

UNIFORM CONVERGENCE OF THE LDG METHOD FOR A SINGULARLY PERTURBED PROBLEM WITH THE EXPONENTIAL BOUNDARY LAYER

HUIQING ZHU AND ZHIMIN ZHANG

ABSTRACT. In this paper, we study a uniform convergence property of the local discontinuous Galerkin method (LDG) for a convection-diffusion problem whose solution has exponential boundary layers. A Shishkin mesh is employed. The trial functions in the LDG method are piecewise polynomials that lies in the space \mathcal{Q}_k , i.e., are tensor product polynomials of degree at most k in one variable, where $k > 0$. We identify that the error of the LDG solution in a DG-norm converges at a rate of $(\ln N/N)^{k+1/2}$; here the total number of mesh points is $O(N^2)$. The numerical experiments show that this rate of convergence is sharp.

1. INTRODUCTION

We consider a singularly perturbed convection-diffusion problem of the form

$$(1.1) \quad \begin{aligned} -\epsilon \Delta u + \boldsymbol{\beta} \cdot \nabla u + cu &= f, & \text{in } \Omega = (0, 1)^2, \\ u &= 0, & \text{on } \partial\Omega, \end{aligned}$$

where $0 < \epsilon \ll 1$ is the diffusion parameter, the convection coefficient $\boldsymbol{\beta}(x, y) > (\beta_1, \beta_2) > (0, 0)$ on $\overline{\Omega}$, and the reaction coefficient $c(x, y) \geq 0$ on $\overline{\Omega}$. The function $f(x, y)$ is a given source term. We assume that $\boldsymbol{\beta}$, c , and f are sufficiently smooth on $\overline{\Omega}$ and

$$(1.2) \quad c - \frac{1}{2} \operatorname{div} \boldsymbol{\beta} \geq c_0 > 0, \quad \text{on } \overline{\Omega},$$

for some constant c_0 . These hypotheses guarantee that (1.1) has a unique solution in $H^2(\Omega) \cap H_0^1(\Omega)$ for any $f \in L^2(\Omega)$. Note that when ϵ is sufficiently small, in which we are interested, the other hypotheses imply that condition (1.2) can always be ensured by a change of variable $v(x, y) = e^{-(x+y)t}u(x, y)$ with a positive constant t that satisfies

$$c(x, y) + (\beta_1(x, y) + \beta_2(x, y))t - 2\epsilon t^2 - \frac{1}{2} \operatorname{div} \boldsymbol{\beta}(x, y) \geq c_0 \geq 0.$$

It is well known that when ϵ is small, the convection plays a dominant role and the solution u of (1.1) usually exhibits boundary layers of width $O(\epsilon \ln \frac{1}{\epsilon})$ at the outflow boundary of Ω (cf. [15, 16]). When $\epsilon \rightarrow 0$, the standard finite element

Received by the editor April 2, 2011 and, in revised form, June 15, 2012.

2010 *Mathematics Subject Classification*. Primary 65N30, 65N12, 65N15.

Key words and phrases. Singularly perturbed, discontinuous Galerkin method, Shishkin mesh, uniform convergence.

Corresponding author: Huiqing Zhu, Huiqing.Zhu@usm.edu.

The second author was supported in part by the US National Science Foundation through grant DMS-1115530.

©2013 American Mathematical Society
Reverts to public domain 28 years from publication

method fails to produce an accurate numerical solution unless the mesh size is comparable or smaller than the singular perturbation parameter ϵ .

There has been significant development in the design and analysis of numerical schemes for convection-dominated problems [9, 10, 16]). Among these methods, the parameter-uniform numerical methods employing layer-adapted meshes, such as a Shishkin or a Bakhvalov mesh, ensure a uniform convergence for the singularly perturbed convection-dominated problems. Here is a list of references related to our current work: [13, 15, 16, 18–20, 23, 24].

On the finite element method (FEM), an optimal error bound of order $O(\ln N/N)$ in an ϵ -weighted norm was established in [18] for linear or bilinear finite element methods with the Shishkin mesh applied to singularly perturbed problems, where the total number of mesh points is $O(N^2)$. A superconvergence rate $O((\ln N/N)^2)$ was obtained in a discrete ϵ -weighted energy norm for bilinear FEM on the Shishkin mesh in [23]. For higher order FEM, in [19], a convergence rate $O((\ln N/N)^{k+1/2})$ in the L^2 -norm was proved for a postprocessed numerical solution of the streamline-diffusion FEM, that used the tensor product approximation space \mathcal{Q}_k .

On the discontinuous Galerkin (DG) method, Castillo et al. [4] proved the first a priori error analysis of the LDG method for purely elliptic problems. It was shown that the orders of convergence of the L^2 error in ∇u and u are k and $k + 1/2$, respectively, when the solution is smooth and polynomials of degree at most k are used. On the other hand, Castillo [3] and then Castillo et al. [5] proved a superconvergence of order $k + 1$ in the energy norm for one-dimensional transient convection–diffusion problems. Later on, Cockburn et al. extended these results to the local discontinuous Galerkin (LDG) method for multidimensional elliptic problems on Cartesian grids and proved orders of superconvergence of $k + 1/2$ and $k + 1$ for $\mathbf{q} = \nabla u$ and u respectively (cf. [8]).

As for ϵ -uniform DG methods, a uniform convergence of order $O((\ln N/N)^{k+1})$ was proved for the LDG method applied to the singularly perturbed problems in a one-dimensional setting (cf. [26]). For two-dimensional problems, Zarin and Roos investigated a nonsymmetric DG method with internal penalties (NIPG method) in [22] and proved that the approximations to the solution and its gradient converge at a rate of $O((\ln N)^{1.5}/N)$ on the Shishkin mesh when bilinear polynomials are used. Under similar assumptions, a convergence rate $O(\ln N/N)$ of the LDG approximation to the solution and its gradient is established in [25]. Moreover, it was numerically observed in [20] that the LDG method superconverges uniformly at nodes of the Shishkin mesh.

However, we have not seen any theoretical results of ϵ -uniform convergence of higher-order DG methods. In this paper, we analyze the LDG method with the tensor-product polynomial space \mathcal{Q}_k on Shishkin mesh. The same regularity assumptions as in [19] are used. We use the mixed formulation of the LDG method as in [8] to discretize the diffusion term. The convection term is handled by the upwinding scheme. We denote by (\mathbf{q}_N, u_N) the LDG approximation of $(\mathbf{q}, u) = (\epsilon \nabla u, u)$. If the numerical fluxes are properly chosen, we establish a uniform error bound

$$|(\mathbf{q} - \mathbf{q}_N, u - u_N)|_{\mathcal{A}} \leq C (\ln N/N)^{k+\frac{1}{2}}.$$

Here $|\cdot|_{\mathcal{A}}$ is a DG-version of the ϵ -weighted energy semi-norm. The difference of this result from that of [19] is that: (1) we obtain the same error bound in approximating both the potential u and flux \mathbf{q} ; and (2) we do not need the postprocessing to obtain this rate for the approximation of u .

The outline of the rest of the paper is as follows. In Section 2, we display the LDG formulation and define the Shishkin mesh. Our main results are presented in Section 3. The proof of these results are provided in Section 4. Section 5 presents numerical experiments that confirm our theoretical results. We end in Section 6 with some concluding remarks.

Throughout the paper, generic positive constants C, C_i are independent of N and ϵ . Constant C can take on different values in the argument. For any measurable subdomain $D \subseteq \Omega$, we use the standard Sobolev spaces $L^2(D), H_0^1(D), H^s(D) = W_2^s(D)$ for some nonnegative integer s , etc. Since the Shishkin mesh \mathcal{T}_N contains highly anisotropic elements, we use subscript ‘ N ’ instead of ‘ h ’.

2. THE LDG METHOD ON A SHISHKIN MESH

For convenience we list the most commonly used notations (see Figure 1). Let $\mathcal{Q}_k(K)$ be the tensor products polynomial space of degree at most k in one variable with $k > 0$. We define \mathcal{T}_N , a rectangular decomposition of domain Ω , and denote by $H^\ell(\mathcal{T}_N)$ ($\ell \geq 0$) the space of functions on Ω whose restriction on each element K belongs to the Sobolev space $H^\ell(K)$. Then, we define the finite element space $\mathbf{V}_N \times M_N$ as

$$M_N = \{v \in L^2(\mathcal{T}_N) \mid v|_K \in \mathcal{Q}_k(K), \forall K \in \mathcal{T}_N\},$$

$$\mathbf{V}_N = M_N \times M_N.$$

We denote by \mathcal{E} , the set of all edges of the mesh. Let $\mathcal{E}^0 \subset \mathcal{E}$ be the set of interior edges, and $\mathcal{E}^B \subset \mathcal{E}$ the set of edges on $\partial\Omega$. Let \mathbf{n} be the outer normal vector and denote

$$\mathcal{E}_-^B = \{(x, y) \in \mathcal{E}^B \mid \beta(x, y) \cdot \mathbf{n}(x, y) < 0\},$$

$$\mathcal{E}_+^B = \{(x, y) \in \mathcal{E}^B \mid \beta(x, y) \cdot \mathbf{n}(x, y) > 0\}.$$

We denote by v^\pm , the function defined on two adjacent elements K^\pm that share a horizontal/vertical edge e , and \mathbf{n}_e^\pm the unit normal vector on e pointing exterior to K^\pm . We further denote by $\gamma_1^\pm, \gamma_2^\pm$ four edges of element K and $h_{x,K} = \frac{1}{2}|\gamma_1^\pm|, h_{y,K} = \frac{1}{2}|\gamma_2^\pm|$ (See Figure 1).

For any $K \in \mathcal{T}_N$, we denote the inner product on $L^2(K)$ as $(\cdot, \cdot)_K$, and $(\cdot, \cdot) = (\cdot, \cdot)_{\mathcal{T}_N}$. For any $e \in \mathcal{E}$, we denote the inner product on $L^2(e)$ as $\langle \cdot, \cdot \rangle_e$. Let $(\cdot, \cdot)_{\mathcal{D}} := \sum_{K \in \mathcal{D}} (\cdot, \cdot)_K$ be the inner product on \mathcal{D} , for any $\mathcal{D} \subseteq \mathcal{T}_N$. $|\cdot|_{s,\mathcal{D}}$ and $\|\cdot\|_{s,\mathcal{D}}$ denote the semi-norm and the norm of Sobolev space $H^s(\mathcal{D})$. We drop the first subscript in the semi-norm and the norm whenever $s = 0$, and drop the second subscript if $\mathcal{D} = \mathcal{T}_N$. Similarly, for any $\mathcal{D} \subseteq \mathcal{E}$ the inner product on $L^2(\mathcal{D})$ is defined as $\langle \cdot, \cdot \rangle_{\mathcal{D}} := \sum_{e \in \mathcal{D}} \langle \cdot, \cdot \rangle_e$.

