}}(x,y) v_{yx^{k-1}}(x_\tau, y) \\
&\quad + \frac{(-2)^k}{(2k+2)!} \int_\tau (E^{k+1}(x))_x w_{x^{k+1}y}(x,y) v_{x^k y}(x_\tau, y) \\
&= O(h_{x,\tau}^{k+1}) ||w_{x^{k+1}y}||_{0,\tau} ||v_y||_{0,\tau}
\end{aligned}
$$

Here $\tau = [x_c - h_{x,\tau}, x_c + h_{x,\tau}] \times [y_c - h_{y,\tau}, y_c + h_{y,\tau}]$ is a rectangular element centered at (x_c, y_c), with width $2h_{x,\tau}$ and length $2h_{y,\tau}$, and

$$E(x) = \frac{1}{2}[(x - x_c)^2 - h_{x,\tau}^2], \quad F(y) = \frac{1}{2}[(y - y_c)^2 - h_{y,\tau}^2].$$

Also we denote $F^k(y) = (F(y))^k$, and $E^k(x) = (E(x))^k$.

In the following, we need to obtain some mixed order derivatives (which are impossible to get by the maximum principle directly) by the technique we developed in [22, 19].

For clarity, we introduce the following short notation:

$$\Omega_1^x = \Omega_1 \cup \Omega_2 \cup \Omega_3,\ \Omega_2^x = \Omega_4 \cup \Omega_5 \cup \Omega_6,\ \Omega_3^x = \Omega_7 \cup \Omega_8 \cup \Omega_9,$$
$$\Omega_1^y = \Omega_1 \cup \Omega_4 \cup \Omega_7,\ \Omega_2^y = \Omega_2 \cup \Omega_5 \cup \Omega_8,\ \Omega_3^y = \Omega_3 \cup \Omega_6 \cup \Omega_9.$$

Lemma 4.3. *For the solution u of (2)-(3) and any $k \geq 1$, we have*

$$\begin{array}{ll} (i) & \mu||u_{xy^k}||_{0,\Omega_1^x} \leq C\mu^{-k} \cdot (\mu|\ln\mu|)^{1/2}, \\ & \mu||u_{xy^k}||_{0,\Omega_2^x} \leq C, \\ & \mu||u_{xy^k}||_{0,\Omega_3^x} \leq C\mu^{-k} \cdot (\mu|\ln\mu|)^{1/2}, \\ (ii) & \mu||u_{x^ky}||_{0,\Omega_1^y} \leq C\mu^{-k} \cdot (\mu|\ln\mu|)^{1/2}, \\ & \mu||u_{x^ky}||_{0,\Omega_2^y} \leq C, \\ & \mu||u_{x^ky}||_{0,\Omega_3^y} \leq C\mu^{-k} \cdot (\mu|\ln\mu|)^{1/2}. \end{array}$$

Proof. (i) Differentiating (2) k times with respect to y gives us

$$-\mu^2 a u_{x^2y^k} = f_{y^k} - (bu)_{y^k} + \mu^2 a u_{y^{k+2}} - \varepsilon u_{y^k t}. \tag{50}$$

Integrating by parts and using the boundary condition $u|_{\partial\Omega\times(0,T]} = 0$ (hence $u_{y^k}|_{x=0,1} = 0$), Theorem 2.1, Lemma 2.10, and (50), we have

$$\begin{aligned} \mu^2 \int_{\Omega_1^x} a u_{xy^k} \cdot u_{xy^k} dxdy &= \mu^2 \int_0^{\sigma_y} (u_{y^k} \cdot a u_{xy^k})|_{x=0}^1 dy - \mu^2 \int_{\Omega_1^x} u_{y^k} \cdot a u_{x^2y^k} dxdy \\ &= -\int_{\Omega_1^x} u_{y^k} \cdot \mu^2 a u_{x^2y^k} dxdy \\ &= \int_{\Omega_1^x} u_{y^k} \cdot (f_{y^k} - (bu)_{y^k} + \mu^2 a u_{y^{k+2}} - \varepsilon u_{y^k t}) \\ &\leq C\mu^{-k} \cdot \mu^{-k} \cdot \text{meas}(\Omega_1^x), \end{aligned}$$

from which and the fact that $meas(\Omega_1^x) = O(\mu|\ln\mu|)$, we obtain

$$\mu||u_{xy^k}||_{0,\Omega_1^x} \leq C\mu^{-k} \cdot (\mu|\ln\mu|)^{1/2}.$$

Similarly, we have

$$\mu^2 \int_{\Omega_2^x} a u_{xy^k} \cdot u_{xy^k} dxdy = \int_{\Omega_2^x} u_{y^k} \cdot (f_{y^k} - (bu)_{y^k} + \mu^2 a u_{y^{k+2}} - \varepsilon u_{y^k t}),$$

