INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 5, Number 4, Pages 729–748

© 2008 Institute for Scientific
Computing and Information

CONSTRUCTION OF BOUNDARY LAYER ELEMENTS FOR SINGULARLY PERTURBED CONVECTION-DIFFUSION EQUATIONS AND L^2- STABILITY ANALYSIS

CHANG-YEOL JUNG AND ROGER TEMAM

Abstract. It has been demonstrated that the ordinary boundary layer elements play an essential role in the finite element approximations for singularly perturbed problems producing ordinary boundary layers. Here we revise the element so that it has a small compact support and hence the resulting linear system becomes sparse, more precisely, block tridiagonal. We prove the validity of the revised element for some singularly perturbed convection-diffusion equations via numerical simulations and via the H^1- approximation error analysis. Furthermore due to the compact structure of the boundary layer we are able to prove the L^2- stability analysis of the scheme and derive the L^2- error approximations.

Key Words. boundary layer, boundary layer element, finite elements, singularly perturbed problem, convection-diffusion, stability, enriched subspaces, exponentially fitted splines.

1. Introduction

In this article we consider linear singularly perturbed boundary value problems of the types:

$$-\epsilon\triangle u^\epsilon - u^\epsilon_x = f \text{ in } \Omega, \tag{1.1a}$$

$$u^\epsilon = 0 \text{ on } \partial\Omega, \tag{1.1b}$$

where $0 < \epsilon << 1,\ \Omega = (0,1)\times(0,1) \subset \mathbb{R}^2$. The function f is assumed to be smooth on $\bar{\Omega}$ but only in Section 3 below we will assume (for the L^2- stability analysis) that f belongs to $L^2(\Omega)$. Problem (1.1) is meant to be a simplified model for a class of problems involving variable coefficients and curved boundaries. However the treatment of these more involved problems only involve additional purely technical difficulties and we thought it would be more appropriate to present our results in the case of this model problem. Variable coefficients equations, curved boundaries and other generalizations will be addressed in separate works.

As ϵ becomes small, the solutions to problem (1.1) generally display, near the boundaries, thin transition layers called boundary layers, which are due to the fact that the boundary conditions of the problem are not the same for $\epsilon > 0$ and $\epsilon = 0$, and then (for $\epsilon > 0$ small) certain derivatives of the solutions become very large near the boundaries. We expect that within these boundary layers, the approximation errors of the discretized system corresponding to problem (1.1)

Received by the editors November 15, 2007 and, in revised form, January 23, 2008.
2000 *Mathematics Subject Classification.* 65N30, 34D15, 76N20.
This work was supported in part by NSF-DMS grants 0305110 and 0604235, and by the Research Fund of Indiana University.

become very large (due to the large H^2- singularities of the boundary layers). When the stiffness of the discretized systems is not properly handled, those large approximation errors at the boundaries propagate in the whole domain due to the convective term, e.g. $-u_x$ in (1.1a), and then the numerical solutions show a highly oscillatory behavior, see e.g. [20], [22], [3], [4], [13], [14] and [15]. Resolving boundary layers by the classical approximation methods requires very fine meshes, which is costly to realize in practice. Indeed, the thickness of the boundary layers (of order $O(\epsilon)$ for ordinary boundary layers (OBL), and of order $O(\epsilon^{1/2})$ for parabolic boundary layers (PBL), see [23], [15]) is usually much smaller than the mesh size h. Notice that our problem (1.1) produces both OBLs at $x = 0$ and PBLs at $y = 0, 1$, which pollute the numerical solutions, globally and locally respectively. In view of properly approximating such problems, it has been suggested by Han and Kellogg, in [10], [11] to add to the Galerkin space suitable profile functions encompassing the main features of the boundary layers, leading to the so-called *enriched subspaces* (ES) method. In this article and related ones [3], [4] we call Boundary Layer Elements (BLE) these profile functions. A related concept is that of *exponentially fitted splines (or L- splines)* (EFS) where the Galerkin basis of spline functions is chosen (constructed) adapted to the operator L_ϵ; see [9] for one-dimensional two-point boundary value problems and [6], [7] and [18] for two-dimensional ones. Our works is closer to the enriched subspaces point of views, and we use asymptotic expansions inspired in part by the work [23] to construct the boundary layer elements using asymptotic expansion techniques. We were not aware of this series of articles on enriched subspaces and exponentially fitted splines when we started our own work in [3], [4], [13] - [16]. Comparisons between these articles and our own past and current work are made below.

Before we proceed, we introduce the notations, the semi-norms and norms for the Sobolev spaces $H^m(\Omega)$, $m \geq 0$ integer (for $m = 0$, it is denoted L^2), which are, respectively, $|u|_{H^m} = \left\{ \sum_{|\alpha|=m} \int_\Omega |D^\alpha u|^2 d\Omega \right\}^{1/2}$ and $\|u\|_{H^m} = \left\{ \sum_{j=0}^{m} |u|_{H^j}^2 \right\}^{1/2}$. The corresponding inner products are $(u, v) = \int_\Omega uv d\Omega$ for L^2, $(\!(u, v)\!) = (u, v) + \int_\Omega \nabla u \cdot \nabla v d\Omega$ for H^1, and $(\!(u, v)\!)_{H^m} = \sum_{|\alpha|\leq m} (D^\alpha u, D^\alpha v)$ for H^m, $m \geq 2$. For the Dirichlet boundary value problem (1.1), we use the Sobolev space $H_0^1(\Omega)$, which is the closure in the space $H^1(\Omega)$ of C^∞ functions compactly supported in Ω. Thanks to the Poincaré inequality the space $H_0^1(\Omega)$ is equipped with the inner product $(\!(\cdot, \cdot)\!) = \int_\Omega \nabla u \cdot \nabla v d\Omega$, and the norm $\|\cdot\| = |\cdot|_{H^1}$.

In [3], [4], [13], [14] and [15], it is demonstrated that the boundary layer elements (BLE), i.e.

$$\phi_0^*(x) = -e^{-x/\epsilon} - (1 - e^{-1/\epsilon})x + 1, \tag{1.2}$$

play an essential role in the finite element approximations for singularly perturbed problems producing the OBLs.

The present article is concerned with two dimensional extensions of [3] and the efficient application of the BLE ϕ_0^*. To solve the problem (1.1) in the finite element context, we consider its weak formulation: *To find* $u \in H_0^1(\Omega)$ *such that*

$$a_\epsilon(u, v) := \epsilon(\!(u, v)\!) - (u_x, v) = (f, v),\ \forall v \in H_0^1(\Omega), \tag{1.3}$$

and then we look for an approximate solution $u_h \in V_h$ such that

$$a_\epsilon(u_h, v_h) = (f, v_h),\ \forall v_h \in V_h, \tag{1.4}$$

where the finite element space V_h will be specified in Section 2.2 below. It contains a classical Q_1 finite element space enriched by a boundary layer element related to ϕ_0^*.

We notice that the BLE ϕ_0^* does not have a compact support and adding the ϕ_0^* in V_h leads to a broad band in the stiffness matrix and hence the corresponding systems are costly to solve. Thus, a first aim in this article is to revise the element ϕ_0^* so that it has a small compact support and to prove that the numerical approximations keep the same accuracy as before. Note that the idea of replacing a given BLE with broad support by one with small support is also advocated in [9]. Then the new system using this revised element ϕ_0 appears to be sparse, more precisely, block tridiagonal and it can be solved very efficiently; it requires essentially the same computing resources as those in the classical methods which use only classical polynomial elements, e.g. Q_1, Q_2. Furthermore, since the stiffness matrix is tridiagonal, via a somehow involved matrix analysis we are able to analyze the L^2- stability; we prove that for any $f \in L^2$, $|u_h|_{L^2(\Omega)} \leq \kappa |f|_{L^2(\Omega)}$, where the positive constant κ is independent of the mesh size h and of the small parameter ϵ, see Section 3 below. Here we denote the mesh size by $h = \max\{h_1, h_2\}$ where $h_1 = 1/M$, $h_2 = 1/N$, and M, N are the number of elements in the x-, and y- directions, respectively. Hence, the number of rectangular elements is MN. In the text κ denotes a generic constant independent of ϵ, h_1, h_2, h, which may be different at different occurrences. But if it needs to be distinguished, we denote it by κ_i, $i = 0, 1, \cdots$, and so on.

Like all linearizations of the Navier-Stokes and Boussinesq equations which are ultimate goal, equation (1.1a) does not contain a reaction term e.g. u^ϵ (unlike equation (1.6) below). Such a reaction term could generate L^2 stability for classical finite element approximations. Indeed the classical schemes which do not use the BLE ϕ_0^* tend to be highly unstable and blow up as $\epsilon \to 0$ as we explain in Remark 3.1 below. To ensure the stability, we could consider a change of variable, e.g. $u = e^{-x} v$, which changes the scheme (1.4) to a slightly more complicated form: To find $v_h \in V_h$ such that for all $w_h \in V_h$,

$$\tilde{a}(v_h, w_h) := \epsilon(\!(v_h, w_h)\!) - (1 - 2\epsilon)(v_{hx}, w_h) + (1 - \epsilon)(v_h, w_h) = (e^x f, w_h). \tag{1.5}$$

Then by the transformation $u_h = e^{-x} v_h$, we can recover the approximate solutions u_h from the modified scheme (1.5). But in the numerical simulations we found that the original scheme (1.4) using the BLE ϕ_0^* indeed attains much better accuracies (e.g. for $\epsilon = 10^{-3}$, $h_1 = 1/M = 1/20$, $h_2 = 1/N = 1/10$, the L^2- errors of the scheme (1.4), (1.5) are respectively 3.7465E-04 and 1.2287E-03, for more detail see Tables 1, 2 in [13]). By numerical simulations we also observed that the scheme (1.4) converges as $\epsilon \to 0$ to a nonsingular system (i.e. the limit linear system is invertible) as explained in Section 3. This limit behavior and the better accuracies of the original scheme (1.4) in the simulations motivate the L^2- stability analysis via the matrix analysis; the L^2- stability analysis of problem (1.4) is involved and we did not find it available in the literature not using a change of variable. Beside this numerical motivation, another reason that we do not change Eq. (1.1) using this change of variable is that we will confront in more general problems (say, $-\epsilon \triangle u^\epsilon - \mathbf{b}(x, y) \cdot \nabla u^\epsilon = f$) many situations that cannot attain the L^2- stability in the numerical simulations by a simple change of variable. These include the cases where $\mathbf{b}$ attains $\mathbf{0}$ in the interior of Ω, e.g. $-\epsilon \triangle u^\epsilon + x u^\epsilon_x = f$ in $(-1, 1) \times (-1, 1)$ and when periodic boundary conditions are imposed, e.g. problem (1.1) with (1.1b) replaced by e.g. $u(x, 0) = u(x, 1) = 0$ and $u^\epsilon(x, y) = u^\epsilon(x + 1, y)$.

This article is organized as follows: We start in Section 2 by modifying ϕ_0^* slightly and we construct a new boundary layer element ϕ_0 which has a small compact support; this will be used in the stability analysis, and it will be used elsewhere in the numerical simulations. In Section 2.2, we consider new finite element schemes using the element ϕ_0; more precisely, the function ϕ_0 will be incorporated into the appropriate finite elements space that we will define. We then perform the L^2-stability analysis via a matrix method in Section 3 and derive error estimates in H^1 and L^2 in Section 4.

A number of technical hypotheses will be needed on h_1, h_2 and ϵ, namely (H0) to (H5) (see (2.6), (2.7), (3.19), (3.29), (3.44), (3.45)).