For a function v and a vector function \mathbf{r} , we define jumps and averages on an edge e as:

$$\begin{aligned} \llbracket \mathbf{r} \rrbracket|_e &= \mathbf{r}^+ \cdot \mathbf{n}^+ + \mathbf{r}^- \cdot \mathbf{n}^-, & \forall e \in \mathcal{E}^0, \\ \llbracket v \rrbracket|_e &= v^+ \mathbf{n}^+ + v^- \mathbf{n}^-, & \forall e \in \mathcal{E}^0, \\ \llbracket v \rrbracket|_e &= v\mathbf{n}, & \forall e \in \mathcal{E}^B, \end{aligned}$$

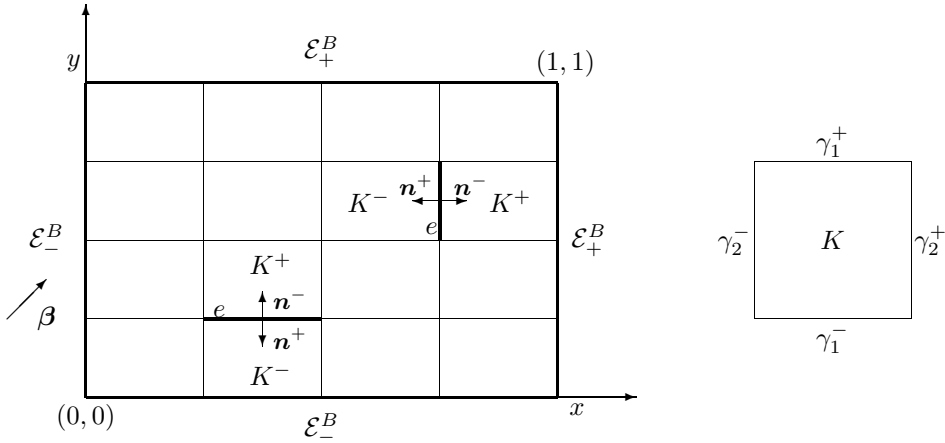


FIGURE 1. Some notations about the domain and the decomposition.

$$\begin{aligned} \{\mathbf{r}\}|_e &= \frac{1}{2}(\mathbf{r}^+ + \mathbf{r}^-), & \forall e \in \mathcal{E}^0, \\ \{\mathbf{r}\}|_e &= \mathbf{r}, & \forall e \in \mathcal{E}^B, \\ \{v\}|_e &= \frac{1}{2}(v^+ + v^-), & \forall e \in \mathcal{E}^0, \\ \{v\}|_e &= v, & \forall e \in \mathcal{E}^B. \end{aligned}$$

2.1. **LDG formulation.** Introducing a new variable $\mathbf{q} := \epsilon \nabla u$, we rewrite (1.1) as a first-order system

$$\begin{aligned} \mathbf{q} - \epsilon \nabla u &= 0, & \text{in } \Omega, \\ -\nabla \cdot \mathbf{q} + \beta \cdot \nabla u + cu &= f, & \text{in } \Omega, \\ u &= 0, & \text{on } \partial\Omega. \end{aligned} \tag{2.1}$$

Then we will look for the approximate solution (\mathbf{q}_N, u_N) of the LDG method from $\mathbf{V}_N \times M_N$. We consider the following LDG formulation ([8], [24]): Find $(\mathbf{q}_N, u_N) \in \mathbf{V}_N \times M_N$, such that for all $(\mathbf{r}, v) \in \mathbf{V}_N \times M_N$,

$$\begin{aligned} \epsilon^{-1} (\mathbf{q}_N, \mathbf{r}) + (u_N, \nabla \cdot \mathbf{r}) - \sum_{K \in \mathcal{T}_N} \langle \hat{u}, \mathbf{r} \cdot \mathbf{n}_K \rangle_{\partial K} &= 0, \\ (\mathbf{q}_N, \nabla v) + (u_N, (c - \beta \cdot \nabla - \text{div} \beta)v) & \\ - \sum_{K \in \mathcal{T}_N} \langle \hat{\mathbf{q}} \cdot \mathbf{n}_K, v \rangle_{\partial K} + \sum_{K \in \mathcal{T}_N} \langle \hat{u}^c, v(\beta \cdot \mathbf{n}_K) \rangle_{\partial K} &= (f, v), \end{aligned} \tag{2.2}$$

where numerical flux \hat{u}^c is the upwinding trace, $\hat{\mathbf{q}}$ and \hat{u} are numerical fluxes defined as

$$\begin{aligned} \hat{u} &:= u_N^-, & \text{on } \mathcal{E}^0, \\ \hat{\mathbf{q}} &:= \mathbf{q}_N^+ - C_{11} \llbracket u_N \rrbracket, & \text{on } \mathcal{E}^0, \\ \hat{u} &:= 0, & \text{on } \mathcal{E}^B, \\ \hat{\mathbf{q}} &:= \mathbf{q}_N - C_{11} u_N \mathbf{n}, & \text{on } \mathcal{E}^B. \end{aligned}$$

Moreover, the penalization parameter C_{11} satisfies $0 \leq C_{11} \leq O(1)$. Notice that \hat{u} , \hat{u}^c , $\hat{\mathbf{q}}$ are conservative, i.e., $[[\hat{u}]] = 0$ on \mathcal{E} , $[[\hat{u}^c]] = 0$, $[[\hat{\mathbf{q}}]] = 0$ on \mathcal{E}^0 . Substituting the definition of numerical fluxes, we can write (2.2) as: for any $(\mathbf{r}, v) \in \mathbf{V}_N \times M_N$,

$$(2.3) \quad \begin{aligned} a(\mathbf{q}_N, \mathbf{r}) + b_1(u_N, \mathbf{r}) &= 0, \\ b_2(v, \mathbf{q}_N) + c(u_N, v) &= f(v), \end{aligned}$$

where

$$\begin{aligned} a(\mathbf{q}_N, \mathbf{r}) &= (\epsilon^{-1} \mathbf{q}_N, \mathbf{r}), \\ b_1(u_N, \mathbf{r}) &= (u_N, \nabla \cdot \mathbf{r}) - \langle u_N^-, [[\mathbf{r}]] \rangle_{\mathcal{E}^0}, \\ b_2(v, \mathbf{q}_N) &= (\nabla v, \mathbf{q}_N) - \langle \mathbf{q}_N^+, [[v]] \rangle_{\mathcal{E}^0} - \langle \mathbf{q}_N, v \mathbf{n} \rangle_{\mathcal{E}^B}, \\ c(u_N, v) &= (u_N, (c - \boldsymbol{\beta} \cdot \nabla - \operatorname{div} \boldsymbol{\beta})v) + \langle C_{11} [[u_N]], [[v]] \rangle_{\mathcal{E}} + \langle u_N^-, \boldsymbol{\beta} \cdot [[v]] \rangle_{\mathcal{E} \setminus \mathcal{E}_+^B}, \\ f(v) &= (f, v). \end{aligned}$$

Remark 2.1. Using integration by parts and direct calculation, we obtain

$$(2.4) \quad b_1(v, \mathbf{r}) = -b_2(v, \mathbf{r})$$

for any $(\mathbf{r}, v) \in \mathbf{V}_N \times M_N$. Meanwhile, applying integration by parts to the convection term in $c(u_N, v)$, we obtain another expression of $c(u_N, v)$:

$$c(u_N, v) = (\boldsymbol{\beta} \cdot \nabla u_N, v) + (u_N, cv) + \langle C_{11} [[u_N]], [[v]] \rangle_{\mathcal{E}} - \langle v^+, \boldsymbol{\beta} \cdot [[u_N]] \rangle_{\mathcal{E} \setminus \mathcal{E}_+^B}.$$

Adding the two equations of (2.3) together gives the compact form of the LDG method:

$$(2.5) \quad \mathcal{A}(\mathbf{q}_N, u_N; \mathbf{r}, v) = f(v), \quad \forall (\mathbf{r}, v) \in \mathbf{V}_N \times M_N,$$

where

$$\mathcal{A}(\mathbf{q}_N, u_N; \mathbf{r}, v) := a(\mathbf{q}_N, \mathbf{r}) + b_1(u_N, \mathbf{r}) - b_1(v, \mathbf{q}_N) + c(u_N, v).$$

Because of the consistency of the numerical fluxes, we have the *Galerkin orthogonality property*

$$(2.6) \quad \mathcal{A}(\mathbf{q} - \mathbf{q}_N, u - u_N; \mathbf{r}, v) = 0, \quad \forall (\mathbf{r}, v) \in \mathbf{V}_N \times M_N.$$

We also introduce an ϵ -weighted norm $|(\mathbf{r}, v)|_{\mathcal{A}}$ which appears in a natural way in the analysis of the LDG method and is defined as

$$(2.7) \quad |(\mathbf{r}, v)|_{\mathcal{A}}^2 := \epsilon^{-1} \|\mathbf{r}\|^2 + \left\| \left(c - \frac{1}{2} \operatorname{div} \boldsymbol{\beta} \right)^{\frac{1}{2}} v \right\|^2 + \left\| \left(C_{11} + \frac{1}{2} |\boldsymbol{\beta} \cdot \mathbf{n}| \right)^{\frac{1}{2}} [[v]] \right\|_{\mathcal{E}}^2,$$

for any $(\mathbf{r}, v) \in [L^2(\mathcal{T}_N)]^2 \times H^1(\mathcal{T}_N)$.

Lemma 2.2. *For any $(\mathbf{r}, v) \in \mathbf{V}_N \times W_N$, we have*

$$(2.8) \quad \mathcal{A}(\mathbf{r}, v; \mathbf{r}, v) = |(\mathbf{r}, v)|_{\mathcal{A}}^2.$$