which along with the fact that

$$\mu^{-k} e^{-\beta_2 y/\mu} + \mu^{-k} e^{-\beta_2(1-y)/\mu} \leq 2\mu^{-k} e^{-\beta_2\sigma_y/\mu} \leq 2, \quad \text{for } \sigma_y \leq y \leq 1 - \sigma_y,$$

gives us

$$\mu||u_{xy^k}||_{0,\Omega_2^x} \leq C.$$

In the same way, it is easy to see that

$$\mu||u_{xy^k}||_{0,\Omega_3^x} \leq C\mu^{-k} \cdot (\mu|\ln\mu|)^{1/2}.$$

(ii) Differentiating (2) k times with respect to x gives us

$$-\mu^2 a u_{x^ky^2} = f_{x^k} - (bu)_{x^k} + \mu^2 a u_{x^{k+2}} - \varepsilon u_{x^k t}. \tag{51}$$

Integrating by parts and using the boundary condition $u|_{\partial\Omega\times(0,T]} = 0$ (hence $u_{x^k}|_{y=0,1} = 0$), Theorem 2.3, Lemma 2.10, and (51), we have

$$\begin{aligned}\mu^2 \int_{\Omega_1^y} au_{x^k y} \cdot u_{x^k y} dxdy &= \mu^2 \int_0^{\sigma_x} (u_{x^k} \cdot au_{x^k y})|_{y=0}^1 dx - \mu^2 \int_{\Omega_1^y} u_{x^k} \cdot au_{x^k y^2} dxdy \\ &= \int_{\Omega_1^y} u_{x^k} \cdot (f_{x^k} - (bu)_{x^k} + \mu^2 au_{x^{k+2}} - \varepsilon u_{x^k t}) \\ &\leq C\mu^{-k} \cdot \mu^{-k} \cdot \mathrm{meas}(\Omega_1^y),\end{aligned}$$

from which we obtain

$$\mu||u_{x^k y}||_{0,\Omega_1^y} \leq C\mu^{-k} \cdot (\mu|\ln\mu|)^{1/2}.$$

The other inequalities can be proved easily by the same technique. ∎

Lemma 4.4. *Let u be the solution of (2)-(3), and $\Pi_h^k u$ be the special interpolant of u defined in section 4. Then for any $k \geq 1$, we have*

(i) $|\mu^2((\Pi_h^k u - u)_x, \chi_x)| \leq CC_\mu N^{-(k+1)} \cdot \mu||\chi_x||_{0,\Omega}, \quad \forall\ \chi \in S_h^k,$

(ii) $|\mu^2((\Pi_h^k u - u)_y, \chi_y)| \leq CC_\mu N^{-(k+1)} \cdot \mu||\chi_y||_{0,\Omega}, \quad \forall\ \chi \in S_h^k,$

(iii) $||\Pi_h^k u - u||_{0,\Omega} \leq CC_\mu N^{-(k+1)},$

(iv) $\varepsilon||\Pi_h^k u_t - u_t||_{0,\Omega} \leq CC_\mu N^{-(k+1)},$

where $C_\mu = 1 + \mu^{1/2}|\ln^{k+3/2}\mu|$.

Proof. (i) Let $h_{x,\tau}$ and $h_{y,\tau}$ be the width and length of element τ, and

$$h_{x,\Omega_i^y} = \max_{\tau\in\Omega_i^y} h_{x,\tau}, \quad h_{y,\Omega_i^x} = \max_{\tau\in\Omega_i^x} h_{y,\tau}.$$

Denote

$$T_i = \mu^2 \int_{\Omega_i^x} (\Pi_h^k u - u)_x \chi_x, \quad i = 1, 2, 3.$$

By the construction of our mesh, we have

$$h_{x,\Omega_1^y} \approx h_{x,\Omega_3^y} \approx h_{y,\Omega_1^x} \approx h_{y,\Omega_3^x} = O(\frac{\mu|\ln\mu|}{N}), \tag{52}$$

$$h_{x,\Omega_2^y} \approx h_{y,\Omega_2^x} = O(\frac{1}{N}), \tag{53}$$

By using Lemmas 4.2 and 4.3, and (52)-(53), we have

$$\begin{aligned}|T_1| &\leq Ch_{y,\Omega_1^x}^{k+1} \cdot \mu||u_{xy^{k+1}}||_{0,\Omega_1^x} \cdot \mu||\chi_x||_{0,\Omega_1^x} \\ &\leq C(\frac{\mu|\ln\mu|}{N})^{k+1} \cdot \mu^{-(k+1)} \cdot (\mu|\ln\mu|)^{1/2} \cdot \mu||\chi_x||_{0,\Omega_1^x} \\ &= CN^{-(k+1)} \cdot \mu^{1/2}|\ln^{k+3/2}\mu| \cdot \mu||\chi_x||_{0,\Omega_1^x}.\end{aligned} \tag{54}$$