Before we proceed, we want to develop a comparison between (EFS), (ES) and (BLE). To compare these methods, we take a simple one dimensional singularly perturbed problem:

$$L_\epsilon u^\epsilon := -\epsilon u^\epsilon_{xx} - u^\epsilon_x + u^\epsilon = f \text{ in } (0,1),\ u^\epsilon(0) = u^\epsilon(1) = 0. \tag{1.6}$$

Its weak formulation and finite elements scheme are, respectively, defined as: *To find* $u = u^\epsilon \in H^1_0(0,1)$, $u_h \in V_h \subset H^1_0(0,1)$ *such that*

$$B(u,v) = (f,v),\ \forall v \in H^1_0(0,1),\ B(u_h,v_h) = (f,v_h),\ \forall v_h \in V_h, \tag{1.7}$$

where $B(u,v) = \epsilon(\!(u,v)\!) - (u_x,v) + (u,v)$. Then it is easy to verify the coercivity $B(u,u) \geq \|u\|^2_\epsilon := \epsilon|u|^2_{H^1} + |u|^2_{L^2}$. From (1.7) we also find that $B(u-u_h,v_h) = 0$ and thus we classically find that $\|u-u_h\|^2_\epsilon = B(u-u_h,u-u_h) = B(u-u_h,u-\tilde{u}_h)$, for any $\tilde{u}_h \in V_h$. By the Poincaré and the Cauchy-Schwarz inequalities, after some elementary calculations, we conclude that

$$\|u-u_h\|_\epsilon \leq \kappa \min\{\epsilon^{1/2}|u-\tilde{u}_h|_{H^1} + \epsilon^{-1/2}|u-\tilde{u}_h|_{L^2}, |u-\tilde{u}_h|_{H^1}\}. \tag{1.8}$$

What is in common in (EFS), (ES) and (BLE) is that all methods attain the ϵ- uniform convergence in the weighted energy norm $\|\cdot\|_\epsilon$ and all of them use the singular functions which absorb the singularities due to the small ϵ. The differences are the construction of the basis in the finite elements space V_h and the method to derive the singular functions. In the (EFS), the finite elements basis are constructed to adapt to the differential operator L_ϵ of Eq. (1.7) in each subinterval. The basis elements φ_i, $i = 1, \cdots, N-1$, are defined as follows: supp $\varphi_i = [x_{i-1}, x_{i+1}]$, $L_\epsilon\varphi_i = 0$ in (x_{i-1}, x_i), $\varphi_i(x_{i-1}) = 0$, $\varphi_i(x_i) = 1$ and $L_\epsilon\varphi_i = 0$ in (x_i, x_{i+1}), $\varphi_i(x_i) = 1$, $\varphi_i(x_{i+1}) = 0$. It is known that $\|u-u_h\|_\epsilon \leq \kappa h^{1/2}$ (see [20]); this estimate can be derived using the fact that $L_\epsilon\varphi_i = 0$ for all (x_{i-1}, x_{i+1}), $i = 1, \cdots, N-1$. The advantage of (EFS) is that it does not assume any a priori knowledge of where the boundary (or interior) layers occur but it is known that the exponentially fitted splines φ_i can introduce spurious interior layers in the numerical simulations where the solutions behave smoothly (see [9]) and since the basis are adapted to L_ϵ, if L_ϵ is complicated, e.g. with variable coefficients, the basis φ_i can be complicated and some approximate form of L_ϵ might be necessary (e.g. $\bar{L}$- splines) (see [20]). To avoid the spurious numerical interior layers, we can apply (EFS) only in a region where a boundary layer occurs and use the classical polynomial elements outside the boundary layers but we then need a priori knowledge on the boundary layers. In this respect, the (ES) and (BLE) which we now describe provide a way to analyze the structures of boundary layers. The (ES) uses the classical polynomial elements enriched with the singular function ϕ^*_0. This method can be justified by the decomposition (see [11]) of the solutions u^ϵ of Eq. (1.1): $u^\epsilon = c_0(\epsilon)\phi^*_0(x) + c_1(x) + c_2(x)$, where $|c_0(\epsilon)| \leq \kappa$, $c_1(x), c_2(x) \in H^1_0(0,1)$ and

$$|c_1(x)| + \epsilon|c_{1x}(x)| + \epsilon^2|c_{1xx}(x)| \leq \kappa\epsilon\ e^{-x/(2\epsilon)}, \tag{1.9a}$$

$$|c_2(x)| + |c_{2x}(x)| + |c_{2xx}(x)| \leq \kappa. \tag{1.9b}$$

We first notice that by the classical interpolation theory applied to $c_2(x)$, there exists a piecewise linear function Πc_2 such that $|c_2 - \Pi c_2|_{H^m} \leq \kappa h^{2-m}|c_2|_{H^2} \leq$

κh^{2-m} and we can also deduce that $|c_1|_{H^m} \le \kappa \epsilon^{3/2-m}$. Setting $\tilde{u}_h = c_0(\epsilon)\phi_0^* + \Pi c_2$, we then find that for $m = 0, 1$,

$$|u - \tilde{u}_h|_{H^m} \le |c_1|_{H^m} + |c_2 - \Pi c_2|_{H^m} \le \kappa(\epsilon^{3/2-m} + h^{2-m}). \tag{1.10}$$

Using the estimate (1.8), for $\epsilon \le h^2$ or $h^2 < \epsilon \le h$, we can deduce that $\|u - u_h\|_\epsilon \le \kappa h$. For $\epsilon > h$, by the classical interpolation theory applied to $c_1(x)$, we find that $|c_1 - \Pi c_1|_{H^m} \le \kappa h^{2-m}|c_1|_{H^2} \le \kappa\epsilon^{-1/2}h^{2-m}$ and thus, setting $\tilde{u}_h = c_0(\epsilon)\phi_0^* + \Pi c_2 + \Pi c_1$, we have

$$|u - \tilde{u}_h|_{H^m} \le |c_1 - \Pi c_1|_{H^m} + |c_2 - \Pi c_2|_{H^m} \le \kappa\epsilon^{-1/2}h^{2-m}. \tag{1.11}$$

Using (1.8) again, we deduce that $\|u - u_h\|_\epsilon \le \kappa h$ for $\epsilon > h$ (thus for all $\epsilon > 0$). Finally, in the (BLE), by the singular perturbation analysis in the H^m space (*correctors* in the context of [17]), we derive that $|u^\epsilon - c_0(\epsilon)\phi_0^*|_{H^2} \le \kappa$ (see [3]) and thus by the classical interpolation theory applied to $u^\epsilon - c_0(\epsilon)\phi_0^*$ we deduce that there exists a function $\tilde{u}_h := c_0(\epsilon)\phi_0^* + \Pi(u^\epsilon - c_0(\epsilon)\phi_0^*)$ such that

$$|u^\epsilon - \tilde{u}_h|_{H^m} \le \kappa h^{2-m}|u^\epsilon - c_0(\epsilon)\phi_0^*|_{H^2} \le \kappa h^{2-m}. \tag{1.12}$$

Using (1.8), we deduce that $\|u - u_h\|_\epsilon \le \kappa h$ for all $\epsilon > 0$. Both (BLE) and (ES) use the singular function ϕ_0^* which is globally smooth. The (EFS) uses φ_i's which have a small compact support and thus are efficient in the numerical implementations. In this article we will modify ϕ_0^* to have small compact support under some smallness assumptions for ϵ. The finite element spaces V_h are $\{\phi_0^*, (\phi_i)_{i=1}^{N-1}\}$, ϕ_i are the hat functions, for (BLE) and (ES) and $\{(\varphi_i)_{i=1}^{N-1}\}$ for (EFS). We note that the functions, ϕ_0^* and φ_i's, absorb the singularities due to the small ϵ. The (BLE) and (ES) try to reveal, by some mathematical analysis, the precise structure of the singularities which cause the instability in the numerical schemes as much as possible. In the H^2 space, ϕ_0^* is the right function to absorb the H^2- singularities, namely, $|u^\epsilon - c_0(\epsilon)\phi_0^*|_{H^2} \le \kappa$. To extend to higher order numerical methods, we will need to find singular functions to absorb the H^m- singularities, $m \ge 3$. For that purpose, the correctors are relatively suitable to reveal the structures of the singularities in the H^m spaces, $\forall m \ge 1$. But the structures are getting complicated and using them is another problem to solve which we aim elsewhere. Furthermore, in higher dimensional spaces, unlike the one-dimensional one, there are many challenging singularities, parabolic boundary layers, interior layers (characteristic layers, turning points) as well as ordinary boundary layers and (one dimensional) turning points as in the one dimensional space. The boundary layers are the easiest to detect and the ordinary boundary layers are the most feasible functions to be discretized in the numerical simulations. The other correctors need to be modified for the computational purpose and this will be a coming subject of study. We believe that our numerical methods, closely connected to the singular perturbation theory (they actually complement each other), have a large potential to explore those challenging problems and the higher-order numerical methods.

2. Boundary Layer Elements (BLE)

Starting with $\phi_0^* = -e^{-x/\epsilon} - (1 - e^{-1/\epsilon})x + 1$, which belongs to $C^\infty([0,1]) \cap H_0^1(0,1)$, we first recall the following lemma from [15] which states that ϕ_0^* absorbs the H^2- singularity of the OBLs. Here we impose the condition

$$f(x, 0) = f(x, 1) = 0 \tag{2.1}$$

so that the PBLs are mild. More precisely, they are $O(1)$- quantity in H^2 space (see [15]).

Lemma 2.1. *Assume that the condition (2.1) holds. Then there exist a positive constant κ independent of ϵ, and a smooth function $g = g^\epsilon(y) \in H_0^1(0,1)$ with $|g|_{H^2(0,1)} \leq \kappa$ such that*

$$\|u^\epsilon - g\phi_0^*\|_{H^2(\Omega)} \leq \kappa. \tag{2.2}$$

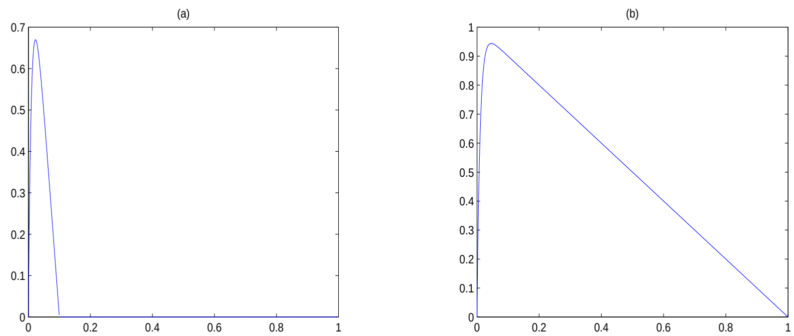

FIGURE 1. Boundary layer elements for $\epsilon = 0.01$, $h_1 = 0.1$ here; (a): ϕ_0, (b): ϕ_0^*.