Proof. It follows from (2.6) and (2.4) that

$$(2.9) \quad \begin{aligned} \mathcal{A}(\mathbf{r}, v; \mathbf{r}, v) &= a(\mathbf{r}, \mathbf{r}) + b_1(v, \mathbf{r}) - b_1(v, \mathbf{r}) + c(v, v) \\ &= a(\mathbf{r}, \mathbf{r}) + c(v, v) \\ &= \epsilon^{-1} \|\mathbf{r}\|^2 + c(v, v), \end{aligned}$$

where the bilinear form $c(v, v)$, which is defined in (2.3), can be transformed into

$$\begin{aligned}
 c(v, v) &= (v, (c - \frac{1}{2} \operatorname{div} \boldsymbol{\beta})v) + \langle C_{11} \llbracket v \rrbracket, \llbracket v \rrbracket \rangle_{\mathcal{E}} - \frac{1}{2} \sum_{K \in \mathcal{T}_N} \langle (\boldsymbol{\beta} \cdot \mathbf{n})v, v \rangle_{\partial K} \\
 &\quad + \sum_{e \in \mathcal{E}^0} \langle (\boldsymbol{\beta} \cdot \mathbf{n})v^-, (v^- - v^+) \rangle_e + \sum_{e \in \mathcal{E}^B_{\pm}} \langle (\boldsymbol{\beta} \cdot \mathbf{n})v, v \rangle_e \\
 &= \left\| (c - \frac{1}{2} \operatorname{div} \boldsymbol{\beta})^{\frac{1}{2}} v \right\|^2 + \left\| C_{11}^{\frac{1}{2}} \llbracket v \rrbracket \right\|_{\mathcal{E}}^2 + \langle |\boldsymbol{\beta} \cdot \mathbf{n}| v^-(v^- - v^+), 1 \rangle_{\mathcal{E}^0} + \langle |\boldsymbol{\beta} \cdot \mathbf{n}| v^2, 1 \rangle_{\mathcal{E}^B_{\pm}} \\
 &\quad - \frac{1}{2} \left[\langle |\boldsymbol{\beta} \cdot \mathbf{n}| [(v^-)^2 - (v^+)^2], 1 \rangle_{\mathcal{E}^0} + \langle |\boldsymbol{\beta} \cdot \mathbf{n}| v^2, 1 \rangle_{\mathcal{E}^B_{\pm}} - \langle |\boldsymbol{\beta} \cdot \mathbf{n}| v^2, 1 \rangle_{\mathcal{E}^B_{-}} \right] \\
 &= \left\| (c - \frac{1}{2} \operatorname{div} \boldsymbol{\beta})^{\frac{1}{2}} v \right\|^2 + \left\| C_{11}^{\frac{1}{2}} \llbracket v \rrbracket \right\|_{\mathcal{E}}^2 + \frac{1}{2} \left\| |\boldsymbol{\beta} \cdot \mathbf{n}|^{\frac{1}{2}} (v^- - v^+) \right\|_{\mathcal{E}^0}^2 + \frac{1}{2} \left\| |\boldsymbol{\beta} \cdot \mathbf{n}|^{\frac{1}{2}} v \right\|_{\mathcal{E}^B}^2 \\
 &= \left\| (c - \frac{1}{2} \operatorname{div} \boldsymbol{\beta})^{\frac{1}{2}} v \right\|^2 + \left\| (C_{11} + |\boldsymbol{\beta} \cdot \mathbf{n}|/2)^{\frac{1}{2}} \llbracket v \rrbracket \right\|_{\mathcal{E}}^2.
 \end{aligned}$$

Here, in the first inequality, we used the following identity

$$(v, \boldsymbol{\beta} \cdot \nabla v) = -\frac{1}{2} (v, (\operatorname{div} \boldsymbol{\beta})v) + \frac{1}{2} \sum_{K \in \mathcal{T}_N} \langle (\boldsymbol{\beta} \cdot \mathbf{n})v, v \rangle_{\partial K}.$$

Comparing (2.7) and (2.9) establishes the stability result (2.9). □

Therefore, by setting $f = 0$ and $(\mathbf{r}, v) = (\mathbf{q}_N, u_N)$, we have $\mathcal{A}(\mathbf{q}_N, u_N; \mathbf{q}_N, u_N) = |(\mathbf{q}_N, u_N)|_{\mathcal{A}}^2 = 0$. It implies $\mathbf{q}_N = 0$ and $u_N = 0$, from which the existence and uniqueness of the LDG solution follow directly.

2.2. Regularities. We assume that the solution u can be decomposed in a precise way that is typical of the behavior in solutions of (1.1) observed when interior layers are excluded [19]. A proof is given in [13] that this decomposition does exist for $k = 1$ under certain compatibility conditions on f . The extension of these results to our case $k > 1$ is nontrivial and can be done by placing a sufficient number of compatibility conditions on f (cf. [13], [19]) or by using cut-off functions to obtain bounds on the error in the computed solution in regions outside of all layers induced by inadequate compatibility of the data (cf. [10]).

Assume that exact solution of (1.1) admits a decomposition of the form

$$(2.10) \quad u = \bar{u} + w_0 + w_1 + w_2$$

where for all $(x, y) \in \Omega$ the component functions have the regularity

$$(2.11) \quad \begin{aligned}
 \left| \frac{\partial^{i+j} \bar{u}}{\partial^i x \partial^j y}(x, y) \right| &\leq C, & \left| \frac{\partial^{i+j} w_0}{\partial^i x \partial^j y}(x, y) \right| &\leq C \epsilon^{-(i+j)} e^{-[\beta_1(1-x) + \beta_2(1-y)]/\epsilon}, \\
 \left| \frac{\partial^{i+j} w_1}{\partial^i x \partial^j y}(x, y) \right| &\leq C \epsilon^{-i} e^{-\beta_1(1-x)/\epsilon}, & \left| \frac{\partial^{i+j} w_2}{\partial^i x \partial^j y}(x, y) \right| &\leq C \epsilon^{-j} e^{-\beta_2(1-y)/\epsilon}
 \end{aligned}$$

for $0 \leq i + j \leq k + 2$. Here constant C depends on $\boldsymbol{\beta}$, c and f .

As a consequence, we get a decomposition of $\mathbf{q} := (q_1, q_2) = \epsilon(u_x, u_y)$:

$$(2.12) \quad \begin{aligned}
 q_1 &= \bar{q}_1 + q_{1,0} + q_{1,1} + q_{1,2} = \epsilon(\bar{u}_x + w_{0,x} + w_{1,x} + w_{2,x}), \\
 q_2 &= \bar{q}_2 + q_{2,0} + q_{2,1} + q_{2,2} = \epsilon(\bar{u}_y + w_{0,y} + w_{1,y} + w_{2,y}).
 \end{aligned}$$

Sometimes we also use $\mathbf{q} = \bar{\mathbf{q}} + \mathbf{q}_0 + \mathbf{q}_1 + \mathbf{q}_2 := (\bar{q}_1, \bar{q}_2) + (q_{1,0}, q_{2,0}) + (q_{1,1}, q_{2,1}) + (q_{1,2}, q_{2,2})$. It follows directly from (2.11) and (2.12) that

$$(2.13) \quad \begin{aligned} \left| \frac{\partial^{i+j} \bar{q}_\ell}{\partial^i x \partial^j y}(x, y) \right| &\leq C\epsilon, & \ell = 1, 2, \\ \left| \frac{\partial^{i+j} q_{\ell,0}}{\partial^i x \partial^j y}(x, y) \right| &\leq C\epsilon^{-(i+j)} e^{-[\beta_1(1-x) + \beta_2(1-y)]/\epsilon}, & \ell = 1, 2, \\ \left| \frac{\partial^{i+j} q_{1,1}}{\partial^i x \partial^j y}(x, y) \right| &\leq C\epsilon^{-i} e^{-\beta_1(1-x)/\epsilon}, & \left| \frac{\partial^{i+j} q_{1,2}}{\partial^i x \partial^j y}(x, y) \right| &\leq C\epsilon^{1-j} e^{-\beta_2(1-y)/\epsilon}, \\ \left| \frac{\partial^{i+j} q_{2,1}}{\partial^i x \partial^j y}(x, y) \right| &\leq C\epsilon^{1-i} e^{-\beta_1(1-x)/\epsilon}, & \left| \frac{\partial^{i+j} q_{2,2}}{\partial^i x \partial^j y}(x, y) \right| &\leq C\epsilon^{-j} e^{-\beta_2(1-y)/\epsilon} \end{aligned}$$

for $0 \leq i + j \leq k + 1$. Here constant C only depends on β , c and f .

Remark 2.3. For the investigation of uniform convergence of a numerical method on layer-adapted meshes, it is practical to assume the above regularities (cf. [16, 22]). So far, a complete regularity analysis on the convection–diffusion problems seems lacking, although the counterpart results for the reaction–diffusion problems are relatively rich. While most of the regularity results for the solution of the problem have been found on the square or rectangular domain (cf. [11, 13, 14, 16]), regularity results on a general polygonal domain have not yet been reported.

2.3. Shishkin mesh. The tensor product Shishkin mesh \mathcal{T}_N that we shall construct in $\bar{\Omega}$ is similar to the mesh used in [19]. Define the transition parameter

$$(2.14) \quad \tau_x = \min\left(\frac{1}{2}, \frac{\kappa}{\beta_1} \epsilon \ln N\right), \quad \tau_y = \min\left(\frac{1}{2}, \frac{\kappa}{\beta_2} \epsilon \ln N\right),$$

where the transition parameter $\kappa = k + 1$. The domain Ω is divided into four parts:

$$\begin{aligned} \Omega_0 &= (0, 1 - \tau_x) \times (0, 1 - \tau_y), & \Omega_x &= (1 - \tau_x, 1) \times (0, 1 - \tau_y), \\ \Omega_y &= (0, 1 - \tau_x) \times (1 - \tau_y, 1), & \Omega_{xy} &= (1 - \tau_x, 1) \times (1 - \tau_y, 1). \end{aligned}$$

Each subdomain is then equally decomposed into $N \times N$ rectangles (see Figure 2). Therefore, \mathcal{T}_N is composed of $4N^2$ elements. Meanwhile, there are $(2N + 1)^2$ nodes (x_i, y_i) , $i = 0, \dots, 2N$. We denote

$$H_x = \frac{1 - \tau_x}{2N}, \quad h_x = \frac{\tau_x}{2N}; \quad H_y = \frac{1 - \tau_y}{2N}, \quad h_y = \frac{\tau_y}{2N}.$$

Let $H = \max\{H_x, H_y\}$, $h = \max\{h_x, h_y\}$. Since β_1 and β_2 are independent of ϵ and N , we assume that $H = \min\{H_x, H_y\}$ and $h = \min\{h_x, h_y\}$.

3. MAIN RESULTS

In this section, we present our main result and some remarks. Detailed proof of the main result is provided in the next section.

Theorem 3.1. *Suppose that (\mathbf{q}, u) is the solution of (2.1) that satisfies the regularity (2.13) and (2.11), $(\mathbf{q}_N, u_N) \in \mathbf{V}_N \times M_N$ is the solution of (2.3). If $0 \leq C_{11} \leq O(1)$, then there exists a constant C independent of ϵ and N such that*

$$(3.1) \quad |(\mathbf{q} - \mathbf{q}_N, u - u_N)|_{\mathcal{A}} \leq C \left(\frac{\ln N}{N}\right)^{k+\frac{1}{2}}.$$

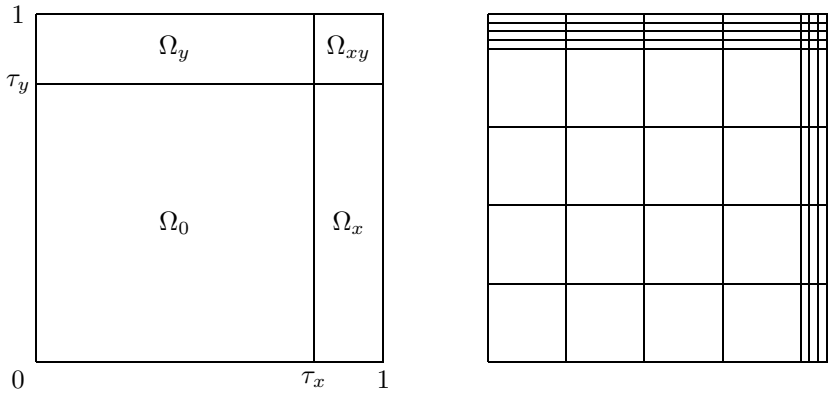


FIGURE 2. Subdomain notation and the Shishkin mesh \mathcal{T}_4 .