By symmetry, we have

$$|T_3| \leq CN^{-(k+1)} \cdot \mu^{1/2}|\ln^{k+3/2}\mu| \cdot \mu||\chi_x||_{0,\Omega_3^x}. \tag{55}$$

Similarly, we obtain

$$\begin{aligned}|T_2| &\leq Ch_{y,\Omega_1^x}^{k+1} \cdot \mu||u_{xy^{k+1}}||_{0,\Omega_2^x} \cdot \mu||\chi_x||_{0,\Omega_2^x} \\ &\leq C(\frac{1}{N})^{k+1} \cdot C \cdot \mu||\chi_x||_{0,\Omega_2^x} \\ &= CN^{-(k+1)} \cdot \mu||\chi_x||_{0,\Omega_2^x},\end{aligned} \tag{56}$$

which along with (54) and (55) gives

$$
\begin{aligned}
|\mu^2((\Pi_h^k u-u)_x,\chi_x)| &= |T_1+T_2+T_3| \\
&\le CN^{-(k+1)}(\mu^{1/2}|\ln^{k+3/2}\mu|+1)\cdot\mu||\chi_x||_{0,\Omega}, \qquad (57)
\end{aligned}
$$

(ii) By the same technique, it is not difficult to see that

$$
\begin{aligned}
|\mu^2((\Pi_h^k u-u)_y,\chi_y)| &= |\sum_{i=1}^{3}\mu^2\int_{\Omega_i^y}(\Pi_h^k u-u)_y\chi_y| \\
&\le C\sum_{i=1}^{3}h_{x,\Omega_i^y}^{k+1}\cdot\mu||u_{x^{k+1}y}||_{0,\Omega_i^y}\cdot\mu||\chi_y||_{0,\Omega_i^y} \\
&\le CN^{-(k+1)}(\mu^{1/2}|\ln^{k+3/2}\mu|+1)\cdot\mu||\chi_y||_{0,\Omega}, \qquad (58)
\end{aligned}
$$

(iii) By using Theorem 2.1 and the construction of our mesh, we observe that

$$ ||u_{x^{k+1}}||_{\infty,\Omega_2^y}\le C,\ ||u_{x^{k+1}}||_{\infty,\Omega_i^y}\le C\mu^{-(k+1)},\ i=1,3, \qquad (59) $$

$$ ||u_{y^{k+1}}||_{\infty,\Omega_2^x}\le C,\ ||u_{y^{k+1}}||_{\infty,\Omega_i^x}\le C\mu^{-(k+1)},\ i=1,3, \qquad (60) $$

then by Lemma 4.1, (52)-(53), and (59)-(60), we have

$$
\begin{aligned}
||\Pi_h^k u-u||_{0,\Omega} &\le \sum_{i=1}^{9}||\Pi_h^k u-u||_{\infty,\Omega_i}\cdot\mathrm{meas}^{1/2}(\Omega_i) \\
&\le C\sum_{i=1}^{3}h_{x,\Omega_i^y}^{k+1}||u_{x^{k+1}}||_{\infty,\Omega_i^y}\cdot\mathrm{meas}^{1/2}(\Omega_i^y) \\
&\quad +C\sum_{i=1}^{3}h_{y,\Omega_i^x}^{k+1}||u_{y^{k+1}}||_{\infty,\Omega_i^x}\cdot\mathrm{meas}^{1/2}(\Omega_i^x) \\
&\le C\cdot(\frac{\mu|\ln\mu|}{N})^{k+1}\cdot\mu^{-(k+1)}\cdot(\mu|\ln\mu|)^{1/2}+CN^{-(k+1)} \\
&= CN^{-(k+1)}\cdot(\mu^{1/2}|\ln^{k+3/2}\mu|+1),
\end{aligned}
$$

which completes the proof of (iii).