2.1. Constructing BLEs. We now sightly modify ϕ_0^*, and derive a new boundary layer element ϕ_0 which has a small compact support, see (a) of Figure 1:

$$\phi_0 = [-e^{-x/\epsilon} - (1 - e^{-h_1/\epsilon})x/h_1 + 1]\chi_{[0,h_1]}(x), \tag{2.3}$$

where $\chi_{[\alpha,\beta]}(x)$ is the characteristic function of $[\alpha, \beta]$. We note that ϕ_0 belongs to $H_0^1(0,1)$. To compare the two elements, ϕ_0^* and ϕ_0, ϕ_0 given in (2.3), we rewrite

$$\phi_0 = \left[-\exp\left(-\frac{x}{\epsilon}\right) + \exp\left(-\frac{h_1}{\epsilon}\right)\frac{x}{h_1}\right]\chi_{[0,h_1]} + \left[1 - \frac{x}{h_1}\right]\chi_{[0,h_1]}. \tag{2.4}$$

Since the last term, $(1 - x/h_1)\chi_{[0,h_1]}$, is exactly a hat function at $x = 0$, we thus easily verify that there exist c_i's such that

$$\left[1 - \frac{x}{h_1}\right]\chi_{[0,h_1]} + \sum_{i=1}^{M-1} c_i\phi_i = 1 - x, \ \forall x \in [0,1], \tag{2.5}$$

where the ϕ_i are the usual hat functions whose definition is recalled in Section 2.2.

We now make two smallness hypotheses for ϵ, namely

$$\text{(H0)} \quad -\epsilon \ln \epsilon \leq \frac{2}{3}h_1, \left(\mathit{or}\ \exp\left(-\frac{h_1}{\epsilon}\right) \leq \epsilon^{3/2}\right), \tag{2.6}$$

$$\text{(H1)} \quad \epsilon \leq \kappa_0 h_1,\ 0 < \kappa_0 \leq 1/4; \tag{2.7}$$

e.g. for $\epsilon = 10^{-2}$, 10^{-3}, respectively, $-\epsilon \ln \epsilon \approx 4.6052 \times 10^{-2}$, 6.9078×10^{-3}. These hypothesis simplify calculations here and later and (2.6) will be used to majorize the expressions (3.5h) - (3.5k) below.

Lemma 2.2. *Assume that (H0)-(2.6), (H1)-(2.7) hold and let $\Phi_0 = \phi_0 + \sum_{i=1}^{M-1} c_i\phi_i$ where the c_i's are as in (2.5). Then the following inequalities hold: for $m = 0, 1$,*

$$|\phi_0^* - \Phi_0|_{H^m(0,1)} \leq \kappa h_1^{2-m}, \tag{2.8a}$$

$$|\Phi_0|_{H^m(0,1)} \leq \kappa\epsilon^{-m/2}. \tag{2.8b}$$

Proof. We first write from (1.2) and (2.5) that $\phi_0^* - \Phi_0 = J_1 + J_2 + J_3$, where

$$J_1 = -e^{-x/\epsilon}\chi_{[h_1,1]}(x),\ J_2 = -e^{-h_1/\epsilon}x/h_1\chi_{[0,h_1]}(x),\ J_3 = e^{-1/\epsilon}x.$$

For J_1, by the assumption (2.6),

$$|J_1|^2_{H^m} \le \kappa\epsilon^{-2m}\int_{h_1}^1 e^{-2x/\epsilon}dx \le \kappa\epsilon^{-2m+1}e^{-2h_1/\epsilon} \le \kappa\epsilon^{-2m+4}.$$

Hence, by (2.7), $|J_1|_{H^m} \le \kappa\epsilon^{2-m} \le \kappa h_1^{2-m}$. For J_2, we find from (2.6), (2.7) that

$$|J_2|_{H^m} \le \kappa h_1^{-1}e^{-h_1/\epsilon}|x^{1-m}|_{L^2(0,h_1)} \le \kappa h_1^{(1-2m)/2}\epsilon^{3/2} \le \kappa h_1^{2-m}.$$

The J_3 is an exponentially small term which is absorbed in the other norm estimates and thus (2.8a) follows. Then (2.8b) follows from (2.8a) observing that $|\phi_0^*|_{L^2(0,1)} \le \kappa$, $|\phi_0^*|_{H^1(0,1)} \le \kappa\epsilon^{-1/2}$, and $|\Phi_0|_{H^m} \le |\phi_0^*|_{H^m} + |\phi_0^* - \Phi_0|_{H^m}$. $\square$

2.2. Finite Element Discretizations. We now define the finite element spaces and introduce the new schemes making use of the classical Q_1 elements (hat functions), ϕ_i and ψ_j, $i = 1, \cdots, M-1$, $j = 1, \cdots, N-1$, to which we add the boundary layer element (BLE) ϕ_0 which absorbs the singularity at $x = 0$ due to the OBLs. We thus introduce the following finite element space for the scheme (1.4) in the form of tensor product of two spaces:

$$V_h = \{\phi_0, (\phi_i)_{i=1}^{M-1}\} \otimes \{(\psi_j)_{j=1}^{N-1}\} \subset H_0^1(\Omega). \tag{2.9}$$

To derive the H^1- and L^2- error estimates below for the scheme (1.4) with (2.9), we will need the following interpolation inequalities.

Lemma 2.3. *Assume that (H0) - (2.6), (H1) - (2.7) hold. Then there exists an interpolant $\tilde{u}_h \in V_h$ such that for $m = 0, 1$,*

$$\|u^\epsilon - \tilde{u}_h\|_{L^2(\Omega)} \le \kappa h^2, \tag{2.10a}$$

$$\|u^\epsilon - \tilde{u}_h\|_{H^1(\Omega)} \le \kappa(h + h_2^2\epsilon^{-1/2}). \tag{2.10b}$$

Proof. From the classical interpolation results, see e.g. [5], [13], [21] applied to $\bar{u}^\epsilon = u^\epsilon - g\phi_0^* \in H_0^1(\Omega)$, and by (2.2), we can find its interpolant $\Pi\bar{u}^\epsilon \in V_h$ such that

$$I_1(m) := |u^\epsilon - g\phi_0^* - \Pi\bar{u}^\epsilon|_{H^m(\Omega)} \le \kappa h^{2-m}|u^\epsilon - g\phi_0^*|_{H^2(\Omega)} \le \kappa h^{2-m}.$$

By (H0), (H1), we easily find from (2.8a) that

$$I_2(m) := |g\phi_0^* - g\Phi_0|_{H^m(\Omega)} \le \kappa h_1^m.$$

Then again, by the classical interpolation results applied this time to $g = g^\epsilon(y)$, we can also find its interpolant, piecewise linear function, $\Pi_y g \in H_0^1(0,1)$ such that

$$|g - \Pi_y g|_{H^m(0,1)} \le \kappa h_2^{2-m};$$

then using the estimates (2.8b), we easily verify the following estimates, observing that Φ_0 depends only on x, and g and ψ_j depend only on y:

$$\begin{aligned} I_3(0) :=& |\Phi_0 g - \Phi_0\Pi_y g|_{L^2(\Omega)} \le |g - \Pi_y g|_{L^2(0,1)}|\Phi_0|_{L^2(0,1)} \le \kappa h_2^2, \\ I_3(1) :=& |\Phi_0 g - \Phi_0\Pi_y g|_{H^1(\Omega)} \le \kappa|g - \Pi_y g|_{H^1(0,1)}|\Phi_0|_{L^2(0,1)} \\ &+ \kappa|g - \Pi_y g|_{L^2(0,1)}|\Phi_0|_{H^1(0,1)} \le \kappa h_2 + \kappa h_2^2\epsilon^{-1/2}. \end{aligned}$$

The lemma follows after setting $\tilde{u}_h = \Pi\bar{u}^\epsilon + \Phi_0\Pi_y g \in V_h$ and observing that

$$|u^\epsilon - \Pi\bar{u}^\epsilon - \Phi_0\Pi_y g|_{H^m} \le I_1(m) + I_2(m) + I_3(m).$$

$\square$

Remark 2.1. For one dimensional problem, since we do not take into account the approximation errors in y due to $g(y)$ in the proof of Lemma 2.3, we can conclude that for $m = 0, 1$, $\|u^\epsilon - \tilde{u}_h\|_{H^1(0,1)} \leq \kappa h_1^{2-m}$ (see [3]).

3. L^2- stability Analysis

When ϵ is small or $\epsilon \to 0$, one would expect that the linear system (1.4) is highly ill-conditioned. However, we will show how the new boundary layer element ϕ_0 stabilizes (or, absorbs the singularities of) the discretized linear system (1.4). Since the limit system (i.e. when $\epsilon = 0$) has a simple structured block matrix which appears in (3.5) and (3.6) below, we are able to analyze the L^2- stability via a matrix method.

Setting

$$u_h = \sum_{i=0}^{M-1} \sum_{j=1}^{N-1} a_{ij} \phi_i \psi_j, \tag{3.1}$$

where ϕ_0 is the BLE as in (2.3), ϕ_i, ψ_j, $i = 1, \cdots, M-1$, $j = 1, \cdots, N-1$, are hat functions, we then write Eq. (1.4) with (2.9) with $F(v) = \int_\Omega f v d\Omega$, for any $f \in L^2$ not necessarily satisfying (2.1), in the matrix from:

$$\Gamma_\epsilon \mathbf{a} = \mathbf{b}. \tag{3.2}$$

The Γ_ϵ and the $\mathbf{b}$ are the stiffness matrix and the load vector specified in (3.5) and (3.41) respectively below, and

$$\mathbf{a} = \begin{pmatrix} \mathbf{a}_0 \\ \mathbf{a}_1 \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{a}_i \\ \cdot \\ \mathbf{a}_{M-2} \\ \mathbf{a}_{M-1} \end{pmatrix}_{M\times 1}, \quad \mathbf{a}_i = \begin{pmatrix} a_{i1} \\ a_{i2} \\ \cdot \\ a_{ij} \\ \cdot \\ a_{i,N-1} \end{pmatrix}_{(N-1)\times 1}; \quad \mathbf{b} = \begin{pmatrix} \mathbf{b}_0 \\ \mathbf{b}_1 \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{b}_k \\ \cdot \\ \mathbf{b}_{M-2} \\ \mathbf{b}_{M-1} \end{pmatrix}_{M\times 1}, \quad \mathbf{b}_k = \begin{pmatrix} b_{k1} \\ b_{k2} \\ \cdot \\ b_{kl} \\ \cdot \\ b_{k,N-1} \end{pmatrix}_{(N-1)\times 1}.$$

Note that the matrix Γ_ϵ is of size $[M \times (N-1)]^2 = [(1-h_2)/(h_1 h_2)]^2$. For the purpose of the analysis below, we introduce the Euclidian inner product $< \cdot, \cdot >$ on $\mathbb{R}^N$, $< \mathbf{a}, \mathbf{b} >= \sum_{i=1}^N a_i b_i$, where $\mathbf{a} = (a_1, \cdots, a_i, \cdots, a_N)^T$, $\mathbf{b} = (b_1, \cdots, b_i, \cdots, b_N)^T$. We also introduce the corresponding matrix norm $\|\Lambda\| = \max_{\|\mathbf{x}\|_2 = 1} \|\Lambda \mathbf{x}\|_2$, where $\|\mathbf{x}\|_2 = \sqrt{< \mathbf{x}, \mathbf{x} >}$. We recall the following well-known facts, see [8], [19].

For a matrix $\Lambda \in \mathbb{R}^{N\times N}$, setting $\overline{\rho}(\Lambda) = \max_{\lambda \in \sigma(\Lambda)} |\lambda|$, $\underline{\rho}(\Lambda) = \min_{\lambda\in\sigma(\Lambda)} |\lambda|$ with $\sigma(\Lambda) = \{\lambda \in \mathbb{C}; \lambda \text{ an eigenvalue of } \Lambda\}$, we have $\|\Lambda\| = \left\{\overline{\rho}(\Lambda^T\Lambda)\right\}^{1/2}$, and if Λ is invertible, $\|\Lambda^{-1}\| = \left\{\underline{\rho}(\Lambda^T\Lambda)\right\}^{-1/2}$, where Λ^T is the transpose of Λ. In particular, if Λ is a symmetric matrix, i.e. $\Lambda^T = \Lambda$, then $\|\Lambda\| = \overline{\rho}(\Lambda)$, $\|\Lambda^{-1}\| = \left\{\underline{\rho}(\Lambda)\right\}^{-1}$.