Proof. The inequality (3.1) follows the triangle inequality

$$|(\mathbf{q} - \mathbf{q}_N, u - u_N)|_{\mathcal{A}} \leq |(\mathbf{q} - \mathbf{q}_I, u - u_I)|_{\mathcal{A}} + |(\mathbf{q}_I - \mathbf{q}_N, u_I - u_N)|_{\mathcal{A}},$$

the interpolation error estimates (4.11b), (4.11d), (4.17a), and the inequality (4.30). Here we used another DG-norm $\|\cdot\|_{\mathcal{A}}$ defined by (4.21). \square

Remark 3.2. The error bound in (3.1) is uniformly valid in terms of singular perturbation parameter ϵ . It was pointed out in [21] that the central point of global error estimates for singularly perturbed problems is the ϵ independence.

Remark 3.3. We could also assume $0 \leq C_{11} \leq O(N)$ on \mathcal{E}_+^B and $C_{11} = 0$ on $\mathcal{E} \setminus \mathcal{E}_+^B$ in the proof of (3.1). This choice of C_{11} is similar to the MD-LDG (cf. [7]), in which the dissipative effect is minimized by setting the penalization parameter C_{11} to zero on \mathcal{E} except the outflow boundary.

If this choice of C_{11} is used, we can actually derive the same error bounds, and nothing needs to be changed in the proof of Theorem 3.1.

4. PROOFS

The proof of Theorem 3.1 will be presented in this section. To this end, we shall proceed in several steps. We first list some preliminary inequalities for future use in Section 4.1 and introduce projection type interpolations (\mathbf{q}_I, u_I) of the exact solution (\mathbf{q}, u) in Section 4.2. Then, we divide the error of the LDG approximation into two parts:

$$(\mathbf{q} - \mathbf{q}_N, u - u_N) = (\mathbf{q} - \mathbf{q}_I, u - u_I) + (\mathbf{q}_I - \mathbf{q}_N, u_I - u_N).$$

Section 4.3 is devoted to the estimation of $\mathbf{q} - \mathbf{q}_I$ and $u - u_I$. Using a stability result proved in Section 4.4, we estimate the error $\mathbf{q}_I - \mathbf{q}_N$ and $u_I - u_N$ in Section 4.5.

4.1. Preliminaries. We introduce some inequalities that will be used in the analysis. Because of the fact $\mathbf{V}_N = M_N \times M_N$, the imbedding inequality and the inverse inequality are valid for $v \in M_N$ and for each component of $\mathbf{r} \in \mathbf{V}_N$. Their proofs are straightforward calculations and are omitted.

Imbedding inequalities:

$$(4.1) \quad \|v\|_{\gamma_2^- \cup \gamma_2^+}^2 \leq Ch_{x,K}^{-1} \|v\|_K^2, \quad \|v\|_{\gamma_1^- \cup \gamma_1^+}^2 \leq Ch_{y,K}^{-1} \|v\|_K^2.$$

Inverse inequalities:

$$(4.2) \quad \|v_x\|_K^2 \leq Ch_{x,K}^{-2} \|v\|_K^2, \quad \|v_y\|_K^2 \leq Ch_{y,K}^{-2} \|v\|_K^2.$$

Trace theorem (cf. Theorem 1.6.6, [2]): If a domain D has a Lipschitz boundary, then for any $v \in H^1(D) \cup C(\overline{D})$ there is a constant C such that

$$(4.3) \quad \|v\|_{\partial D} \leq C \left(\|v\|_D^{\frac{1}{2}} \|v\|_{1,D}^{\frac{1}{2}} \right) \leq C (|v|_{1,D} + \|v\|_D).$$

Next, we list two frequently used inequalities about the exponential boundary layer functions:

$$(4.4) \quad \sum_{i=1}^N e^{-2\beta_1(1-x_i)/\epsilon} \leq CN^{-2\kappa}, \quad \max\{e^{-2\beta_1\tau_x/\epsilon}, e^{-2\beta_2\tau_y/\epsilon}\} \leq CN^{-2\kappa}.$$

We list some notation about subsets of \mathcal{E} . This notation will be used in interpolation error estimates. Let

$$\begin{aligned} \Theta^v &= \{e \in \mathcal{E} \mid e \text{ is vertical}\}, & \Theta^h &= \{e \in \mathcal{E} \mid e \text{ is horizontal}\}, \\ \Theta_0^v &= \{e \in \Theta^v \mid e \subset \overline{\Omega_0}\}, & \Theta_0^h &= \{e \in \Theta^h \mid e \subset \overline{\Omega_0}\}, \\ \Theta_x^v &= \{e \in \Theta^v \mid e \subset \overline{\Omega_0 \cup \Omega_x} \setminus \overline{\Omega_0}\}, & \Theta_x^h &= \{e \in \Theta^h \mid e \subset \overline{\Omega_0 \cup \Omega_x} \setminus \overline{\Omega_0}\}, \\ \Theta_y^v &= \{e \in \Theta^v \mid e \subset \overline{\Omega_0 \cup \Omega_y} \setminus \overline{\Omega_0}\}, & \Theta_y^h &= \{e \in \Theta^h \mid e \subset \overline{\Omega_0 \cup \Omega_y} \setminus \overline{\Omega_0}\}, \\ \Theta_{xy}^v &= \Theta^v \setminus (\Theta_0^v \cup \Theta_x^v \cup \Theta_y^v), & \Theta_{xy}^h &= \Theta^h \setminus (\Theta_0^h \cup \Theta_x^h \cup \Theta_y^h), \\ \Theta_1^h &= \Theta_0^h \cup \Theta_y^h, \quad \Theta_2^h = \Theta_x^h \cup \Theta_{xy}^h, & \Theta_3^v &= \Theta_0^v \cup \Theta_x^v, \quad \Theta_4^v = \Theta_y^v \cup \Theta_{xy}^v, \\ \Theta_0 &= \Theta_0^h \cup \Theta_0^v, \quad \Theta_x = \Theta_x^h \cup \Theta_x^v, & \Theta_y &= \Theta_y^h \cup \Theta_y^v, \quad \Theta_{xy} = \Theta_{xy}^h \cup \Theta_{xy}^v, \\ \mathcal{E}_1^B &= \{e \in \mathcal{E}_+^B : e \text{ is horizontal}\}, & \mathcal{E}_2^B &= \{e \in \mathcal{E}_+^B : e \text{ is vertical}\}. \end{aligned}$$

4.2. Interpolation. To define interpolation operators, we start from some one-dimensional projections (cf. § 3.2,[8]). Let $\mathcal{P}_k(\hat{I})$ be the space of the polynomials of degree at most k on $\hat{I} = (-1, 1)$. We denote by $\hat{\pi}$ the $L^2(\hat{I})$ projection onto $\mathcal{P}_k(\hat{I})$, i.e., for a function $\hat{w} \in L^2(\hat{I})$, $\hat{\pi}\hat{w}$ is the unique polynomial in $\mathcal{P}_k(\hat{I})$ satisfying

$$\int_{\hat{I}} [\hat{w}(\xi) - \hat{\pi}\hat{w}(\xi)]p(\xi)d\xi = 0, \quad \forall \hat{p} \in \mathcal{P}_k(\hat{I}).$$

Furthermore, for $\hat{w} \in H^{\frac{1}{2}+\mu}$ with $\mu > 0$, we define the projection $\hat{\pi}^\pm \hat{w} \in \mathcal{P}_k(\hat{I})$ by the following conditions:

$$\hat{\pi}^\pm \hat{w}(\pm 1) = \hat{w}(\pm 1), \quad \int_{\hat{I}} [\hat{w}(x) - \hat{\pi}^\pm \hat{w}(\xi)] \hat{p}(\xi) d\xi = 0 \quad \forall \hat{p} \in \mathcal{P}_{k-1}(\hat{I}).$$

On a rectangle $\hat{K} = \hat{I} \times \hat{I}$, we define the following tensor product operators for $v \in C^0(\overline{\hat{K}})$ and $\hat{r} = (\hat{r}_1, \hat{r}_2) \in H^1(\hat{K})^2$:

$$(4.5) \quad \begin{aligned} \hat{\Pi}^\pm \hat{v} &:= \hat{\pi}_1^\pm \otimes \hat{\pi}_2^\pm \hat{v}, \\ \hat{\Pi}^\pm \hat{r} &:= (\hat{\pi}_1^\pm \otimes \hat{\pi}_2 \hat{r}_1, \hat{\pi}_1 \otimes \hat{\pi}_2^\pm \hat{r}_2). \end{aligned}$$

The above definitions actually define the interpolants in a unique way (see §3.2, [8]). The interpolation operators satisfy

$$\begin{aligned}
 (4.6) \quad & \int_{\widehat{K}} (\widehat{\mathbf{r}} - \widehat{\mathbf{\Pi}}^\pm \widehat{\mathbf{r}}) \cdot \nabla \widehat{p} = 0, & \forall \widehat{p} \in \mathcal{Q}_k(\widehat{K}), \\
 & \int_{\widehat{\gamma}_i^\pm} (\widehat{\mathbf{r}} - \widehat{\mathbf{\Pi}}^\pm \widehat{\mathbf{r}}) \cdot \mathbf{n} \widehat{p} = 0, & \forall \widehat{p} \in \mathcal{Q}_k(\widehat{\gamma}_i^\pm), \quad i = 1, 2, \\
 & \int_{\widehat{K}} (\widehat{v} - \Pi^\pm \widehat{v}) \widehat{p} = 0, & \forall \widehat{p} \in \mathcal{Q}_{k-1}(\widehat{K}), \\
 & \int_{\widehat{\gamma}_i^\pm} (\widehat{v} - \widehat{\Pi}^\pm \widehat{v}) \widehat{p} = 0, & \forall \widehat{p} \in \mathcal{Q}_{k-1}(\widehat{\gamma}_i^\pm), \quad i = 1, 2.
 \end{aligned}$$

Using the transformation $x = x_K + h_{x,K}\xi/2$, $y = y_K + h_{y,K}\eta/2$ to map from \widehat{K} to an arbitrary $K \in \mathcal{T}_N$, we obtain the corresponding approximation operators $\mathbf{\Pi}^\pm : C(K)^2 \rightarrow (\mathcal{Q}_k(K))^2$ and $\Pi^\pm : C(K) \rightarrow \mathcal{Q}_k(K)$.

Furthermore, we denote by $I : C(K) \rightarrow \mathcal{Q}_k(K)$ the standard Lagrange interpolation operator.

We define $u_I|_K = I\bar{u}|_K + \Pi^+ w|_K$ and $\mathbf{q}_I|_K = \mathbf{\Pi}^- \mathbf{q}|_K$, for any $K \in \mathcal{T}_N$. It is clear that $(\mathbf{q}_I, u_I) \in \mathbf{V}_N \times W_N$. Let $\xi_{\mathbf{q}} = \mathbf{q} - \mathbf{q}_I$, $\xi_u = u - u_I$, $\eta_{\mathbf{q}} = \mathbf{q}_I - \mathbf{q}_N$, $\eta_u = u_I - u_N$. As a consequence,

$$(4.7) \quad (\mathbf{q} - \mathbf{q}_N, u - u_N) = (\xi_{\mathbf{q}}, \xi_u) + (\eta_{\mathbf{q}}, \eta_u).$$

For simplicity, let $\xi_{\mathbf{r}} := \mathbf{r} - \Pi^+ \mathbf{r}$, $\xi_{\bar{u}} := \bar{u} - I\bar{u}$, and $\xi_v := v - \Pi^- v$ for any $v \in \{w, w_0, w_1, w_2\}$.