(iv) By using Lemma 2.10 and the construction of our mesh in section 3, we observe that

$$ ||u_{x^{k+1}t}||_{\infty,\Omega_2^y}\le C\varepsilon^{-1}(1+2\mu^{-(k+1)}e^{-\beta_2\sigma_x/\mu})=3C\varepsilon^{-1}, \qquad (61) $$

$$ ||u_{y^{k+1}t}||_{\infty,\Omega_2^x}\le C\varepsilon^{-1}(1+2\mu^{-(k+1)}e^{-\beta_2\sigma_y/\mu})=3C\varepsilon^{-1}, \qquad (62) $$

$$ ||u_{x^{k+1}t}||_{\infty,\Omega_i^y}\le C\varepsilon^{-1}\mu^{-(k+1)},\quad ||u_{y^{k+1}t}||_{\infty,\Omega_i^x}\le C\varepsilon^{-1}\mu^{-(k+1)},\ i=1,3, \qquad (63) $$

then by Lemma 4.1, (52)-(53), and (61)-(63), we have

$$\begin{aligned}
\varepsilon||\Pi_h^k u_t - u_t||_{0,\Omega} &\le \varepsilon \sum_{i=1}^{9} ||\Pi_h^k u_t - u_t||_{\infty,\Omega_i} \cdot \text{meas}^{1/2}(\Omega_i) \\
&\le C\varepsilon \sum_{i=1}^{3} h_{x,\Omega_i^y}^{k+1} ||u_{x^{k+1}t}||_{\infty,\Omega_i^y} \cdot \text{meas}^{1/2}(\Omega_i^y) \\
&\quad + C\varepsilon \sum_{i=1}^{3} h_{y,\Omega_i^x}^{k+1} ||u_{y^{k+1}t}||_{\infty,\Omega_i^x} \cdot \text{meas}^{1/2}(\Omega_i^x) \\
&\le C\varepsilon \cdot (\frac{\mu|\ln\mu|}{N})^{k+1} \cdot \varepsilon^{-1}\mu^{-(k+1)} \cdot (\mu|\ln\mu|)^{1/2} \\
&\quad + C\varepsilon \cdot N^{-(k+1)} \cdot \varepsilon^{-1} \\
&= CN^{-(k+1)} \cdot (\mu^{1/2}|\ln^{k+3/2}\mu| + 1),
\end{aligned}$$

which completes the proof of (iv). ∎

5. The error estimate

By letting $v = v_h$ and $t = t_{n+1}$ in (44), and subtracting (45) from (44), we obtain the error equation

$$\begin{aligned}
\varepsilon(\frac{\partial u^{n+1}}{\partial t} - \frac{u_{h,k}^{n+1} - u_{h,k}^{n}}{\delta t_n}, v_h) &+ \mu^2(a(\nabla u^{n+1} - \nabla u_{h,k}^{n+1}), \nabla v_h) \\
&+ (b(u^{n+1} - u_{h,k}^{n+1}), v_h) = 0, \quad v_h \in S_h^k.
\end{aligned} \tag{64}$$

Theorem 5.1. *Let u and $u_{h,k}^{n+1}$ be the solutions of (2)-(3) and (45), respectively. Then we have*

$$(\sum_{n=0}^{M-1} \delta t_n ||u^{n+1} - u_{h,k}^{n+1}||_{0,\Omega}^2)^{1/2} \le C(N^{-(k+1)} + M^{-1}),$$

where C is independent of the small parameters ε and μ.

We remark that this estimate is optimal in the discrete L^2-norm [10, pp.151] with respect to the order of N and M. Estimates for more sophisticated discrete time schemes [35, 10, 33] can be pursued accordingly.

Proof. By the Taylor expansion, we have

$$\frac{\partial u^{n+1}}{\partial t} = \frac{u^{n+1} - u^n}{\delta t_n} + \frac{1}{2}\delta t_n \hat{u}_{t^2}^{n+1}, \tag{65}$$

where $\hat{u}_{t^2}^{n+1} = \frac{\partial^2 u}{\partial t^2}(x, y, \hat{t})$, for some $\hat{t}$ between t_n and t_{n+1}.

Substituting (65) into (64) gives us

$$\begin{aligned}
&\varepsilon(\frac{(u^{n+1} - u_{h,k}^{n+1}) - (u^n - u_{h,k}^n)}{\delta t_n}, v_h) + \mu^2(a(\nabla u^{n+1} - \nabla u_{h,k}^{n+1}), \nabla v_h) \\
&+ (b(u^{n+1} - u_{h,k}^{n+1}), v_h) = -\varepsilon\frac{1}{2}\delta t_n(\hat{u}_{t^2}^{n+1}, v_h), \quad v_h \in S_h^k.
\end{aligned} \tag{66}$$

Denote

$$\chi_{h,k}^{n+1} = \Pi_h^k u^{n+1} - u_{h,k}^{n+1}, \quad \eta_k^{n+1} = \Pi_h^k u^{n+1} - u^{n+1}.$$