We will explicitly calculate the entries of the stiffness matrix Γ_ϵ. For that purpose and for the analysis later on, it is convenient here to define the identity matrices, I and $\tilde{I}$, and the tridiagonal matrices, S, U which will be used repeatedly.

$$I = \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & \cdot & & \\ & & & 1 & \\ & & & & 1 \end{pmatrix}_{(N-1)\times(N-1)}, \quad \tilde{I} = \begin{pmatrix} I & & & & \\ & I & & & \\ & & \cdot & & \\ & & & I & \\ & & & & I \end{pmatrix}_{M\times M},$$

$$\mathrm{S} = \begin{pmatrix} 4 & 1 & & & \\ 1 & 4 & 1 & & \\ & \cdot & \cdot & \cdot & \\ & & 1 & 4 & 1 \\ & & & 1 & 4 \end{pmatrix}_{(N-1)\times(N-1)}, \quad \mathrm{U} = \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \cdot & \cdot & \cdot & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix}_{(N-1)\times(N-1)}.$$

From the Gershgorin circle theorem, we easily find that

$$\|\mathrm{S}\| \le 6,\ \|\mathrm{S}^{-1}\| \ \le 1/2,\ \|\mathrm{U}\| \le 4; \tag{3.3}$$

however for $\|\mathrm{U}^{-1}\|$, since the Gershgorin discs of U contain the point 0, we are not able to find the upper bound of $\|\mathrm{U}^{-1}\|$. The relations (3.3) for S follow also from the explicit expression of its eigenvalues which are, see e.g. [1]: $\lambda_i = 6-4\left(\sin\frac{i\pi}{2N}\right)^2$, $i = 1, \cdots, N-1$.

We now compute the $(l,j)^{th}$ entry of each $(k,i)^{th}$ block, $a_\epsilon(\phi_i\psi_j, \phi_k\psi_l)$, of the stiffness matrix Γ_ϵ:

$$\begin{aligned} a_\epsilon(\phi_i\psi_j, \phi_k\psi_l) &= \epsilon \int_0^1 \frac{d\phi_i}{dx}\frac{d\phi_k}{dx}dx \int_0^1 \psi_j\psi_l dy \\ &+ \epsilon \int_0^1 \phi_i\phi_k dx \int_0^1 \frac{d\psi_j}{dy}\frac{d\psi_l}{dy}dy - \int_0^1 \frac{d\phi_i}{dx}\phi_k dx \int_0^1 \psi_j\psi_l dy; \end{aligned} \tag{3.4}$$

the indices k, i range from 0 to $M-1$, and the indices l, j range from 1 to $N-1$. Using e.g. MAPLE we explicitly compute the integrals in (3.4) and we find that:

$$\Gamma_\epsilon = \begin{pmatrix} \mathrm{A}_\epsilon & \mathrm{B}_\epsilon & & & & & \\ \mathrm{C}_\epsilon & \mathrm{D}_\epsilon & \mathrm{E}_\epsilon & & & & \\ & \mathrm{F}_\epsilon & \mathrm{D}_\epsilon & \mathrm{E}_\epsilon & & & \\ & & \cdot & \cdot & \cdot & & \\ & & & \cdot & \cdot & \cdot & \\ & & & & \mathrm{F}_\epsilon & \mathrm{D}_\epsilon & \mathrm{E}_\epsilon \\ & & & & & \mathrm{F}_\epsilon & \mathrm{D}_\epsilon \end{pmatrix}_{M\times M}, \tag{3.5a}$$

where $\mathrm{A}_\epsilon = \mathrm{A} + \tilde{\mathrm{A}}_\epsilon,\ \mathrm{B}_\epsilon = -\mathrm{A} + \tilde{\mathrm{B}}_\epsilon,\ \mathrm{C}_\epsilon = \mathrm{A} + \tilde{\mathrm{C}}_\epsilon,\ \mathrm{E}_\epsilon = -\mathrm{A} + \tilde{\mathrm{E}}_\epsilon,\ \mathrm{F}_\epsilon = \mathrm{A} + \tilde{\mathrm{E}}_\epsilon$, and

$$\mathrm{A} = \left(\frac{h_2}{12}\right)\mathrm{S}, \tag{3.5b}$$

$$\tilde{\mathrm{A}}_\epsilon = \left(\frac{(\xi_1 - 2\epsilon)h_2}{12h_1}\right)\mathrm{S} + \epsilon\left(\frac{\xi_2 + 2h_1^2 - 9h_1\epsilon + 12\epsilon^2}{6h_1h_2}\right)\mathrm{U}, \tag{3.5c}$$

$$\tilde{\mathrm{B}}_\epsilon = \left(\frac{(\xi_3 + 2\epsilon)h_2}{12h_1}\right)\mathrm{S} + \epsilon\left(\frac{\xi_4 - 6\epsilon^2 + h_1^2}{6h_1h_2}\right)\mathrm{U}, \tag{3.5d}$$

$$\tilde{\mathrm{C}}_\epsilon = -\left(\frac{(\xi_3 + 2\epsilon)h_2}{12h_1}\right)\mathrm{S} + \epsilon\left(\frac{\xi_4 - 6\epsilon^2 + h_1^2}{6h_1h_2}\right)\mathrm{U}, \tag{3.5e}$$

$$\tilde{\mathrm{E}}_\epsilon = -\epsilon\left(\frac{h_2}{6h_1}\right)\mathrm{S} + \epsilon\left(\frac{h_1}{6h_2}\right)\mathrm{U}, \tag{3.5f}$$

$$\mathrm{D}_\epsilon = \epsilon\left(\frac{h_2}{3h_1}\right)\mathrm{S} + \epsilon\left(\frac{2h_1}{3h_2}\right)\mathrm{U}, \tag{3.5g}$$

where the ξ_i are all exponentially small as $\epsilon \to 0$:

$$\xi_1 = (-2\epsilon - h_1)e^{-2h_1/\epsilon} + 4\epsilon e^{-h_1/\epsilon}, \tag{3.5h}$$

$$\xi_2 = (9h_1\epsilon + 2h_1^2 + 12\epsilon^2)e^{-2h_1/\epsilon} + (2h_1^2 - 24\epsilon^2)e^{-h_1/\epsilon}, \tag{3.5i}$$

$$\xi_3 = (-2\epsilon - h_1)e^{-h_1/\epsilon}, \tag{3.5j}$$

$$\xi_4 = (6h_1\epsilon + 6\epsilon^2 + 2h_1^2)e^{-h_1/\epsilon}. \tag{3.5k}$$

The matrix Γ_ϵ is block tridiagonal, and each block, A_ϵ to F_ϵ, is tridiagonal too. The blank blocks are $(N-1)\times(N-1)$- zero matrices, we will sometimes denote a zero matrix by 0 if it needs to be distinguished.

It is noteworthy that the matrices $\tilde{\mathrm{A}}_\epsilon$ to $\tilde{\mathrm{E}}_\epsilon$ converge to 0 when $\epsilon \to 0$ and

$$\mathrm{A}_\epsilon,\ \mathrm{C}_\epsilon,\ \mathrm{F}_\epsilon \to \mathrm{A}, \quad \mathrm{B}_\epsilon,\ \mathrm{E}_\epsilon \to -\mathrm{A}, \quad \mathrm{D}_\epsilon \to 0. \tag{3.6}$$

The limit matrix, i.e. Γ_ϵ when $\epsilon = 0$, is a simple structured block matrix ($= \mathrm{A}\Lambda_0$ as in (3.16) below) and, furthermore, its inverse matrix can be found explicitly, see Lemma 3.2 below.

We now consider the two following cases, namely when $\epsilon \le \kappa h_2^2$ and when $\kappa h_2^2 \le \epsilon \le \kappa h_2$. The case $\epsilon \le \kappa h_2^2$, which is presented in the next section, gives us some insights on why the new scheme (1.4) with (2.9) is stable and the classical scheme is not.

3.1. Case $\epsilon \le \kappa h_2^2$. We will use Lemma 3.1 below which estimates the matrix norm of a (block) band matrix or dense matrix.

Let A_{ki} denote the $(k+1, i+1)^{th}$ block in the block matrix Λ. We define its bandwidth w as follows: $w = p+q-1$ if the entry blocks $A_{ki} = 0$ whenever $k+p \le i$ or $i+q \le k$.

Lemma 3.1. *Let w be the bandwidth of a block matrix Λ with blocks $\{A_{ki}\}$. Then*

$$\|\Lambda\| \le w \times \max_{k,i}\{\|A_{ki}\|\}. \tag{3.7}$$

Proof. Let $\mathbf{x} = (\mathbf{x}_0, \mathbf{x}_1, \cdots, \mathbf{x}_i, \cdots, \mathbf{x}_{M-1})^T$. We then easily verify that

$$\|\Lambda\mathbf{x}\|_2^2 = \sum_{k=0}^{M-1} \left\| \sum_{i=k-q+1}^{k+p-1} A_{ki}\mathbf{x}_i \right\|_2^2 \le \max_{k,i}\{\|A_{ki}\|^2\} \sum_{k=0}^{M-1} \left[\sum_{i=k-q+1}^{k+p-1} \|\mathbf{x}_i\|_2 \right]^2, \tag{3.8}$$

where p, q are as above. By the Cauchy-Schwarz inequality

$$\left[\sum_{i=k-q+1}^{k+p-1} \|\mathbf{x}_i\|_2 \right]^2 \le (p+q-1) \sum_{i=k-q+1}^{k+p-1} \|\mathbf{x}_i\|_2^2, \tag{3.9}$$

and hence, permuting the summation

$$\sum_{k=0}^{M-1} \left[\sum_{i=k-q+1}^{k+p-1} \|\mathbf{x}_i\|_2 \right]^2 \le (p+q-1)^2 \sum_{i=0}^{M-1} \|\mathbf{x}_i\|_2^2. \tag{3.10}$$

Therefore, from (3.8) and (3.10),

$$\|\Lambda\| = \max_{\|\mathbf{x}\|_2 = 1} \|\Lambda\mathbf{x}\|_2 \le (p+q-1) \max_{k,i}\{\|A_{ki}\|\}; \tag{3.11}$$

the lemma follows. □

Remark 3.1. The norm of each $\|A_{ki}\|$ can be estimated as in Lemma 3.1. More precisely, if A_{ki} is a band matrix with a bandwidth $\bar{w}$, then

$$\|A_{ki}\| \le \bar{w} \times \max_{l,j}\{|a_{lj}^{ki}|\}, \tag{3.12}$$

where a_{lj}^{ki} is the $(l, j)^{th}$ entry of A_{ki}.