The rest of this section is devoted to the estimation of the interpolation error $(\xi_{\mathbf{q}}, \xi_u)$ and the error $(\eta_{\mathbf{q}}, \eta_u)$.

Remark 4.1. The interpolant $\Pi^- \mathbf{q}$ and $\Pi^+ w$ simplifies the error estimation of $|(\eta_{\mathbf{q}}, \eta_u)|_{\mathcal{A}}$ (see Lemma 4.7) because of its property (4.6). It seems that standard Lagrange interpolants and L^2 projections do not have these advantages. However, these interpolants could not simplify $\llbracket \xi_u \rrbracket$, while $\llbracket \xi_u \rrbracket = 0$ if a standard Lagrange interpolant is used.

Actually, the error bound of $\|\llbracket \bar{u} - \Pi^+ \bar{u} \rrbracket\|_{\mathcal{E}}$ is not ϵ -uniform because of the anisotropic property of Shishkin mesh. So we adopt standard Lagrange interpolation only for \bar{u} without causing any loss in the rate of convergence, in view of the regularity of the exact solution and its decomposition. Moreover, it is not necessary to change the interpolation for $\bar{\mathbf{q}}$ because $\bar{\mathbf{q}} = \epsilon \nabla \bar{u}$, which is of $O(\epsilon)$ and will not cause any trouble in the error estimation on anisotropic mesh.

4.3. Interpolation error estimates. To obtain satisfactory approximation error estimates on $\Omega_x \cup \Omega_y$, where the Shishkin mesh is highly anisotropic we use some technical results of Apel [1] (see Lemma 2.14, Theorem 2.7).

Lemma 4.2. *Assume that \widehat{K} is a reference square. Then, for all $\widehat{v} \in C(\widehat{K}) \cap H^{k+1}(\widehat{K})$, $\widehat{\mathbf{r}} \in C(\widehat{K})^2 \cap H^{k+1}(\widehat{K})^2$, there exists a constant C independent of \widehat{v} and*

\hat{r} such that

$$\begin{aligned}
 (4.8a) \quad & \|\Pi^+ \hat{v}\|_{L^2(\hat{K})} \leq C \|\hat{v}\|_{C^0(\overline{\hat{K}})}, \\
 (4.8b) \quad & \|\hat{v} - \Pi^+ \hat{v}\|_{L^2(\hat{K})} \leq C [\hat{v}]_{H^{k+1}(\hat{K})}, \\
 (4.8c) \quad & \|\hat{D}^\alpha(\hat{v} - \Pi^+ \hat{v})\|_{L^2(\hat{K})^2} \leq C [\hat{D}^\alpha \hat{v}]_{H^k(\hat{K})^2}, \quad \alpha \in \{(1, 0), (0, 1)\}, \\
 (4.8d) \quad & \|\Pi^- \hat{r}\|_{H^1(\hat{K})} \leq C \|\hat{r}\|_{H^1(\overline{\hat{K}})}, \\
 (4.8e) \quad & \|\hat{r} - \Pi^- \hat{r}\|_{L^2(\hat{K})} \leq C [\hat{r}]_{H^{k+1}(\hat{K})}, \\
 (4.8f) \quad & \|\hat{D}^\alpha(\hat{r} - \Pi^- \hat{r})\|_{L^2(\hat{K})^2} \leq C [\hat{D}^\alpha \hat{r}]_{H^k(\hat{K})^2}, \quad \alpha \in \{(1, 0), (0, 1)\},
 \end{aligned}$$

where we denote

$$(4.9) \quad [\hat{r}]_{H^s(\hat{K})}^2 := \sum_{i=1,2} \sum_{|\alpha|=1} \|\hat{D}^{s\alpha} \hat{r}_i\|_{\hat{K}}^2, \quad [\hat{v}]_{H^s(\hat{K})}^2 := \sum_{|\alpha|=1} \|\hat{D}^{s\alpha} \hat{v}\|_{\hat{K}}^2,$$

for $s = k, k + 1$.

Proof. The inequalities (4.8a) and (4.8d) could be proved by using Lemma 3.3, [8] and embedding theory of Sobolev spaces. Next, we will prove (4.8b). It is clear that the interpolation operator Π^+ is a linear operator. Define $(k + 1)^2$ continuous linear functionals $F_i : H^k(\hat{K}) \rightarrow \mathbb{R}$ by the definition of the interpolation

$$\begin{aligned}
 \hat{v} & \mapsto \hat{v}(1, 1); \\
 \hat{v} & \mapsto \int_{\hat{\gamma}_1^+} \xi^i \hat{v} d\xi, & \text{for } i = 0, 1, \dots, k - 1; \\
 \hat{v} & \mapsto \int_{\hat{\gamma}_2^+} \eta^i \hat{v} d\eta, & \text{for } i = 0, 1, \dots, k - 1; \\
 \hat{v} & \mapsto \int_{\hat{K}} \xi^j \eta^\ell \hat{v} d\xi d\eta, & \text{for } j = 0, 1, \dots, k - 1, \ell = 0, 1, \dots, k - 1.
 \end{aligned}$$

Recalling the definition of Π^+ , it is straightforward to verify that

$$F_i(\hat{v} - \Pi^+ \hat{v}) = 0 \quad \forall \hat{v} \in \mathcal{C}(\hat{K}) \cap H^k(\hat{K}) \quad \text{for } i = 1, 2, \dots, (k + 1)^2,$$

and

$$F_i(\hat{v}) = 0 \quad \text{for } i = 1, 2, \dots, (k + 1)^2 \quad \Rightarrow \quad \hat{v} = 0 \quad \forall \hat{v} \in \mathcal{Q}_k(\hat{K}).$$

Then the inequality (4.8b) follows immediately from Lemma 2.14 in [1]. The inequalities (4.8c), (4.8e) and (4.8f) can be verified by taking similar linear functionals and applying Lemma 2.14 in [1]. \square

Using scaling arguments on an element K (or an edge e), we obtain the interpolation error estimates on K (or e) with respect to its size. All these inequalities will be used in estimating ξ_q and ξ_u .

Lemma 4.3. *If $v \in H^{k+1}(\Omega)$ and $\mathbf{r} \in (H^{k+1}(\Omega))^2$, then there exists constant C such that for any $K \in \mathcal{T}_N$,*

$$(4.10a) \quad \|v - Iv\|_K \leq C \sum_{|\alpha|=1} \bar{h}^{(k+1)\alpha} \|D^{(k+1)\alpha} v\|_K,$$

$$(4.10b) \quad \|D^\alpha(v - Iv)\|_K \leq C \sum_{|\beta|=1} \bar{h}^{k\beta} \|D^{k\beta+\alpha} v\|_K, \quad |\alpha| = 1,$$

$$(4.10c) \quad \|v - \Pi^+ v\|_K \leq C \sum_{|\alpha|=1} \bar{h}^{(k+1)\alpha} \|D^{(k+1)\alpha} v\|_K,$$

$$(4.10d) \quad \|v - \Pi^+ v\|_e \leq Ch_{x,K}^{k+1} \|D^{(k+1,0)} v\|_e, \quad e = \gamma_1^+ \in \partial K,$$

$$(4.10e) \quad \|v - \Pi^+ v\|_e \leq Ch_{y,K}^{k+1} \|D^{(0,k+1)} v\|_e, \quad e = \gamma_2^+ \in \partial K,$$

$$(4.10f) \quad \|v - \Pi^+ v\|_e \leq C|e|^{\frac{1}{2}}|K|^{-\frac{1}{2}} \sum_{\alpha \in \Lambda} \bar{h}^\alpha \|D^\alpha v\|_K, \quad e \in \{\gamma_1^-, \gamma_2^-\} \subset \partial K,$$

$$(4.10g) \quad \|\mathbf{r} - \Pi^- \mathbf{r}\|_K \leq C \sum_{|\alpha|=1} \bar{h}^{(k+1)\alpha} \|D^{(k+1)\alpha} \mathbf{r}\|_K,$$

$$(4.10h) \quad \|\mathbf{r} - \Pi^- \mathbf{r}\|_e \leq C|e|^{\frac{1}{2}}|K|^{-\frac{1}{2}} \sum_{\alpha \in \Lambda} \bar{h}^\alpha \|D^\alpha \mathbf{r}\|_K, \quad e \in \{\gamma_1^+, \gamma_2^+\} \subset \partial K,$$

where $h_{x,K} = \frac{1}{2}|\gamma_1^\pm|$ and $h_{y,K} = \frac{1}{2}|\gamma_2^\pm|$. Here we used the notation $\bar{h} = (h_{x,K}, h_{y,K})$, and the multi-index set $\Lambda = \{(k+1, 0), (k, 1), (1, k), (0, k+1)\}$.

Proof. The error estimates of standard Lagrange interpolation on anisotropic Cartesian grids (4.10a)-(4.10b) have been proved in [1]. The inequalities (4.10c) and (4.10g) directly follow (4.8b), (4.8e) and the transformation $x = x_K + h_{x,K}\xi/2$, $y = y_K + h_{y,K}\eta/2$ and the scaling arguments. The inequalities (4.10d) and (4.10e) were proved in Lemma 3.3, [5]. For (4.10f) and (4.10h), we apply the trace inequality, (4.8b)-(4.8c), and (4.8e)-(4.8f) on the reference element to obtain

$$\begin{aligned} \|\xi_{\hat{v}}\|_{\hat{\gamma}_1^- \cup \hat{\gamma}_2^-} &\leq C \|\xi_{\hat{v}}\|_{H^1(\hat{K})} \leq C \sum_{|\alpha|=1} \left[\hat{D}^\alpha \hat{v} \right]_{H^k(\hat{K})}, \\ \|\xi_{\hat{\mathbf{r}}}\|_{\hat{\gamma}_1^+ \cup \hat{\gamma}_2^+} &\leq C \|\xi_{\hat{\mathbf{r}}}\|_{H^1(\hat{K})^2} \leq C \sum_{|\alpha|=1} \left[\hat{D}^\alpha \hat{\mathbf{r}} \right]_{H^k(\hat{K})^2}. \end{aligned}$$

Then, (4.10f) and (4.10h) follows the standard scaling arguments. □

The next two lemmas collect several results that involve the approximants of the various terms in the decomposition (2.10) and (2.12). Briefly, in the proof, u and q will be separated according to the regularity properties of each component on each subdomain of Ω . The purpose of employing this technique is to use the exponential decreasing properties of w and to avoid any negative power of singular parameter ϵ . This technique will be used again in the proof of Lemma 4.9.