By letting $v_h = \chi_{h,k}^{n+1}$ in (66) and reorganizing it, we have

$$\begin{aligned}
&\frac{\varepsilon}{\delta t_n}(\chi_{h,k}^{n+1}, \chi_{h,k}^{n+1}) + \mu^2(a\nabla\chi_{h,k}^{n+1}, \nabla\chi_{h,k}^{n+1}) + (b\chi_{h,k}^{n+1}, \chi_{h,k}^{n+1}) \\
= \;& \frac{\varepsilon}{\delta t_n}(\chi_{h,k}^{n}, \chi_{h,k}^{n+1}) + \frac{\varepsilon}{\delta t_n}(\eta_k^{n+1} - \eta_k^n, \chi_{h,k}^{n+1}) \\
&+\mu^2(a\nabla\eta_k^{n+1}, \nabla\chi_{h,k}^{n+1}) + (b\eta_k^{n+1}, \chi_{h,k}^{n+1}) - \varepsilon\cdot\frac{1}{2}\delta t_n(\hat{u}_{t^2}^{n+1}, \chi_{h,k}^{n+1}),
\end{aligned}$$

from which we obtain

$$\varepsilon||\chi_{h,k}^{n+1}||_{0,\Omega}^2 + \delta t_n\cdot\mu^2||\nabla\chi_{h,k}^{n+1}||_{0,\Omega}^2 + \delta t_n||\chi_{h,k}^{n+1}||_{0,\Omega}^2 \le \sum_{i=1}^{5} E_i, \tag{67}$$

where we used the properties of a and b, and the notations

$$\begin{aligned}
E_1 &= C|\varepsilon(\chi_{h,k}^n, \chi_{h,k}^{n+1})|, \\
E_2 &= C|\varepsilon(\eta_k^{n+1} - \eta_k^n, \chi_{h,k}^{n+1})|, \\
E_3 &= C|\delta t_n\cdot\mu^2(a\nabla\eta_k^{n+1}, \nabla\chi_{h,k}^{n+1})|, \\
E_4 &= C|\delta t_n(b\eta_k^{n+1}, \chi_{h,k}^{n+1})|, \\
E_5 &= C|\varepsilon\cdot(\delta t_n)^2(\hat{u}_{t^2}^{n+1}, \chi_{h,k}^{n+1})|.
\end{aligned}$$

By the Cauchy-Schwarz inequality, we have

$$E_1 \le \frac{\varepsilon C^2}{2}||\chi_{h,k}^n||_{0,\Omega}^2 + \frac{\varepsilon}{2}||\chi_{h,k}^{n+1}||_{0,\Omega}^2. \tag{68}$$

By the Taylor expansion, we have

$$\begin{aligned}
\eta_k^{n+1} - \eta_k^n &= \delta t_n\cdot\frac{\partial\eta_k}{\partial t}(x,y,\tilde{t}), \quad \text{for some } \tilde{t} \text{ between } t_n \text{ and } t_{n+1} \\
&= \delta t_n\cdot(\Pi_h^k u_t - u_t)(x,y,\tilde{t}),
\end{aligned}$$

which along with Lemma 4.4 (iv) gives us

$$\begin{aligned}
E_2 &\le C\delta t_n\cdot\varepsilon||(\Pi_h^k u_t - u_t)(x,y,\tilde{t})||_{0,\Omega}||\chi_{h,k}^{n+1}||_{0,\Omega} \\
&\le \delta t_n\cdot CC_\mu N^{-(k+1)}\cdot||\chi_{h,k}^{n+1}||_{0,\Omega} \\
&\le \delta t_n[\frac{1}{4}||\chi_{h,k}^{n+1}||_{0,\Omega}^2 + C^2C_\mu^2N^{-2(k+1)}]. \qquad (69)
\end{aligned}$$

By Lemma 4.4 (i)-(iii), we have

$$\begin{aligned}
E_3 &\le \delta t_n\cdot CC_\mu N^{-(k+1)}\cdot\mu||\nabla\chi_{h,k}^{n+1}||_{0,\Omega} \\
&\le \delta t_n[\frac{1}{2}\mu^2||\nabla\chi_{h,k}^{n+1}||_{0,\Omega}^2 + \frac{1}{2}C^2C_\mu^2N^{-2(k+1)}]. \qquad (70)
\end{aligned}$$

By Lemma 4.4 (iii), we have

$$\begin{aligned}
E_4 &\le \delta t_n\cdot CC_\mu N^{-(k+1)}\cdot||\chi_{h,k}^{n+1}||_{0,\Omega} \\
&\le \delta t_n[\frac{1}{4}||\chi_{h,k}^{n+1}||_{0,\Omega}^2 + C^2C_\mu^2N^{-2(k+1)}). \qquad (71)
\end{aligned}$$