In particular, if the matrix Λ is of size $M \times M$ and its bandwidth w depends on $M = 1/h_1$, e.g. a matrix with no zero entries, we easily see that since $w \le 2M$,

$$\|\Lambda\| \le \frac{\kappa}{h_1} \times \max_{k,i}\{\|A_{ki}\|\}. \tag{3.13}$$

Furthermore, if each block A_{ki} is of size $(N-1)\times(N-1)$ and its bandwidth $\bar{w}$ depends on $N = 1/h_2$, since $\bar{w} \leq 2(N-1)$, we infer from (3.12), (3.13) that

$$\|\Lambda\| \leq w \times \bar{w} \times \max_{l,j}\{|a_{lj}^{ki}|\} \leq \frac{\kappa}{h_1 h_2} \times \max_{l,j,k,i}\{|a_{lj}^{ki}|\}. \tag{3.14}$$

Thanks to (3.5), multiplying equation (3.2) by A^{-1}, we write this equation as follows:

$$\mathrm{A}^{-1}\Gamma_\epsilon \mathbf{a} = (\Lambda_0 + \Lambda_\epsilon)\mathbf{a} = \tilde{\mathbf{b}}, \tag{3.15}$$

where

$$\Lambda_0 = \begin{pmatrix} I & -I & & & & & \\ I & 0 & -I & & & & \\ & I & 0 & -I & & & \\ & & \cdot & \cdot & \cdot & & \\ & & & \cdot & \cdot & \cdot & \\ & & & & I & 0 & -I \\ & & & & & I & 0 \end{pmatrix}_{M\times M}, \tag{3.16}$$

$$\Lambda_\epsilon = \begin{pmatrix} \mathrm{A}^{-1}\tilde{\mathrm{A}}_\epsilon & \mathrm{A}^{-1}\tilde{\mathrm{B}}_\epsilon & & & & & \\ \mathrm{A}^{-1}\tilde{\mathrm{C}}_\epsilon & \mathrm{A}^{-1}\mathrm{D}_\epsilon & \mathrm{A}^{-1}\tilde{\mathrm{E}}_\epsilon & & & & \\ & \mathrm{A}^{-1}\tilde{\mathrm{E}}_\epsilon & \mathrm{A}^{-1}\mathrm{D}_\epsilon & \mathrm{A}^{-1}\tilde{\mathrm{E}}_\epsilon & & & \\ & & \cdot & \cdot & \cdot & & \\ & & & \cdot & \cdot & \cdot & \\ & & & & \mathrm{A}^{-1}\tilde{\mathrm{E}}_\epsilon & \mathrm{A}^{-1}\mathrm{D}_\epsilon & \mathrm{A}^{-1}\tilde{\mathrm{E}}_\epsilon \\ & & & & & \mathrm{A}^{-1}\tilde{\mathrm{E}}_\epsilon & \mathrm{A}^{-1}\mathrm{D}_\epsilon \end{pmatrix}_{M\times M}, \tag{3.17}$$

and

$$\tilde{\mathbf{b}} = (\mathrm{A}^{-1}\mathbf{b}_0, \mathrm{A}^{-1}\mathbf{b}_1, \cdots, \mathrm{A}^{-1}\mathbf{b}_{M-1})^T. \tag{3.18}$$

We now assume that the ratio h_2/h_1 is bounded:

$$\text{(H2)} \quad h_2 \leq \kappa h_1. \tag{3.19}$$

We then notice that

$$\|\tilde{\mathrm{A}}_\epsilon\| \leq \kappa \frac{\epsilon h_1}{h_2}; \tag{3.20}$$

indeed, from the entries of the matrix $\tilde{\mathrm{A}}_\epsilon$ shown in (3.5c), using the hypotheses (H0)-(2.6), (H1)-(2.7) and (H2)-(3.19), after some elementary calculations, we find that the entries of the block $\tilde{\mathrm{A}}_\epsilon$ are majorized by $\kappa\epsilon h_1/h_2$. Furthermore, since $\tilde{\mathrm{A}}_\epsilon$ is banded, (3.20) follows from estimate (3.12). Similarly, from (3.5d)-(3.5g) we find

$$\|\tilde{\mathrm{B}}_\epsilon\|,\ \|\tilde{\mathrm{C}}_\epsilon\|,\ \|\tilde{\mathrm{E}}_\epsilon\|,\ \|\mathrm{D}_\epsilon\| \leq \kappa \frac{\epsilon h_1}{h_2}. \tag{3.21}$$

From (3.3) and (3.5b), we find

$$\|\mathrm{A}^{-1}\| \leq \frac{\kappa}{h_2}\|\mathrm{S}^{-1}\| \leq \frac{\kappa}{h_2}. \tag{3.22}$$

Using estimates (3.20), we then find

$$\|\mathrm{A}^{-1}\tilde{\mathrm{A}}_\epsilon\| \leq \|\mathrm{A}^{-1}\|\|\tilde{\mathrm{A}}_\epsilon\| \leq \kappa \frac{\epsilon h_1}{h_2^2}, \tag{3.23}$$

and similarly, by (3.21), we find

$$\|\mathrm{A}^{-1}\tilde{\mathrm{B}}_\epsilon\|,\ \|\mathrm{A}^{-1}\tilde{\mathrm{C}}_\epsilon\|,\ \|\mathrm{A}^{-1}\tilde{\mathrm{E}}_\epsilon\|,\ \|\mathrm{A}^{-1}\mathrm{D}_\epsilon\| \leq \kappa \frac{\epsilon h_1}{h_2^2}. \tag{3.24}$$

Hence, from Lemma 3.1 and (3.17), since Λ_ϵ is banded, we easily find

$$\|\Lambda_\epsilon\| \leq \kappa \frac{\epsilon h_1}{h_2^2}, \tag{3.25}$$

and it is not hard to see that

$$\|\tilde{\mathbf{b}}\|_2 \le \|\mathrm{A}^{-1}\|\|\mathbf{b}\|_2 \le \frac{\kappa}{h_2}\|\mathbf{b}\|_2. \tag{3.26}$$

Taking the norm of each side of equation (3.15), we find

$$\|(\Lambda_0 + \Lambda_\epsilon)\mathbf{a}\|_2 = \|\tilde{\mathbf{b}}\|_2 \le \frac{\kappa}{h_2}\|\mathbf{b}\|_2. \tag{3.27}$$

We are now able to estimate the norm $\|\mathbf{a}\|_2$ as follows. Firstly,

$$\begin{aligned} \|\Lambda_0\mathbf{a}\|_2 \le \|(\Lambda_0 + \Lambda_\epsilon)\mathbf{a}\|_2 + \| - \Lambda_\epsilon\mathbf{a}\|_2 \ &\le \frac{\kappa}{h_2}\|\mathbf{b}\|_2 + \|\Lambda_\epsilon\|\|\Lambda_0^{-1}\|\|\Lambda_0\mathbf{a}\|_2 \\ \le \text{(by (3.25) and Lemma 3.2 below)} \ &\le \frac{\kappa}{h_2}\|\mathbf{b}\|_2 + \kappa_1\frac{\epsilon}{h_2^2}\|\Lambda_0\mathbf{a}\|_2. \end{aligned} \tag{3.28}$$

Note here that we named the constant κ_1 and we now assume that

$$\text{(H3)} \quad \kappa_1\frac{\epsilon}{h_2^2} \le \frac{1}{2}, \left(or\ \epsilon \le \frac{h_2^2}{2\kappa_1}\right). \tag{3.29}$$

We then deduce from (3.28) that

$$\|\Lambda_0\mathbf{a}\|_2 \le \frac{\kappa}{h_2}\|\mathbf{b}\|_2, \tag{3.30}$$

and hence

$$\|\mathbf{a}\|_2 \le \|\Lambda_0^{-1}\|\|\Lambda_0\mathbf{a}\|_2 \le \frac{\kappa}{h_1 h_2}\|\mathbf{b}\|_2. \tag{3.31}$$

We now justify the estimate of the norm of Λ_0^{-1} and derive the relations between $\|\mathbf{a}\|_2$ and $|u_h|_{L^2}$, and between $\|\mathbf{b}\|_2$ and $|f|_{L^2}$ in the subsequent lemmas.

Lemma 3.2. *The inverse Λ_0^{-1} of Λ_0 is given by formulas (3.33) below and we have:*

$$\|\Lambda_0^{-1}\| \le \frac{\kappa}{h_1}. \tag{3.32}$$

Proof. The inverse matrix Λ_0^{-1} can be found recursively as follows. Set

$$\Xi(1) = I,\ \Xi(2) = \begin{pmatrix} 0 & I \\ -I & I \end{pmatrix}. \tag{3.33a}$$

Then we claim that

$$\Lambda_0^{-1} = \Xi(M) = \begin{pmatrix} \Xi(M-2) & \Xi_1 \\ \Xi_3 & \Xi_2 \end{pmatrix}, \tag{3.33b}$$

where

$$\Xi_1 = \begin{pmatrix} 0 & 0 & 0 & \cdot & 0 & 0 & 0 \\ I & I & I & \cdot & I & I & I \end{pmatrix}^T_{2\times(M-2)},\ \Xi_2 = \begin{pmatrix} 0 & I \\ -I & I \end{pmatrix}_{2\times 2}, \tag{3.33c}$$

and for $M = 2m$ and $M = 2m+1$, respectively,

$$\begin{aligned} \Xi_3 &= \begin{pmatrix} 0 & 0 & 0 & \cdot & 0 & 0 & 0 \\ -I & I & -I & \cdot & I & -I & I \end{pmatrix}_{2\times(2m-2)}, \\ \Xi_3 &= \begin{pmatrix} 0 & 0 & 0 & \cdot & 0 & 0 & 0 \\ I & -I & I & \cdot & I & -I & I \end{pmatrix}_{2\times(2m-1)}; \end{aligned} \tag{3.33d}$$

note that the minus sign is alternating[1]. Then since the entries of Λ_0^{-1} are 0 or $\pm I$ blocks, the estimate (3.32) follows from (3.13).

To prove that the matrix $\Xi(M)$ in (3.33) is indeed the inverse of the matrix Λ_0, we first consider the case $M = 2m$ and use an induction on m (the case $M = 2m+1$ can be done similarly). For $m = 1$, we easily verify that $\Xi(2)$ is the inverse of Λ_0. Suppose that for $M = 2m$, $m \geq 1$, $\Xi(M)$ is the inverse of Λ_0. We then verify that for $M = 2(m+1)$, $\Xi(M)$ is the inverse of Λ_0: indeed we rewrite Λ_0 in (3.16) in the form:

$$\Lambda_0 = \Lambda(M) = \begin{pmatrix} \Lambda(M-2) & \Lambda_1 \\ \Lambda_3 & \Lambda_2 \end{pmatrix}_{M\times M}, \tag{3.34}$$

where

$$\Lambda(M-2) = \begin{pmatrix} I & -I & & & \\ I & 0 & -I & & \\ & I & 0 & \cdot & \\ & & \cdot & \cdot & -I \\ & & & I & 0 \end{pmatrix}_{(M-2)\times(M-2)}, \tag{3.35}$$

and

$$\Lambda_1 = \begin{pmatrix} 0 & 0 & 0 & \cdot & 0 & 0 & -I \\ 0 & 0 & 0 & \cdot & 0 & 0 & 0 \end{pmatrix}^T_{2\times(M-2)}, \tag{3.36}$$

$$\Lambda_2 = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}_{2\times 2}, \ \Lambda_3 = \begin{pmatrix} 0 & 0 & 0 & \cdot & 0 & 0 & I \\ 0 & 0 & 0 & \cdot & 0 & 0 & 0 \end{pmatrix}_{2\times(M-2)}. \tag{3.37}$$

Then by explicit calculations, we find that

$$\Lambda_0\Xi(M) = \begin{pmatrix} \Lambda(M-2)\Xi(M-2) + \Lambda_1\Xi_3 & \Lambda(M-2)\Xi_1 + \Lambda_1\Xi_2 \\ \Lambda_3\Xi(M-2) + \Lambda_2\Xi_3 & \Lambda_3\Xi_1 + \Lambda_2\Xi_2 \end{pmatrix}_{M\times M} = \tilde{I}. \tag{3.38}$$