Lemma 4.4. *Let (\mathbf{q}, u) be the solution of (2.1) that satisfies the regularity (2.13) and (2.11). There exists a constant C , independent of ϵ and N , such that*

$$(4.11a) \quad \|\nabla \xi_{\bar{u}}\| \leq CN^{-k},$$

$$(4.11b) \quad \|\xi_u\| \leq C \left[\frac{1}{N^{k+1}} + \sqrt{\epsilon} \left(\frac{\ln N}{N} \right)^{k+1} \right],$$

$$(4.11c) \quad \|\xi_{u^-}\|_{\mathcal{E} \setminus \mathcal{E}^B} \leq C \left[\frac{1}{N^{k+\frac{1}{2}}} + \sqrt{\epsilon \ln N} \left(\frac{\ln N}{N} \right)^{k+\frac{1}{2}} \right],$$

$$(4.11d) \quad \epsilon^{-\frac{1}{2}} \|\xi_{\mathbf{q}}\| \leq C \left(\frac{\ln N}{N} \right)^{k+1}.$$

Proof. (1) The first inequality (4.11a) is a standard approximation result of Lagrange interpolation on an anisotropic rectangle element that was proved in [1].

(2) Consider $\|\xi_u\|$. For each component of $u = \bar{u} + w_1 + w_2 + w_0$ on each subdomain of Ω , we have

$$(4.12) \quad \|\xi_u\|_{\mathcal{T}_N}^2 = \sum_{K \in \mathcal{T}_N} \|\xi_u\|_K^2 = C \sum_{s \in \{0, x, y, xy\}} \sum_{\varphi \in \{\bar{u}, w_1, w_2, w_0\}} X_s(\varphi)$$

where $X_s(\varphi)$ represents the L^2 error of ξ_φ on each subdomain of Ω , i.e., Ω_0 , Ω_x , Ω_y , and Ω_{xy} . First, it follows directly from the regularity of \bar{u} and (4.10a) that

$$(4.13) \quad \begin{aligned} \|\xi_{\bar{u}}\|_{\Omega_0}^2 &= X_0(\bar{u}) \leq CH^{2(k+1)}, \\ \|\xi_{\bar{u}}\|_{\Omega \setminus \Omega_0}^2 &= \sum_{s \in \{x, y, xy\}} X_s(\bar{u}) \leq C\epsilon H^{2(k+1)}. \end{aligned}$$

Second, we have to find the upper bounds of $X_0(w)$, $X_{xy}(w)$, $X_x(w_1)$, $X_x(w_2)$, and $X_x(w_0)$. Because of the symmetry, $X_y(w_2)$, $X_y(w_1)$, and $X_y(w_0)$ will have the same bounds as $X_x(w_1)$, $X_x(w_2)$, and $X_x(w_0)$, respectively. Recalling the definition of Shishkin mesh, the mesh is uniform on Ω_0 and Ω_{xy} . Therefore, by the inverse inequality (4.2), Lemma 4.3, and (2.11), one has

$$\begin{aligned} X_0(w) &= \|w\|_{\Omega_0}^2 + \|\Pi^+ w\|_{\Omega_0}^2 \leq C\epsilon e^{2\beta_1 \tau_x / \epsilon} + \|w\|_{C^0(\bar{\Omega}_0)}^2 \leq CH^{2(k+1)}, \\ X_{xy}(w) &= \|\xi_w\|_{\Omega_{xy}}^2 \leq Ch^{2(k+1)} \left(\|D^{(k+1,0)} w\|_{\Omega_{xy}}^2 + \|D^{(0,k+1)} w\|_{\Omega_{xy}}^2 \right) \\ &\leq Ch^{2(k+1)} \epsilon^{-2k} \leq C\epsilon^2 (\ln N / N)^{2(k+1)}, \\ X_x(w_1) &= \|\xi_{w_1}\|_{\Omega_x}^2 \leq Ch^{2(k+1)} \|D^{(k+1,0)} w_1\|_{\Omega_x}^2 + CH^{2(k+1)} \|D^{(0,k+1)} w_1\|_{\Omega_x}^2 \\ &\leq Ch^{2(k+1)} \epsilon^{-2k-1} + CH^{2(k+1)} \epsilon \leq C\epsilon (\ln N / N)^{2(k+1)}, \\ X_x(w_2) &= \|w_2\|_{\Omega_x}^2 + \|\Pi^+ w_2\|_{\Omega_x}^2 \leq C\epsilon^2 e^{2\beta_2 \tau_y / \epsilon} + \|w_2\|_{C^0(\bar{\Omega}_0)}^2 \leq CH^{2(k+1)}, \\ X_x(w_0) &= \|w_0\|_{\Omega_x}^2 + \|\Pi^+ w_0\|_{\Omega_x}^2 \leq C\epsilon^2 e^{2\beta_2 \tau_y / \epsilon} + \|w_0\|_{C^0(\bar{\Omega}_0)}^2 \leq CH^{2(k+1)}. \end{aligned}$$

Substituting above inequalities and (4.13) into (4.12) yields

$$(4.14) \quad \begin{aligned} \|\xi_w\|_{\Omega_0} &\leq CN^{-(k+1)}, & \|\xi_{\bar{u}}\|_{\Omega_0} &\leq CN^{-(k+1)}, \\ \|\xi_w\|_{\mathcal{T}_N \setminus \Omega_0} &\leq C\sqrt{\epsilon} \left(\frac{\ln N}{N} \right)^{k+1}, & \|\xi_{\bar{u}}\|_{\mathcal{T}_N \setminus \Omega_0} &\leq C \frac{\sqrt{\epsilon}}{N^{k+1}}. \end{aligned}$$

The combination of the above inequalities gives rise to (4.11b).

(3) As for $\|\xi_{u^-}\|_{\mathcal{E}\setminus\mathcal{E}_-^B}$, we have

$$(4.15) \quad \|\xi_{u^-}\|_{\mathcal{E}\setminus\mathcal{E}_-^B}^2 = \|\xi_{u^-}\|_{\Theta^v\setminus\mathcal{E}_-^B}^2 + \|\xi_{u^-}\|_{\Theta^h\setminus\mathcal{E}_-^B}^2.$$

By the symmetric property of the regularity of u , we only consider the error estimates on $\Theta^v \setminus \mathcal{E}_-^B$. For any $\varphi \in \{\bar{u}, w_1\}$, applying (4.10d)-(4.10e) yields

$$(4.16) \quad \begin{aligned} \|\xi_{\varphi^-}\|_{\Theta^v\setminus\mathcal{E}_-^B}^2 &\leq C \sum_{K \in \mathcal{T}_N} h_{y,k}^{2(k+1)} \left\| D^{(0,k+1)}\varphi \right\|_{\gamma_2^+}^2 \\ &\leq C \sum_{e \in \Theta_3^v\setminus\mathcal{E}_-^B} H^{2(k+1)} \left\| D^{(0,k+1)}\varphi \right\|_e^2 + C \sum_{e \in \Theta_4^v\setminus\mathcal{E}_-^B} h^{2(k+1)} \left\| D^{(0,k+1)}\varphi \right\|_e^2 \\ &\leq CH^{2k+1}. \end{aligned}$$

The estimates of $\xi_{w_1^-}$ and $\xi_{w_0^-}$ can be proved as follows. Lemma 3.7 in [17] states that the one-dimensional operator π^+ satisfies $\|\pi^+\hat{\varphi}\|_{\hat{I}}^2 \leq C\|\hat{\varphi}\|_{\hat{I}} + C|\hat{\varphi}(1)|^2$ with $\hat{I} = (-1, 1)$. Therefore, by applying scaling argument to this inequality and by using (4.10d)-(4.10e), and (2.11) one has

$$\begin{aligned} \|\xi_{\varphi^-}\|_{\Theta^v\setminus\mathcal{E}_-^B}^2 &\leq C \sum_{K \in \Omega_0 \cup \Omega_x} \left(\|\varphi\|_{\gamma_2^+}^2 + \|\pi^+\varphi\|_{\gamma_2^+}^2 \right) + \sum_{K \in \Omega_y \cup \Omega_{xy}} \|\xi_{\varphi}\|_{\gamma_2^+}^2 \\ &\leq C \left[\|\varphi\|_{\Theta_3^v\setminus\mathcal{E}_-^B}^2 + N \sum_{j=1}^N H e^{-2\beta_2(1-y_j)/\epsilon} + \sum_{e \in \Theta_4^v\setminus\mathcal{E}_-^B} h^{2(k+1)} \left\| D^{(0,k+1)}\varphi \right\|_e^2 \right] \\ &\leq C \left[N^{-(2k+1)} + \epsilon \ln N \left(\frac{\ln N}{N} \right)^{2k+1} \right], \end{aligned}$$

for any $\varphi \in \{w_2, w_0\}$. Substituting the above inequality and (4.16) into (4.15) gives rise to (4.11c).

(4) Using similar arguments as in (2) and the regularity for each component \mathbf{q} in (2.13) we can prove (4.11d). □

Lemma 4.5. *Let (\mathbf{q}, u) be the solution of (2.1) that satisfies the regularity (2.11). Then there exists a constant C , independent of ϵ and N , such that*

$$(4.17a) \quad \|\llbracket \xi_w \rrbracket\|_{\mathcal{E}} = \|\llbracket \xi_u \rrbracket\|_{\mathcal{E}} \leq C \left(\frac{\ln N}{N} \right)^{k+\frac{1}{2}},$$

$$(4.17b) \quad \|\xi_{\mathbf{q}}\|_{\mathcal{E}_+^B} \leq C \left(\frac{\ln N}{N} \right)^{k+1},$$

$$(4.17c) \quad \|\xi_u\|_{\mathcal{E}_+^B} \leq C \left(\frac{\ln N}{N} \right)^{k+1}.$$

Proof. (1) Consider $\|\llbracket \xi_u \rrbracket\|_{\mathcal{E}}$. According to the definition of the jump $\llbracket \xi_u \rrbracket$, we have

$$\|\llbracket \xi_u \rrbracket\|_{\mathcal{E}}^2 \leq 2 \sum_{e \in \mathcal{E}} \left(\|\xi_{u^+}\|_{e \in K^+}^2 + \|\xi_{u^-}\|_{e \in K^-}^2 \right) = 2 \left(\|\xi_{u^+}\|_{\mathcal{E}\setminus\mathcal{E}_+^B}^2 + \|\xi_{u^-}\|_{\mathcal{E}\setminus\mathcal{E}_-^B}^2 \right).$$

Since $\|\xi_{u-}\|_{\mathcal{E} \setminus \mathcal{E}_-^B}^2$ have already been bounded in (4.11c), we only need to estimate $\|\xi_{u+}\|_{\mathcal{E} \setminus \mathcal{E}_+^B}^2$, which could be divided into several parts:

$$\begin{aligned} \|\xi_{u+}\|_{\mathcal{E} \setminus \mathcal{E}_+^B}^2 &= \sum_{K \in \mathcal{T}_N} \sum_{e \in \{\gamma_1^-, \gamma_2^-\}} \|\xi_u\|_e^2 \\ &= C \sum_{s \in \{0, x, y, xy\}} \sum_{\varphi \in \{\bar{u}, w_1, w_2, w_0\}} Y_s(\varphi), \end{aligned}$$

where $Y_s(\varphi)$ represents the above error expression of $\|\xi_\varphi\|_e$ on each subset of \mathcal{E} , i.e., $\Theta_0, \Theta_x, \Theta_y$, and Θ_{xy} .