Also by Lemma 2.9, we have

$$\begin{aligned}
E_5 &\le C\varepsilon\cdot(\delta t_n)^2||\hat{u}_{t^2}^{n+1}||_{\infty,\Omega}||\chi_{h,k}^{n+1}||_{0,\Omega} \\
&\le C\cdot\delta t_n\cdot\varepsilon\delta t_n(1+\varepsilon^{-2}e^{-\beta_1\hat{t}/\varepsilon})||\chi_{h,k}^{n+1}||_{0,\Omega} \\
&\le \delta t_n[\frac{1}{4}||\chi_{h,k}^{n+1}||_{0,\Omega}^2 + C^2\cdot(\varepsilon\delta t_n)^2\cdot(1+\varepsilon^{-2}e^{-\beta_1\hat{t}/\varepsilon})^2]. \qquad (72)
\end{aligned}$$

Combining (67) with (68)-(72), we obtain

$$\begin{aligned}
&\frac{\varepsilon}{2}(||\chi_{h,k}^{n+1}||_{0,\Omega}^2 - ||\chi_{h,k}^{n}||_{0,\Omega}^2) + \frac{1}{2}\delta t_n \cdot \mu^2 ||\nabla \chi_{h,k}^{n+1}||_{0,\Omega}^2 + \frac{1}{4}\delta t_n ||\chi_{h,k}^{n+1}||_{0,\Omega}^2 \\
\leq\ & \delta t_n [\frac{5}{2} C^2 C_\mu^2 N^{-2(k+1)} + C^2 \cdot (\varepsilon \delta t_n)^2 \cdot (1 + \varepsilon^{-2} e^{-\beta_1 \hat{t}/\varepsilon})^2]. \qquad (73)
\end{aligned}$$

Summing up the above inequality from $n = 0$ to $M - 1$, we have

$$\begin{aligned}
&\frac{\varepsilon}{2}(||\chi_{h,k}^{M}||_{0,\Omega}^2 - ||\chi_{h,k}^{0}||_{0,\Omega}^2) \\
&+\frac{1}{2}\sum_{n=0}^{M-1} \delta t_n \cdot \mu^2 ||\nabla \chi_{h,k}^{n+1}||_{0,\Omega}^2 + \frac{1}{4}\sum_{n=0}^{M-1} \delta t_n ||\chi_{h,k}^{n+1}||_{0,\Omega}^2 \\
\leq\ & \frac{5}{2} C^2 C_\mu^2 N^{-2(k+1)} T + E_6, \qquad (74)
\end{aligned}$$

where we used the fact that $\sum_{n=0}^{M-1} \delta t_n = T$, and

$$E_6 = \sum_{n=0}^{M-1} \delta t_n [C^2 \cdot (\varepsilon \delta t_n)^2 \cdot (1 + \varepsilon^{-2} e^{-\beta_1 \hat{t}/\varepsilon})^2].$$

Using the facts that

$$\begin{aligned}
\delta t_n &= \frac{\varepsilon \beta_1^{-1} |\ln \varepsilon|}{M/2}, \quad \text{for } n = 0, \cdots, \frac{M}{2} - 1, \\
\delta t_n &= \frac{T - \varepsilon \beta_1^{-1} |\ln \varepsilon|}{M/2} \leq \frac{T}{M/2}, \quad \text{for } n = \frac{M}{2}, \cdots, M - 1,
\end{aligned}$$

and

$$\begin{aligned}
\varepsilon^{-1} e^{-\beta_1 \hat{t}/\varepsilon} &\leq \varepsilon^{-1}, \quad \text{for } n = 0, \cdots, \frac{M}{2} - 1, \\
\varepsilon^{-1} e^{-\beta_1 \hat{t}/\varepsilon} &\leq \varepsilon^{-1} e^{-\beta_1 \sigma_t/\varepsilon} = 1, \quad \text{for } n = \frac{M}{2}, \cdots, M - 1.
\end{aligned}$$

Hence

$$\begin{aligned}
E_6 &\leq \varepsilon^{-2} \sum_{n=0}^{M/2-1} (\delta t_n)^3 + \sum_{n=M/2}^{M-1} (\delta t_n)^3 \\
&\leq \varepsilon^{-2} \sum_{n=0}^{M/2-1} (\frac{\varepsilon \beta_1^{-1} |\ln \varepsilon|}{M/2})^3 + \sum_{n=M/2}^{M-1} (\frac{T}{M/2})^3 \\
&\leq CM^{-2}(1 + \varepsilon |\ln^3 \varepsilon|),
\end{aligned}$$

which along with (74) and the fact that $\chi_{h,k}^0 = 0$ gives us

$$\begin{aligned}
&\frac{\varepsilon}{2}||\chi_{h,k}^{M}||_{0,\Omega}^2 + \frac{1}{2}\sum_{n=0}^{M-1} \delta t_n \cdot \mu^2 ||\nabla \chi_{h,k}^{n+1}||_{0,\Omega}^2 + \frac{1}{4}\sum_{n=0}^{M-1} \delta t_n ||\chi_{h,k}^{n+1}||_{0,\Omega}^2 \\
\leq\ & \frac{5}{2} C^2 C_\mu^2 N^{-2(k+1)} T + CM^{-2}(1 + \varepsilon |\ln^3 \varepsilon|),
\end{aligned}$$

from which we obtain

$$(\sum_{n=0}^{M-1} \delta t_n ||\chi_{h,k}^{n+1}||_{0,\Omega}^2)^{1/2} \leq CC_\mu N^{-(k+1)} T^{1/2} + CM^{-1}(1 + \varepsilon^{1/2} |\ln^{3/2} \varepsilon|). \qquad (75)$$