Indeed, by our assumption, we find that $\Lambda(M-2)\Xi(M-2) = \tilde{I}_{(M-2)\times(M-2)}$ and the other entries are computed explicitly. $\square$

Lemma 3.3. *For* $\mathbf{a}$, $\mathbf{b}$, u_h *as in (3.1)-(3.2), the following relations hold:*

$$|u_h|_{L^2} \leq \kappa(h_1h_2)^{1/2}\|\mathbf{a}\|_2, \tag{3.39a}$$

$$\|\mathbf{b}\|_2 \leq \kappa(h_1h_2)^{1/2}|f|_{L^2}. \tag{3.39b}$$

Proof. We easily verify that, by the compact supports of the elements ϕ_i and ψ_j,

$$\begin{aligned} |u_h|^2_{L^2} &= \int_\Omega \left(\sum_{i=0}^{M-1}\sum_{j=1}^{N-1} a_{ij}\phi_i\psi_j\right)^2 d\Omega = \sum_{i=0}^{M-1}\sum_{j=1}^{N-1} \sum_{\substack{k\in\{i-1,i,i+1\},\\ l\in\{j-1,j,j+1\}}} a_{ij}a_{kl}\cdot \\ &\cdot \int_0^1 \phi_i\phi_k dx \int_0^1 \psi_j\psi_l dy \ \leq \kappa h_1h_2 \sum_{i=0}^{M-1}\sum_{j=1}^{N-1} a_{ij}^2, \end{aligned} \tag{3.40}$$

[1]For example, for $m = 4$,

$$\Lambda_0^{-1} = \begin{pmatrix} 0 & I & 0 & I & 0 & I & 0 & I \\ -I & I & 0 & I & 0 & I & 0 & I \\ 0 & 0 & 0 & I & 0 & I & 0 & I \\ -I & I & -I & I & 0 & I & 0 & I \\ 0 & 0 & 0 & 0 & 0 & I & 0 & I \\ -I & I & -I & I & -I & I & 0 & I \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & I \\ -I & I & -I & I & -I & I & -I & I \end{pmatrix}_{8\times 8}, \quad \begin{pmatrix} I & 0 & I & 0 & I & 0 & I & 0 & I \\ 0 & 0 & I & 0 & I & 0 & I & 0 & I \\ I & -I & I & 0 & I & 0 & I & 0 & I \\ 0 & 0 & 0 & 0 & I & 0 & I & 0 & I \\ I & -I & I & -I & I & 0 & I & 0 & I \\ 0 & 0 & 0 & 0 & 0 & 0 & I & 0 & I \\ I & -I & I & -I & I & -I & I & 0 & I \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & I \\ I & -I & I & -I & I & -I & I & -I & I \end{pmatrix}_{9\times 9}.$$

and this is exactly (3.39a). To verify (3.39b), we notice that

(3.41)
$$\mathbf{b} = \left(\int_{\Omega_{01}} f\phi_0\psi_1 d\Omega, \cdots, \int_{\Omega_{kl}} f\phi_k\psi_l d\Omega, \cdots, \int_{\Omega_{M-1,N-1}} f\phi_{M-1}\psi_{N-1} d\Omega \right)^T,$$

where the Ω_{kl} are the compact supports of the elements $\phi_k\psi_l$, more precisely,

$$\Omega_{0l} = (0, h_1) \times ((l-1)h_2, (l+1)h_2) \text{ for } k = 0,$$
$$\Omega_{kl} = ((k-1)h_1, (k+1)h_1) \times ((l-1)h_2, (l+1)h_2) \text{ for } k \geq 1.$$

We then find by the Cauchy-Schwarz inequality that

(3.42)
$$\begin{aligned} \|\mathbf{b}\|_2^2 &= \sum_{k=0}^{M-1} \sum_{l=1}^{N-1} \left(\int_{\Omega_{kl}} f\phi_k\psi_l d\Omega \right)^2 \\ &\leq \sum_{k=0}^{M-1} \sum_{l=1}^{N-1} \int_{\Omega_{kl}} f^2 d\Omega \int_{\Omega_{kl}} (\phi_k\psi_l)^2 d\Omega \leq \kappa h_1 h_2 \int_{\Omega} f^2 d\Omega; \end{aligned}$$

hence, (3.39b) follows. □

Using the estimate (3.31) and Lemma 3.3, we can directly deduce the following theorem.

Theorem 3.1. *Assume that (H0)-(H3) hold, that is [(2.6), (2.7), (3.19), (3.29)]. Let u_h be the solution of problem (1.4) with (2.9). Then for any data $f = f(x, y) \in L^2(\Omega)$ (not necessarily satisfying (2.1)), there exists a constant $\kappa > 0$ independent of ϵ, h_1, and h_2 such that*

(3.43)
$$|u_h|_{L^2(\Omega)} \leq \kappa |f|_{L^2(\Omega)}.$$

Remark 3.2. For a classical scheme not using a BLE, the stiffness matrix Γ_ϵ is as in (3.5a), after deleting the first row and the first column of the matrix in (3.5a). Hence, since from (3.6), as $\epsilon \to 0$, $\mathrm{D}_\epsilon \to 0$, $\mathrm{E}_\epsilon \to -\mathrm{A}$ and $\mathrm{F}_\epsilon \to \mathrm{A}$, it is obvious that the system tends to a singular system. On the other hand, for the new scheme (1.4) with (2.9), the entries A_ϵ, B_ϵ, and C_ϵ in Γ_ϵ stabilize our system as we have seen in this section, i.e. the BLE ϕ_0 absorbs the singularity due to the small ϵ of the linear system (3.2).

3.2. Case $\kappa h_2^2 \leq \epsilon \leq \kappa h_2$. If we assume that

(3.44)
$$\text{(H4)} \quad \kappa_2 h_2^2 \leq \epsilon \leq \kappa h_2, \ \kappa_2 = 1/(2\kappa_1),$$

we easily see that we cannot derive (3.30) from (3.28). We will need some more delicate analysis which we introduce in this section; in particular we need to investigate more carefully the BLE ϕ_0 introduced in (2.3).

To obtain the L^2- stability in this range of values of ϵ, we will utilize quasi-uniform elements, namely we assume

(3.45)
$$\text{(H5)} \quad \sqrt{2} h_1 \leq h_2 \leq \kappa h_1;$$

the $\sqrt{2}$ will be justified later; note that (H5)-(3.45) implies (H2)-(3.19).

We first derive in Lemma 3.4 below, a Poincaré-like inequality for any v_h^{BL}. For $v_h \in V_h$, we write $v_h = v_h^{BL} + v_h^{LI}$, where

(3.46)
$$v_h^{BL} = \sum_{j=1}^{N-1} a_{0j}\phi_0\psi_j, \ v_h^{LI} = \sum_{i=1}^{M-1} \sum_{j=1}^{N-1} a_{ij}\phi_i\psi_j,$$

and we have:

Lemma 3.4. *Assume only that (H0)-(2.6), (H1)-(2.7) hold. Then there exists a positive constant κ independent of ϵ, h_1, and h_2 such that, for any $v_h \in V_h$,*

$$|v_h^{BL}|_{L^2(\Omega)} \le \kappa \epsilon^{1/2} h_1^{1/2} |v_h|_{H^1(\Omega)}. \tag{3.47}$$

Proof. Firstly, we notice that $Mh_1 = Nh_2 = 1$, and due to the boundary conditions,

$$a_{M,l} = a_{k,0} = a_{k,N} = 0, \tag{3.48}$$

and by explicit calculations,

$$\int_{(l-1)h_2}^{lh_2} \psi_l^2 dy = \int_{(l-1)h_2}^{lh_2} \psi_{l-1}^2 dy = 2\int_{(l-1)h_2}^{lh_2} \psi_l \psi_{l-1} dy = \frac{h_2}{3}. \tag{3.49}$$

We now derive the estimates (3.47) as follows. Taking into consideration the supports of the elements ϕ_k and ψ_l, we see that

(3.50)

$$\int_\Omega \left(\frac{\partial v_h}{\partial x}\right)^2 d\Omega = \sum_{k=1}^{M}\sum_{l=1}^{N} \int_{\tilde{\Omega}_{kl}} \left(\frac{\partial v_h}{\partial x}\right)^2 d\Omega = \sum_{k=1}^{M}\sum_{l=1}^{N} \int_{\tilde{\Omega}_{kl}} \left(\psi_l \frac{d\tilde{\phi}_{k,l}}{dx} + \psi_{l-1}\frac{d\tilde{\phi}_{k,l-1}}{dx}\right)^2 d\Omega,$$

where

$$\tilde{\Omega}_{kl} = ((k-1)h_1, kh_1) \times ((l-1)h_2, lh_2), \tag{3.51}$$

$$\tilde{\phi}_{k,j} = a_{k,j}\phi_k + a_{k-1,j}\phi_{k-1},\ j = l,\ l-1. \tag{3.52}$$

We find from (3.49) that

$$\begin{aligned} &\int_{\tilde{\Omega}_{kl}} \left(\psi_l \frac{d\tilde{\phi}_{k,l}}{dx} + \psi_{l-1}\frac{d\tilde{\phi}_{k,l-1}}{dx}\right)^2 d\Omega \\ &= \frac{h_2}{3}\int_{(k-1)h_1}^{kh_1} \left\{ \left(\frac{d\tilde{\phi}_{k,l}}{dx}\right)^2 + \frac{d\tilde{\phi}_{k,l}}{dx}\frac{d\tilde{\phi}_{k,l-1}}{dx} + \left(\frac{d\tilde{\phi}_{k,l-1}}{dx}\right)^2 \right\} dx \\ &\ge I_{kl} + I_{k,l-1}, \quad (\text{using } a^2 + ab + b^2 \ge (a^2+b^2)/2), \end{aligned} \tag{3.53}$$

where

$$I_{kj} = \frac{h_2}{6}\int_{(k-1)h_1}^{kh_1} \left(\frac{d\tilde{\phi}_{k,l}}{dx}\right)^2 dx,\ j = l,\ l-1. \tag{3.54}$$

For $k = 1$, since $\int_0^{h_1} (d\phi_0/dx)dx = \phi_0(h_1) - \phi_0(0) = 0$,

$$\int_0^{h_1} \left(\frac{d\tilde{\phi}_{1,l}}{dx}\right)^2 dx = \frac{a_{1,l}^2}{h_1} + a_{0,l}^2 \int_0^{h_1}\left(\frac{d\phi_0}{dx}\right)^2 dx \ge a_{0,l}^2 \frac{\xi_1 - 2\epsilon + h_1}{2h_1\epsilon}. \tag{3.55}$$

We then notice from (H0)-(2.6) that

$$\xi_1 = \{4\epsilon - (2\epsilon + h_1)e^{-h_1/\epsilon}\}e^{-h_1/\epsilon} \ge (4 - 2\epsilon - h_1)\epsilon e^{-h_1/\epsilon} \ge 0, \tag{3.56}$$

and thus using the fact that $2\epsilon \le h_1/2$ from (H1)-(2.7) we find that

$$I_{1,l} = \frac{h_2}{6}\int_0^{h_1}\left(\frac{d\tilde{\phi}_{1,l}}{dx}\right)^2 dx \ge \frac{a_{0,l}^2 h_2(-2\epsilon + h_1)}{12h_1\epsilon} \ge a_{0,l}^2 \frac{h_2}{24\epsilon}. \tag{3.57}$$