First, it follows directly from the regularity of \bar{u} and approximation theory of 1-D Lagrange interpolation that

$$\begin{aligned} \sum_{s \in \{0, x, y, xy\}} Y_s(\bar{u}) &\leq CH^{2(k+1)} \left(\|D^{(k+1,0)}\bar{u}\|_{\Theta_h}^2 + \|D^{(0,k+1)}\bar{u}\|_{\Theta_v}^2 \right) \\ &\leq CN^{-(2k+1)}. \end{aligned}$$

Second, we are going to find the upper bounds of $Y_0(w), Y_{xy}(w), Y_x(w_1), Y_x(w_2)$, and $Y_x(w_0)$. Because of the symmetry, $Y_y(w_2), Y_y(w_1)$, and $Y_y(w_0)$ have the same bounds as $Y_x(w_1), Y_x(w_2)$, and $Y_x(w_0)$, respectively. Recalling the definition of Shishkin mesh, the mesh is uniform on Ω_0 and Ω_{xy} . Therefore, by (4.8a), (4.1), (2.11), and (4.4) one has

$$\begin{aligned} Y_0(w) &\leq \sum_{e \in \Theta_0} \|w\|_e^2 + C \sum_{K \subset \Omega_0} \left(h_{y,K}^{-1} + h_{x,K}^{-1} \right) \|\Pi^+ w\|_K^2 \\ &\leq CN \left(\int_0^{1-\tau_x} e^{-2\beta_1(1-x)/\epsilon} dx + \int_0^{1-\tau_y} e^{-2\beta_2(1-y)/\epsilon} dy \right) + CH^{-1} \|\Pi^+ w\|_{\Omega_0}^2 \\ &\leq CN \left(e^{-2\beta_1\tau_x/\epsilon} + e^{-2\beta_2\tau_y/\epsilon} \right) + CH^{-1} \max \left\{ e^{-2\beta_1\tau_x/\epsilon}, e^{-2\beta_2\tau_y/\epsilon} \right\} \\ &\leq CN^{-(2k+1)}. \end{aligned}$$

By (4.10f) and (2.11) one has

$$\begin{aligned} Y_{xy}(w) &\leq C \sum_{K \subset \Omega_{xy}} \frac{h_{x,K} + h_{y,K}}{|K|} \sum_{\alpha \in \Lambda} \tilde{h}^{2\alpha} \|D^\alpha w\|_K^2 \\ &\leq Ch^{2k+1} \sum_{\alpha \in \Lambda} \|D^\alpha w\|_{\Omega_{xy}}^2 \\ &\leq C\epsilon (\ln N/N)^{2k+1}, \end{aligned}$$

where $\Lambda = \{(k+1, 0), (k, 1), (1, k), (0, k+1)\}$. Similarly, on Θ_x one has

$$\begin{aligned} Y_x(w_1) &\leq C \sum_{K \subset \Omega_x} \frac{h_{x,K} + h_{y,K}}{|K|} \sum_{\alpha \in \Lambda} \tilde{h}^{2\alpha} \|D^\alpha w_1\|_K^2 \\ &\leq Ch^{-1} h^{k+1} \left\| D^{(k+1,0)} w_1 \right\|_{\Omega_x}^2 + Ch^{-1} H h^k \left\| D^{(k,1)} w_1 \right\|_{\Omega_x}^2 \\ &\quad + Ch^{-1} H^k h \left\| D^{(1,k)} w_1 \right\|_{\Omega_x}^2 + Ch^{-1} H^{k+1} \left\| D^{(0,k+1)} w_1 \right\|_{\Omega_x}^2 \\ &\leq C (\ln N/N)^{2k+1}. \end{aligned}$$

To estimate ξ_{w_2} on Θ_x , we use (4.8a), (4.1), (2.11), and (4.4) and arrive at

$$\begin{aligned} Y_x(w_2) &\leq \sum_{\epsilon \in \Theta_x} \|w_2\|_\epsilon^2 + C \sum_{K \subset \Omega_x} \left(h_{y,K}^{-1} + h_{x,K}^{-1} \right) \|\Pi^+ w_2\|_K^2 \\ &\leq C \left(\epsilon \sum_{i=1}^N e^{-2\beta_2(1-y_i)/\epsilon} + N \int_0^{1-\tau_y} e^{-2\beta_2(1-y)/\epsilon} dy \right) + Ch^{-1} \|\Pi^+ w_2\|_{\Omega_x}^2 \\ &\leq C \left(\epsilon N^{-2\kappa} + N e^{-2\beta_2\tau_y/\epsilon} \right) + Ch^{-1}\epsilon \|w_2\|_{C^0(\bar{\Omega}_x)}^2 \\ &\leq C \left(\epsilon N^{-2\kappa} + N^{-2\kappa+1} \right) + Ch^{-1}\epsilon e^{-2\beta_2\tau_y/\epsilon} \\ &\leq CN^{-(2\kappa+1)}. \end{aligned}$$

Using a similar argument, we can show that the bound of $Y_x(w_0)$ is exactly the same as $Y_x(w_2)$. Then, (4.17a) follows the combination of the above error estimates.

(2) To prove (4.17b) we separate the boundary set \mathcal{E}_+^B into \mathcal{E}_1^B and \mathcal{E}_2^B as defined in Section 4.1. We only consider $\xi_{\mathbf{q}}$ on \mathcal{E}_1^B . The proof on \mathcal{E}_2^B is omitted due to the symmetry property. Let $\Omega_1 = (0, 1 - \tau_x) \times (1 - h_y, 1)$, $\Omega_2 = (1 - \tau_x, 1) \times (1 - h_y, 1)$. It follows from (4.10h) that

$$(4.18) \quad \|\xi_{\mathbf{q}}\|_{\mathcal{E}_1^B \cap \Omega_2}^2 \leq Ch^{2k+1} \sum_{|\alpha| \in \Lambda} \|D^\alpha \mathbf{q}\|_{\Omega_2}^2 \leq C\epsilon(\ln N/N)^{2(k+1)}.$$

Recalling the decomposition of \mathbf{q} in (2.12) and the regularity (2.13), we apply (4.10h) for $\mathbf{p} \in \{\bar{\mathbf{q}}, \mathbf{q}_2\}$,

$$\begin{aligned} (4.19) \quad \|\xi_{\mathbf{p}}\|_{\mathcal{E}_1^B \cap \Omega_1}^2 &\leq CH^{2(k+1)}h^{-1} \|D^{(k+1,0)} \mathbf{p}\|_{\Omega_1}^2 + CH^{2k}h \|D^{(k,1)} \mathbf{p}\|_{\Omega_1}^2 \\ &\quad + CH^2h^{2k-1} \|D^{(1,k)} \mathbf{p}\|_{\Omega_1}^2 + Ch^{2k+1} \|D^{(0,k+1)} \mathbf{p}\|_{\Omega_1}^2 \\ &\leq C(\ln N/N)^{2(k+1)}. \end{aligned}$$

On the other hand, by (4.3) and (4.8d) and the fact $\mathbf{q}_0 = \epsilon \nabla w_0$ we obtain that the interpolation on the boundary of the reference element satisfies

$$\begin{aligned} \|\xi_{\hat{\mathbf{q}}_0}\|_{\hat{\gamma}_1^-}^2 &\leq C \|\xi_{\hat{\mathbf{q}}_0}\|_{H^1(\hat{K})^2}^2 \leq C \left(\|\hat{\mathbf{q}}_0\|_{H^1(\hat{K})^2}^2 + \|\Pi^- \hat{\mathbf{q}}_0\|_{H^1(\hat{K})^2}^2 \right) \\ &\leq C \|\hat{\mathbf{q}}_0\|_{H^1(\hat{K})^2}^2 \\ &\leq C\epsilon^2 \|\hat{w}_0\|_{C(\bar{K})}^2. \end{aligned}$$

Similarly, we have $\|\xi_{\hat{\mathbf{q}}_1}\|_{\hat{\gamma}_1^-}^2 \leq C\epsilon^2 \|\hat{w}_1\|_{C(\bar{K})}^2$. Therefore, for $\mathbf{p} \in \{\mathbf{q}_0, \mathbf{q}_1\}$ we use the above inequality and the regularity (2.13) to get that

$$\begin{aligned} (4.20) \quad \|\xi_{\mathbf{p}}\|_{\mathcal{E}_1^B \cap \Omega_1}^2 &\leq CH\epsilon^2 \max\{\|\hat{w}_0\|_{C(\Omega_1)}^2, \|\hat{w}_1\|_{C(\Omega_1)}^2\} \\ &\leq CH\epsilon^2 e^{-2\beta_1\tau_x/\epsilon} \\ &\leq C\epsilon^2 N^{-(2k+3)}. \end{aligned}$$

The combination of (4.18)–(4.20) establishes (4.17b).

(3) The inequality (4.17c) can be proved in a similar way to the proof in (2). \square

4.4. **The stability of $\mathcal{A}(\phi, \varphi; \mathbf{r}, v)$.** To estimate (η_q, η_u) , we utilize a stability result of the compact form $\mathcal{A}(\phi, \varphi; \mathbf{r}, v)$ in another mesh-dependent energy norm $\|(\mathbf{r}, v)\|_{\mathcal{A}}$, which is defined as

$$(4.21) \quad \|(\mathbf{r}, v)\|_{\mathcal{A}} := \left[\epsilon^{-1} \|\mathbf{r}\|^2 + \left(H \|\boldsymbol{\beta} \cdot \nabla v\|_{\Omega_0}^2 + h \|\nabla v\|_{\mathcal{T}_N \setminus \Omega_0}^2 \right) + c(\psi, \psi) \right]^{\frac{1}{2}}$$

for $(\mathbf{r}, v) \in [L^2(\mathcal{T}_N)]^2 \times H^1(\mathcal{T}_N)$. This norm is clearly stronger than $|(\cdot, \cdot)|_{\mathcal{A}}$.

The stability in $|(\mathbf{r}, v)|_{\mathcal{A}}$ is not sufficient to estimate $\mathcal{A}(\xi_{\mathbf{q}}, \xi_u; r, v)$ because it comprises one term $(\xi_u, \boldsymbol{\beta} \cdot \nabla v)$ (or $(\boldsymbol{\beta} \cdot \nabla \xi_u, v)$). We should avoid estimating $\|\nabla \xi_u\|$ because it doesn't have a ϵ -uniform upper bound. However, if we start to estimate $(\xi_u, \boldsymbol{\beta} \cdot \nabla v)$ by considering

$$(\xi_u, \boldsymbol{\beta} \cdot \nabla v) \leq \|\xi_u\|_{\Omega_0} \|\boldsymbol{\beta} \cdot \nabla v\|_{\Omega_0} + \|\xi_u\|_{\Omega \setminus \Omega_0} \|\boldsymbol{\beta} \cdot \nabla v\|_{\Omega \setminus \Omega_0},$$

it seems that we can use inverse inequality to obtain $CH^{-1} \|v\|_{\Omega_0}$ and $Ch^{-1} \|v\|_{\Omega \setminus \Omega_0}$, which could be bounded by $Ch^{-1} |(\mathbf{r}, v)|_{\mathcal{A}}$. (Note that $|(\mathbf{r}, v)|_{\mathcal{A}}$ does not contain gradient terms.) However, it is unlikely to get a uniform convergence rate on $\Omega \setminus \Omega_0$ because $\|\xi_u\|_{\Omega \setminus \Omega_0}$ is not small enough to cancel out ϵ^{-1} .