By Lemma 4.4 (iii), we have

$$\sum_{n=0}^{M-1} \delta t_n ||\eta_{h,k}^{n+1}||_{0,\Omega}^2 \leq \sum_{n=0}^{M-1} \delta t_n \cdot (CC_\mu N^{-(k+1)})^2 = C^2 C_\mu^2 N^{-2(k+1)} T,$$

which along with (75) and the triangle inequality gives us

$$(\sum_{n=0}^{M-1} \delta t_n ||u^{n+1} - u_{h,k}^{n+1}||_{0,\Omega}^2)^{1/2} \leq CC_\mu N^{-(k+1)} T^{1/2} + CM^{-1}(1 + \varepsilon^{1/2} |\ln^{3/2} \varepsilon|). \tag{76}$$

It is easy to see that $\mu^{1/2} | \ln^{k+3/2} \mu|$ is actually uniformly (independent of μ) bounded for all $\mu \in (0, 1]$, for example

$$\mu^{1/2} | \ln^{3/2} \mu| < 1.5, \quad \mu^{1/2} | \ln^{5/2} \mu| < 4.6, \quad \mu^{1/2} | \ln^{7/2} \mu| < 28, \quad \text{for all } \mu \in (0, 1],$$

from which we see that (76) can be bounded independent of ε and μ. That completes the proof. □

6. Conclusions

Optimal uniform convergence is proved for a two-dimensional parabolic equation with two small parameters. First we used the maximum principle with carefully chosen barrier functions to obtain the pointwise arbitrary order derivative estimates, from which an anisotropic mesh is constructed. The mesh is much finer inside the boundary layer regions than elsewhere. Note that the standard finite element analysis technique can not be used directly for such highly nonuniform anisotropic meshes because of the violation of the quasi-uniformity assumption [8]. The optimal uniform convergence is obtained by using the integral identity technique [23]. Generalization of our results to other singular perturbation problems with two or more small parameters can be pursued similarly.

Acknowledgments

The author thanks the anonymous referee for helpful comments on the improvement of this paper.

References

[1] M. Ainsworth and I. Babuska, Reliable and robust a posteriori error estimation for singularly perturbed reaction-diffusion problems, SIAM J. Numer. Anal., 36 (1999), 331-353.

[2] T. Apel, Anisotropic Finite Elements: Local Estimates and Applications, B.G. Teubner Stuttgart, Leipzig, 1999.

[3] T. Apel and G. Lube, Anisotropic mesh refinement for a singularly perturbed reaction diffusion model problem, Appl. Numer. Math., 26 (1998), 415-433.

[4] U. M. Ascher, R. M. M. Mattheij and R. D. Russell, Numerical Solution of Boundary Value Problems for Ordinary Differential Equations, SIAM, Philadelphia, PA, 1995.

[5] S. C. Brenner and L. R. Scott, The Mathematical Theory of Finite Element Methods, Springer-Verlag, New York, 1994.

[6] V. F. Butuzov and A. B. Vasil'eva, Singularly perturbed differential equations of parabolic type, in: Lecture Notes in Mathematics, No.985 (1983), 35–75.

[7] C. M. Chen and Y. Q. Huang, High Accuracy Theory of Finite Element Methods, Hunan Science Press, China, 1995. (in Chinese)

[8] P. G. Ciarlet, The Finite Element Methods for Elliptic Problems, North-Holland, Amsterdam, 1978.

[9] R. E. Ewing, Y. Lin, T. Sun, J. Wang and S. Zhang, L^2-error estimates and superconvergence of mixed finite element methods for non-Fickian flows in porous media, SIAM J. Numer. Anal. 40 (2002), 1538-1560.

[10] G. Fairweather, Finite Element Galerkin Methods for Differential Equations, Marcel Dekker, New York, 1978.

[11] P. A. Farrell, A. F. Hegarty, J. J. H. Miller, E. O'Riordan and G. I. Shishkin, Robust Computational Techniques for Boundary Layers, CRC Press, Boca Raton, Florida, 2000.