For $k \ge 2$, we observe that

$$I_{k,l} = \frac{h_2}{6}\int_{(k-1)h_1}^{kh_1} \left(\frac{d\tilde{\phi}_{k,l}}{dx}\right)^2 dx = \frac{h_2(a_{k,l} - a_{k-1,l})^2}{6h_1}. \tag{3.58}$$

Now using (3.50), (3.53), (3.57) and the positivity of the $I_{k,l}$, we find that

$$\int_\Omega \left(\frac{\partial v_h}{\partial x}\right)^2 d\Omega \geq \sum_{k=1}^{M}\sum_{l=1}^{N}\Big\{I_{kl} + I_{k,l-1}\Big\} \geq \sum_{l=1}^{N} I_{1,l} \geq \frac{h_2}{24\epsilon}\sum_{l=1}^{N} a_{0,l}^2, \tag{3.59}$$

and thus

$$h_2 \sum_{j=1}^{N-1} a_{0,j}^2 \leq \kappa\epsilon |v_h|_{H^1}^2. \tag{3.60}$$

Thanks to (H0)-(2.6), (H1)-(2.7), we easily verify that $\int_0^1 \phi_0^2 dx \leq \kappa h_1$ and we thus have

$$\begin{aligned} |v_h^{BL}|_{L^2}^2 &= \int_\Omega \left(\sum_{j=1}^{N-1} a_{0j}\phi_0\psi_j\right)^2 d\Omega = \sum_{j=1}^{N-1}\sum_{l\in\{j-1,j,j+1\}} a_{0j}a_{0l}\cdot \\ &\cdot \int_0^1 \psi_j\psi_l dy \int_0^1 \phi_0^2 dx \leq \kappa h_1 h_2 \sum_{j=1}^{N-1} a_{0j}^2 \leq (\text{by } (3.60)) \leq \kappa h_1 \epsilon |v_h|_{H^1}^2, \end{aligned} \tag{3.61}$$

and the lemma follows. □

Remark 3.3. Later on we will use the following inequality: from (3.58) and (3.59) with $v_h = u_h$, where u_h is the solution of equation (1.4) with (2.9), we can write

$$\frac{h_2}{6h_1}\sum_{k=2}^{M} \|\mathbf{a}_{k-1} - \mathbf{a}_k\|_2^2 \leq \sum_{k=2}^{M}\sum_{l=1}^{N} I_{k,l} \leq \int_\Omega \left(\frac{\partial u_h}{\partial x}\right)^2 d\Omega. \tag{3.62}$$

We now estimate $\|\mathbf{a}\|_2$ to obtain the upper bound of $|u_h|_{L^2}$ as indicated in Lemma 3.3. For that purpose we write the system (3.2) in the more explicit form:

$$\mathrm{A}_\epsilon \mathbf{a}_0 + \mathrm{B}_\epsilon \mathbf{a}_1 = \mathbf{b}_0, \tag{3.63a}$$

$$\mathrm{C}_\epsilon \mathbf{a}_0 + \mathrm{D}_\epsilon \mathbf{a}_1 + \mathrm{E}_\epsilon \mathbf{a}_2 = \mathbf{b}_1, \tag{3.63b}$$

$$\mathrm{F}_\epsilon \mathbf{a}_{i-2} + \mathrm{D}_\epsilon \mathbf{a}_{i-1} + \mathrm{E}_\epsilon \mathbf{a}_i = \mathbf{b}_{i-1}, \text{ for } i = 3, \cdots, M-1, \tag{3.63c}$$

$$\mathrm{F}_\epsilon \mathbf{a}_{M-2} + \mathrm{D}_\epsilon \mathbf{a}_{M-1} = \mathbf{b}_{M-1}. \tag{3.63d}$$

Using (3.5) and setting $\mathbf{a}_M = 0$, we rewrite (3.63c) and (3.63d): for $k = 3, \cdots, M$,

$$(\mathrm{A} + \tilde{\mathrm{E}}_\epsilon)\mathbf{a}_{k-2} + \mathrm{D}_\epsilon \mathbf{a}_{k-1} + (-\mathrm{A} + \tilde{\mathrm{E}}_\epsilon)\mathbf{a}_k = \mathbf{b}_{k-1}. \tag{3.64}$$

Taking the inner product of (3.64) with $\mathbf{a}_{k-1}$, using the symmetry of A and $\tilde{\mathrm{E}}_\epsilon$, and summing over $k = i, \cdots, M$, $i \geq 3$, we find after some elementary calculations:

$$< (\mathrm{A} + \tilde{\mathrm{E}}_\epsilon)\mathbf{a}_{i-2}, \mathbf{a}_{i-1} > + J = \sum_{k=i}^{M} < \mathbf{b}_{k-1}, \mathbf{a}_{k-1} >, \tag{3.65a}$$

where

$$J = \sum_{k=i}^{M-1} < 2\tilde{\mathrm{E}}_\epsilon \mathbf{a}_{k-1}, \mathbf{a}_k > + \sum_{k=i}^{M} < \mathrm{D}_\epsilon \mathbf{a}_{k-1}, \mathbf{a}_{k-1} > . \tag{3.65b}$$

We then claim that $J \geq 0$. We firstly notice that $2\tilde{\mathrm{E}}_\epsilon + \mathrm{D}_\epsilon = \epsilon h_1/h_2 \mathrm{U}$, and thanks to the Gershgorin circle theorem, we find that the eigenvalues of U are nonnegative. Hence, $< \mathrm{U}\mathbf{a}_{k-1}, \mathbf{a}_{k-1} > \geq 0$, and

$$J \geq -2\sum_{k=i}^{M-1} < \mathrm{G}\mathbf{a}_{k-1}, \mathbf{a}_k > + 2\sum_{k=i}^{M} < \mathrm{G}\mathbf{a}_{k-1}, \mathbf{a}_{k-1} >, \tag{3.66}$$

where $\mathrm{G} = -\tilde{\mathrm{E}}_\epsilon$. The quasi-uniform mesh hypothesis shows that G is positive semidefinite. Indeed, from (3.5f), we have $\mathrm{G} = -\tilde{\mathrm{E}}_\epsilon = \epsilon\,(h_2/6h_1)\,\mathrm{S} - \epsilon\,(h_1/6h_2)\,\mathrm{U}$. Then its Gershgorin discs belong to

$$\mathcal{G} = \left\{ z \in \mathbb{C}; \left| z - \frac{\epsilon}{3}\left(\frac{2h_2}{h_1} - \frac{h_1}{h_2}\right)\right| \leq \frac{\epsilon}{3}\left(\frac{h_2}{h_1} + \frac{h_1}{h_2}\right)\right\}. \tag{3.67}$$

From (H5)-(3.45) we find that $h_2/h_1 - 2h_1/h_2 \geq 0$ which guarantees that the Gershgorin discs belong to $\mathbb{C}$ with nonnegative real parts. Since G is symmetric and thus its eigenvalues are real numbers, all eigenvalues of G are nonnegative. By the spectral property of G, we then write $< \mathrm{G}\xi, \eta > = < \mathrm{G}^{1/2}\xi, \mathrm{G}^{1/2}\eta >$ and hence we rewrite (3.66):

(3.68)

$$J \geq \sum_{k=i}^{M-1} \|\mathrm{G}^{1/2}\mathbf{a}_{k-1} - \mathrm{G}^{1/2}\mathbf{a}_k\|_2^2 + < \mathrm{G}\mathbf{a}_{i-1}, \mathbf{a}_{i-1} > + < \mathrm{G}\mathbf{a}_{M-1}, \mathbf{a}_{M-1} > \geq 0.$$

Hence from (3.65) we find that

$$< (\mathrm{A} + \tilde{\mathrm{E}}_\epsilon)\mathbf{a}_{i-2}, \mathbf{a}_{i-1} > \leq \sum_{k=i}^{M} < \mathbf{b}_{k-1}, \mathbf{a}_{k-1} > . \tag{3.69}$$

Since $< \mathrm{U}\xi, \xi > \geq 0$ and $< \mathrm{S}\xi, \xi > \geq 2\|\xi\|_2^2$, we find that by (H1)-(2.7)

$$< (\mathrm{A} + \tilde{\mathrm{E}}_\epsilon)\xi, \xi > = \frac{h_2}{12}\left(1 - \frac{2\epsilon}{h_1}\right) < \mathrm{S}\xi, \xi > + \epsilon\frac{h_1}{6h_1} < \mathrm{U}\xi, \xi > \geq \frac{h_2}{12}\|\xi\|_2^2, \tag{3.70}$$

and from (3.69) and the fact that $\|\mathrm{A} + \tilde{\mathrm{E}}_\epsilon\|_2 \leq \kappa h_2$ we find

$$\begin{aligned} \frac{h_2}{12}\|\mathbf{a}_{i-2}\|_2^2 &\leq < (\mathrm{A} + \tilde{\mathrm{E}}_\epsilon)\mathbf{a}_{i-2}, \mathbf{a}_{i-2} > \\ &\leq \sum_{k=i}^{M} < \mathbf{b}_{k-1}, \mathbf{a}_{k-1} > \ + < (\mathrm{A} + \tilde{\mathrm{E}}_\epsilon)\mathbf{a}_{i-2}, \mathbf{a}_{i-2} - \mathbf{a}_{i-1} > \\ &\leq \|\mathbf{a}\|_2\|\mathbf{b}\|_2 + \frac{h_2}{24}\|\mathbf{a}_{i-2}\|_2^2 + \kappa h_2\|\mathbf{a}_{i-2} - \mathbf{a}_{i-1}\|_2^2. \end{aligned} \tag{3.71}$$

Hence, summing (3.71) over $i = 3$ to $M + 1$ and multiplying by $24h_1$, we find that

$$h_1 h_2 \sum_{i=3}^{M+1} \|\mathbf{a}_{i-2}\|_2^2 \leq \kappa h_1 M\|\mathbf{a}\|_2\|\mathbf{b}\|_2 + \kappa h_1 h_2 \sum_{i=3}^{M+1} \|\mathbf{a}_{i-2} - \mathbf{a}_{i-1}\|_2^2. \tag{3.72}$$

Thanks to (3.60), adding to (3.72) $h_1 h_2\|\mathbf{a}_0\|_2^2 \leq \kappa h_1 \epsilon |u_h|_{H^1}^2$, and since $|u_h|_{H^1}^2 \leq \epsilon^{-1}|f|_{L^2}|u_h|_{L^2}$ by letting $v = u_h$ in (1.4), we find that

$$\begin{aligned} h_1 h_2\|\mathbf{a}\|_2^2 &\leq \kappa\|\mathbf{a}\|_2\|\mathbf{b}\|_2 + \kappa h_1\epsilon|u_h|_{H^1}^2 + \kappa h_1 h_2 \sum_{i=3}^{M+1} \|\mathbf{a}_{i-2} - \mathbf{a}_{i-1}\|_2^2 \\ &\leq (\text{by } (3.62)) \leq \kappa\|\mathbf{a}\|_2\|\mathbf{b}\|_2 + \kappa(h_1^2 + h_1\epsilon)|u_h|_{H^1}^2 \\ &\leq (\text{by } (3.44),\ (3.45)) \leq \kappa\|\mathbf{a}\|_2\|\mathbf{b}\|_2 + \kappa|f|_{L^2}|u_h|_{L^2}. \end{aligned} \tag{3.73}$$

Hence from Lemma 3.3 valid in all cases:

$$|u_h|_{L^2}^2 \leq h_1 h_2\|\mathbf{a}\|_2^2 \leq \kappa h_1^{-1} h_2^{-1}\|\mathbf{b}\|_2^2 + \kappa|f|_{L^2}|u_h|_{L^2} \leq \kappa|f|_{L^2}^2 + \kappa|f|_{L^2}|u_h|_{L^2}.$$

By the Cauchy-Schwarz inequality, we thus deduce the following theorem.

Theorem 3.2. *Assume that the hypotheses (H0)-(H1), (H4)-(H5) hold, that is [(2.6), (2.7), (3.44), (3.45)]. Let u_h be the solution of problem (1.4) with (2.9). Then for any data $f = f(x,y) \in L^2(\Omega)$ (not necessarily satisfying (2.1)), there exists a constant $\kappa > 0$ independent of ϵ, h_1, and h_2 such that*

$$|u_h|_{L^2(\Omega)} \leq \kappa |f|_{L^2(\Omega)}. \tag{3.74}$$

Remark 3.4. For the problem (1.1a) with different boundary conditions, e.g. $u = 0$ at $x = 0, 1$ and $\partial u/\partial y = 0$ at $y = 0, 1$, or $u = 0$ at $x = 0, 1$ and $u(x, y) = u(x, y+1)$, which lead to a slight change of each block A_ϵ to F_ϵ, we can similarly verify Theorem 3.1 - 3.2.

4. H^1- and L^2- Approximation Errors

The following Theorem 4.1 - 4.2 give the H^1 and L^2- behavior of the convergence errors for the approximate solutions.

Theorem 4.1 (H^1- error)**.** *Assume only that (H0)-(2.6), (H1)-(2.7) hold. Let $u = u^\epsilon$ be the exact solution of (1.3), and u_h the solution of (1.4) with (2.9), and let f be smooth on $\bar{\Omega}$ satisfying (2.1). Then*

$$|u - u_h|_{H^1(\Omega)} \leq \kappa(h + h^2\epsilon^{-1}). \tag{4.1}$$

Proof. Subtracting (1.4) from (1.3), we find

$$a_\epsilon(u - u_h, v_h) = 0 \quad \text{for all } v_h \in V_h, \tag{4.2}$$

and thus for an interpolant $\tilde{u}_h \in V_h$, $a_\epsilon(u - u_h, u - u_h) = a_\epsilon(u - u_h, u - \tilde{u}_h)$. We thus find by the Cauchy-Schwarz inequality that

$$|u - u_h|_{H^1} \leq \kappa |u - \tilde{u}_h|_{H^1} + \kappa\epsilon^{-1}|u - \tilde{u}_h|_{L^2}. \tag{4.3}$$

Hence (4.1) follows from the interpolation inequalities as in Lemma 2.3. □

Theorem 4.2 (L^2- error)**.** *Assume only that (H0)-(2.6), (H1)-(2.7) hold. Let $u = u^\epsilon$ be the exact solution of (1.3), and u_h the solution of (1.4) with (2.9), and let f be smooth on $\bar{\Omega}$ satisfying (2.1). Then there exist positive constants λ and κ independent of ϵ, h_1, h_2 such that*

$$|u - u_h|_{L^2(\Omega)} \leq \kappa \begin{cases} h + h_2^2\epsilon^{-1/2} & \text{if } \epsilon \leq \lambda h_2^2,\ h_2 \leq \kappa h_1, \\ h & \text{if } \lambda h_2^2 \leq \epsilon \leq \kappa h_2,\ \sqrt{2}h_1 \leq h_2 \leq \kappa h_1. \end{cases} \tag{4.4}$$

Proof. Set $\lambda = 1/(2\kappa_1) = \kappa_2$, where $\kappa_1,\ \kappa_2$ are as in (H3)-(3.29), (H4)-(3.44). From (4.2), we have for all $v_h \in V_h$,

$$a_\epsilon(u_h - \tilde{u}_h, v_h) = a_\epsilon(u - \tilde{u}_h, v_h) = \epsilon((u - \tilde{u}_h, v_h)) - ((u - \tilde{u}_h)_x, v_h). \tag{4.5}$$

By the uniqueness of solutions, we can decompose $u_h - \tilde{u}_h$ as:

$$u_h - \tilde{u}_h = v_h^1 + v_h^2, \tag{4.6a}$$

$$a_\epsilon(v_h^1, v_h) = \epsilon((u - \tilde{u}_h, v_h)),\ \forall v_h \in V_h, \tag{4.6b}$$

$$a_\epsilon(v_h^2, v_h) = -((u - \tilde{u}_h)_x, v_h),\ \forall v_h \in V_h. \tag{4.6c}$$

From equation (4.6b) with $v_h = v_h^1$, we then easily find that

$$|v_h^1|_{H^1} \leq |u - \tilde{u}_h|_{H^1}. \tag{4.7}$$

If $\epsilon \leq \lambda h_2^2$, $h_2 \leq \kappa h_1$, then from Theorem 3.1 applied to equation (4.6c) with $f = -(u - \tilde{u}_h)_x$, we find

$$|v_h^2|_{L^2} \leq \kappa |f|_{L^2} \leq \kappa |u - \tilde{u}_h|_{H^1}. \tag{4.8}$$

The first estimate in (4.4) follows from the interpolation inequality (2.10b) observing that due to the Poincaré inequality, (4.6), (4.7) and (4.8), we have

$$\begin{aligned} |u-u_h|_{L^2} &\le |u-\tilde{u}_h|_{L^2} + |u_h-\tilde{u}_h|_{L^2} \\ &\le |u-\tilde{u}_h|_{L^2} + \kappa|v_h^1|_{H^1} + |v_h^2|_{L^2} \le \kappa|u-\tilde{u}_h|_{H^1}. \end{aligned} \tag{4.9}$$

If $\lambda h_2^2 \le \epsilon \le \kappa h_2$, $\sqrt{2}h_1 \le h_2 \le \kappa h_1$, i.e. quasi-uniform elements. The second estimate in (4.4) similarly follows from Theorem 3.2 with $f = -(u-\tilde{u}_h)_x$ applied to equation (4.6c) again:

$$|u-u_h|_{L^2} \le \kappa|u-\tilde{u}_h|_{H^1} \le \kappa(h + h_2^2\epsilon^{-1/2}) \le \kappa h. \tag{4.10}$$

□

Remark 4.1. From Theorem 4.1 and Theorem 4.2, we find that for the new scheme (1.4) with (2.9) to be effective, we should require the space mesh to be of order $h = o(\epsilon^{1/2})$ in the H^1 approximation and of order h_1 small, $h_2 = o(\epsilon^{1/4})$ in the L^2 approximation (thus $o(\epsilon^{1/4})$ in the weighted norm $\|\cdot\|_\epsilon$). These mesh restrictions come from the approximation errors in y due to $g^\epsilon(y)$ which appeared in (2.2) unlike the one dimensional example (1.7). To relax these restrictions, we can employ higher order polynomials, or to remove the restrictions, we might need a finite element space slightly different from the V_h as in (2.9) which will appear elsewhere. Extensive numerical simulations for (1.1) (with various boundary conditions) appear in [13] - [16] and elsewhere.

Acknowledgments

The authors thank the referees for many suggestions which helped improve the article and they thank Bruce Kellogg for discussions during the revision of this article. This work was supported in part by NSF grants DMS0604235, and by the Research Fund of Indiana University.

References

[1] Richard L. Burden and J.Douglas Faires, *Numerical Analysis.* Brooks/Cole, CA, 2001.

[2] W. Cai, D. Gottlieb and C.-W. Shu, *Essentially nonoscillatory spectral Fourier methods for shock wave calculations.* Math. Comp. 52 (1989), pp 389–410.

[3] W. Cheng and R. Temam, *Numerical approximation of one-dimensional stationary diffusion equations with boundary layers.* Computers and Fluids., 31(2002), pp. 453-466.

[4] W. Cheng, R. Temam and X. Wang, *New approximation algorithms for a class of partial differential equations displaying boundary layer behavior. Methods and Applications of Analysis.*, 7(2000), pp. 363-390.

[5] P. G. Ciarlet, *The finite element method for elliptic problems.* North-Holland, New York, 1978.

[6] W. Dörfler, *Uniform error estimates for an exponentially fitted finite element method for singularly perturbed elliptic equations.* SIAM J. Numer. Anal. 36 (1999), pp 1709–1738.

[7] W. Dörfler, *Uniform a priori estimates for singularly perturbed elliptic equations in multidimensions.* SIAM J. Numer. Anal. 36 (1999), pp 1878–1900.

[8] S. H. Friedberg, A. J. Insel, and L. E. Spence, *Linear Algebra.* Prentice Hall, New Jersey 07458, 1999.

[9] P. W. Hemker, *A numerical study of stiff two-point boundary value problems.* Mathematical Centre Tracts, No. 80. Mathematisch Centrum, Amsterdam, 1977.

[10] H. Han and R.B. Kellogg, *A method of enriched subspaces for the numerical solution of a parabolic singular perturbation problem.* Computational and asymptotic methods for boundary and interior layers (Dublin, 1982), pp 46–52

[11] H. Han and R.B. Kellogg, *The use of enriched subspaces for singular perturbation problems.* Proceedings of the China-France symposium on finite element methods (Beijing, 1982), 293–305, Science Press, Beijing, 1983.

[12] T. Y. Hou and X.-H. Wu *A multiscale finite element method for elliptic problems in composite materials and porous media.* J. Comput. Phys. 134(1997), pp 169-189.

[13] C. Jung, *Numerical approximation of two-dimensional convection-diffusion equations with boundary layers.* Numer. Methods Partial Differential Equations, 21(2005), pp. 623-648.

[14] C. Jung, *Numerical approximation of convection-diffusion equations in a channel using boundary layer elements.* Appl. Numer. Math. 56(2006), pp 756-777.

[15] C. Jung and R. Temam, *Numerical approximation of two-dimensional convection-diffusion equations with multiple boundary layers.* Internat. J. Numer. Analysis and Modeling. 2(2005), pp 367-408.

[16] C. Jung and R. Temam, *On parabolic boundary layers for convection-diffusion equations in a channel: Analysis and Numerical applications.* J. Sci. Comput. 28(2006), pp 361-410.

[17] J. L. Lions, *Perturbations singulières dans les problèmes aux limites et en contrôle optimal.* (French) Lecture Notes in Mathematics, Vol. 323. Springer-Verlag, Berlin-New York, 1973.

[18] 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), pp 47–62.

[19] R. Plato, *Concise Numerical Mathematics.* AMS, Providence, Rhode Island, 2003.

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

[21] M. H. Schultz, *Spline Analysis*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1973.

[22] M. Stynes, *Steady-state convection-diffusion problems.* Acta Numerica, Cambridge University Press, Cambridge, 2005, pp 445-508.

[23] S. Shih and R. B. Kellogg, *Asymptotic analysis of a singular perturbation problem.* Siam J. Math. Anal. 18(1987), pp 1467-1511.

[24] R. Temam, *Infinite dimensional dynamical systems in mechanics and physics*, Applied Mathematical Sciences Series, Vol. 68, Springer-Verlag, New-York, 1997.

The Institute for Scientific Computing and Applied Mathematics, Indiana University, Bloomington, IN 47405, USA

E-mail: `chajung@indiana.edu` and `temam@indiana.edu`

URL: `http://mypage.iu.edu/~chajung/` and `http://mypage.iu.edu/~temam/`