[12] M. Garbey and H. G. Kaper, Heterogeneous domain decomposition for singularly perturbed elliptic boundary value problems, SIAM J. Numer. Anal., 34 (1997), 1513-1544.

[13] V. Girault and P. A. Raviart, Finite Element Methods for Navier-Stokes Equations: Theory and Algorithms, Springer-Verlag, Berlin, 1986.

[14] W. Hackbusch, Multigrid Methods and Applications, Springer-Verlag, Berlin, 1985.

[15] H. Han and R.B. Kellogg, Differentiability properties of solutions of the equation $-\varepsilon^2\triangle u + ru = f(x, y)$ in a square, SIAM J. Math. Anal., 21 (1990), 394-408.

[16] M. H. Holmes, Introduction to Perturbation Methods, Springer-Verlag, New York, 1995.

[17] G. Kunert and R. Verfurth, Edge residuals dominate a posteriori error estimates for linear finite element methods on anisotropic triangular and tetrahedral meshes, Numer. Math., 86 (2000), 283-303.

[18] J. Li, Full-order convergence of a mixed finite element method for fourth-order elliptic equations, J. Math. Anal. Appl., 230 (1999), 329-349.

[19] J. Li, Convergence and superconvergence analysis of finite element methods on highly nonuniform anisotropic meshes for singularly perturbed reaction-diffusion problems, Appl. Numer. Math., 36 (2001), 129-154.

[20] J. Li, Optimal uniform convergence analysis for a singularly perturbed quasilinear reaction-diffusion problem, Journal of Numerical Mathematics, 12 (2004), 39-54.

[21] J. Li and I. M. Navon, Global uniformly convergent finite element methods for singularly perturbed elliptic boundary value problems: higher-order elements, Comput. Methods Appl. Mech. Engrg., 171 (1999), 1-23.

[22] J. Li and M. F. Wheeler, Uniform convergence and superconvergence of mixed finite element methods on anisotropically refined grids, SIAM J. Numer. Anal., 38 (2000), 770-798.

[23] Q. Lin and N. N. Yan, The Construction and Analysis of High Accurate Finite Element Methods, Hebei University Press, Hebei, China, 1996. (in Chinese)

[24] J. J. H. Miller, E. O'Riordan and G. I. Shishkin, Fitted Numerical Methods for Singular Perturbation Problems, World Scientific, Singapore, 1995.

[25] J. T. Oden, Optimal h-p finite element methods, Comput. Methods Appl. Mech. Engrg., 112 (1994), 309-331.

[26] M. H. Protter and H. F. Weinberger, Maximum Principles in Differential Equations, Prentice-Hall, Englewood Cliffs, N.J., 1967.

[27] E. O'Riordan and M. Stynes, A globally uniformly convergent finite element method for a singularly perturbed elliptic problem in two dimensions, Math. Comp., 57 (1991), 47-62.

[28] H. G. Roos, Layer-adapted grids for singular perturbation problems, Z. Angew. Math. Mech., 78 (1998), 291-310.

[29] H. G. Roos, D. Adam and A. Felgenhauer, A novel nonconforming uniformly convergent finite element method in two dimensions, J. Math. Anal. Appl., 201 (1996), 715-755.

[30] H. G. Roos, M. Stynes and L. Tobiska, Numerical Methods for Singularly Perturbed Differential Equations, Springer-Verlag, Berlin, 1996.

[31] A. H. Schatz and L. B. Wahlbin, On the finite element method for singularly perturbed reaction-diffusion problems in two and one dimensions, Math. Comp., 40 (1983), 47-89.

[32] C. Schwab, M. Suri and C. Xenophontos, The hp finite element method for problems in mechanics with boundary layers, Comp. Meth. Appl. Mech. Eng., 57 (1998), 311-334.

[33] V. Thomée, Galerkin finite element methods for parabolic problems, Springer-Verlag, Berlin, 1997.

[34] P. Wesseling, Uniform convergence of discretization error for a singular perturbation problem, Numer. Methods PDEs, 12 (1996), 657-671.

[35] M. F. Wheeler, A priori L_2 error estimates for Galerkin approximations to parabolic partial differential equations, SIAM J. Numer. Anal., 10 (1973), 723-759.

[36] Z. Zhang, Finite element superconvergence approximation for one-dimensional singularly perturbed problems, Numer. Methods PDEs, 18 (2002), 374-395.
[37] Z. Zhang, Finite element superconvergence on Shishkin mesh for 2-d convection-diffusion problems, Math. Comp., 72 (2003), 1147-1177.

Department of Mathematical Sciences, University of Nevada, Las Vegas, NV 89154, USA
E-mail: jichun@unlv.nevada.edu
URL: http://www.nevada.edu/∼jichun