On the other hand, if we use the stability in norm $\|(\cdot, \cdot)\|_{\mathcal{A}}$, these two gradient terms $\|\boldsymbol{\beta} \cdot \nabla v\|_{\Omega_0}$ and $\|\boldsymbol{\beta} \cdot \nabla v\|_{\Omega \setminus \Omega_0}$ can be absorbed by $\|(\mathbf{r}, v)\|_{\mathcal{A}}$. This is the reason why we need the following stability result.

Lemma 4.6 (Stability). *If $0 \leq C_{11} \leq O(1)$ on \mathcal{E} for some constant C independent ϵ and N , there exists a positive constant C_D such that the estimate*

$$(4.22) \quad \sup_{(\mathbf{r}, v) \in \mathbf{V}_N \times M_N} \frac{\mathcal{A}(\phi, \psi; \mathbf{r}, v)}{\|(\mathbf{r}, v)\|_{\mathcal{A}}} \geq C_D \|(\phi, \psi)\|_{\mathcal{A}}$$

holds for any $(\phi, \psi) \in \mathbf{V}_N \times M_N$. Here C_D is independent of ϵ and N .

Proof. (1) We start by testing the bilinear form \mathcal{A} defined in (2.5) with $(\mathbf{r}, v) = (\phi, \psi)$. It is straightforward to see that

$$(4.23) \quad \mathcal{A}(\phi, \psi; \phi, \psi) = \epsilon^{-1} \|\phi\|^2 + c(\psi, \psi).$$

(2) Next, we choose $(\mathbf{r}, v) = (-\gamma h \nabla \psi, \frac{\gamma h}{\epsilon} \psi)$ for some positive constant γ , it follows from the definition of the compact form (2.5) and (2.4) that

$$(4.24) \quad \mathcal{A}(\phi, \psi; \mathbf{r}, v) = a(\phi, -\gamma h \nabla \psi) - b_2(\psi, -\gamma h \nabla \psi) + b_2\left(\frac{\gamma h}{\epsilon} \psi, \phi\right) + c\left(\psi, \frac{\gamma h}{\epsilon} \psi\right),$$

where

$$\begin{aligned} a(\phi, -\gamma h \nabla \psi) &= \epsilon^{-1} (\phi, -\gamma h \nabla \psi), \\ -b_2(\psi, -\gamma h \nabla \psi) &= \gamma h \|\nabla \psi\|_{\mathcal{T}_N}^2 - \left\langle \gamma h (\nabla \psi)^+, [\psi] \right\rangle_{\mathcal{E}^0} - \langle \gamma h \nabla \psi, \psi \rangle_{\mathcal{E}^B}, \\ b_2(\gamma h \epsilon^{-1} \psi, \phi) &= \epsilon^{-1} (\gamma h \nabla \psi, \phi) - \langle \phi^+, \gamma h \epsilon^{-1} [\psi] \rangle_{\mathcal{E}^0} - \langle \phi, \gamma h \epsilon^{-1} \psi \rangle_{\mathcal{E}^B}, \\ c\left(\psi, \frac{\gamma h}{\epsilon} \psi\right) &= \gamma h \epsilon^{-1} c(\psi, \psi). \end{aligned}$$

For $b_2(\psi, -\gamma h \nabla \psi)$, applying the Schwarz inequality and (4.1) yields

$$- \left\langle \gamma h (\nabla \psi)^+, [\psi] \right\rangle_{\mathcal{E}^0} - \langle \gamma h \nabla \psi, \psi \rangle_{\mathcal{E}^B} \geq -C \gamma^2 h \|\nabla \psi\|^2 - \frac{1}{4} \left\| |\boldsymbol{\beta} \cdot \mathbf{n}|^{1/2} [\psi] \right\|_{\mathcal{E}}^2.$$

Similarly, for $b_2(\gamma h \epsilon^{-1} \psi, \phi)$ we have

$$- \langle \phi^+, \gamma h \epsilon^{-1} [\psi] \rangle_{\mathcal{E}^0} - \langle \phi, \gamma h \epsilon^{-1} \psi \rangle_{\mathcal{E}^B} \geq -\frac{1}{4\epsilon} \|\phi\|^2 - C \frac{\gamma^2 h^2}{\epsilon} \left\| |\boldsymbol{\beta} \cdot \mathbf{n}|^{1/2} [\psi] \right\|_{\mathcal{E}}^2.$$

Substituting the above two inequalities into (4.24) yields

$$\begin{aligned} \mathcal{A}(\phi, \psi; -\gamma h \nabla \psi, \gamma h \epsilon^{-1} \psi) &\geq \gamma h (1 - C\gamma) \|\nabla \psi\|^2 - \frac{1}{4\epsilon} \|\phi\|^2 \\ &\quad - \left[\frac{1}{4} - \gamma h \epsilon^{-1} (1 - C\gamma h) \right] \|\beta \cdot \mathbf{n}\|^{1/2} \|\psi\|_{\mathcal{E}}^2. \end{aligned}$$

Here constant C depends on $\min\{\beta_1, \beta_2\}$. Choosing sufficiently small constant γ such that $0 < C_d \leq \gamma(1 - C\gamma) \leq C$ and $(1 - C\gamma h) > 0$, we get

$$(4.25) \quad \mathcal{A}(\phi, \psi; -\gamma h \nabla \psi, \gamma h \epsilon^{-1} \psi) \geq C_d h \|\nabla \psi\|^2 - \frac{1}{4\epsilon} \|\phi\|^2 - \frac{1}{4} c(\psi, \psi).$$

(3) Furthermore, let $\bar{\beta}$ be a piecewise constant vector function defined by

$$\bar{\beta}|_K = \frac{1}{|K|} \int_K \beta dx dy, \quad \forall K \in \mathcal{T}_N.$$

Set $(\mathbf{r}, v) = (0, \gamma \delta \bar{\beta} \cdot \nabla \psi)$ for some positive constant γ and

$$\delta = \begin{cases} H, & \text{on } \Omega_0, \\ 0, & \text{on } \mathcal{T}_N \setminus \Omega_0. \end{cases}$$

From the definition of the compact form (2.5), one has

$$(4.26) \quad \mathcal{A}(\phi, \psi; \mathbf{r}, v) = b_2(\gamma \delta \bar{\beta} \cdot \nabla \psi, \phi) + c(\psi, \gamma \delta \bar{\beta} \cdot \nabla \psi),$$

where

$$\begin{aligned} b_2(\gamma \delta \bar{\beta} \cdot \nabla \psi, \phi) &= (\gamma \delta \nabla(\bar{\beta} \cdot \nabla \psi), \phi) - \langle \phi^+, \llbracket \gamma \delta \bar{\beta} \cdot \nabla \psi \rrbracket \rangle_{\mathcal{E}^0} - \langle \phi, \gamma \delta \bar{\beta} \cdot \nabla \psi \rangle_{\mathcal{E}^B}, \\ c(\psi, \gamma \delta \bar{\beta} \cdot \nabla \psi) &= (\beta \cdot \nabla \psi, \gamma \delta \bar{\beta} \cdot \nabla \psi) + (\psi, c\gamma \delta \bar{\beta} \cdot \nabla \psi) \\ &\quad + \langle C_{11} \llbracket \psi \rrbracket, \llbracket \gamma \delta \bar{\beta} \cdot \nabla \psi \rrbracket \rangle_{\mathcal{E}} - \langle \beta \llbracket \psi \rrbracket, (\gamma \delta \bar{\beta} \cdot \nabla \psi)^+ \rangle_{\mathcal{E} \setminus \mathcal{E}^B}. \end{aligned}$$

Next, we are going to deal with each term of $b_2(\gamma \delta \bar{\beta} \cdot \nabla \psi, \phi)$ and $c(\psi, \gamma \delta \bar{\beta} \cdot \nabla \psi)$. Applying the Schwarz inequality, (4.1) and (4.2) to $b_2(\gamma \delta \bar{\beta} \cdot \nabla \psi, \phi)$ yields

$$\begin{aligned} (\gamma \delta \nabla(\bar{\beta} \cdot \nabla \psi), \phi) &\geq -\frac{1}{8\epsilon} \|\phi\|_{\Omega_0}^2 - 2\gamma^2 H^2 \epsilon \|\nabla(\bar{\beta} \cdot \nabla \psi)\|_{\Omega_0}^2 \\ &\geq -\frac{1}{8\epsilon} \|\phi\|_{\Omega_0}^2 - 2\gamma^2 \epsilon \|\bar{\beta} - \beta\| \cdot \nabla \psi\|_{\Omega_0}^2 - 2\gamma^2 \epsilon \|\beta \cdot \nabla \psi\|_{\Omega_0}^2 \\ &\geq -\frac{1}{8\epsilon} \|\phi\|_{\Omega_0}^2 - 2\gamma^2 \epsilon \left(1 + CH^2 \|\beta\|_{1,\infty}^2\right) \|\beta \cdot \nabla \psi\|_{\Omega_0}^2, \end{aligned}$$

and

$$\begin{aligned} -\langle \phi^+, \gamma \delta \llbracket \bar{\beta} \cdot \nabla \psi \rrbracket \rangle_{\mathcal{E}^0} - \langle \phi, \gamma \delta \bar{\beta} \cdot \nabla \psi \rangle_{\mathcal{E}^B} &\geq -\left(\|\phi^+\|_{\mathcal{E}^0 \cap \Omega_0} + \|\phi\|_{\mathcal{E}^B \cap \Omega_0}\right) \|\llbracket \gamma \delta \bar{\beta} \cdot \nabla \psi \rrbracket\|_{\mathcal{E} \cap \Omega_0} \\ &\geq -C\gamma \|\phi\|_{\Omega_0} \cdot \|\bar{\beta} \cdot \nabla \psi\|_{\Omega_0} \\ &\geq -\frac{1}{8\epsilon} \|\phi\|_{\Omega_0}^2 - C\gamma^2 \epsilon \left[\|\beta \cdot \nabla \psi\|_{\Omega_0}^2 + \|\bar{\beta} - \beta\| \cdot \nabla \psi\|_{\Omega_0}^2\right] \\ &\geq -\frac{1}{8\epsilon} \|\phi\|_{\Omega_0}^2 - C\gamma^2 \epsilon \left(1 + CH^2 \|\beta\|_{1,\infty}^2\right) \|\beta \cdot \nabla \psi\|_{\Omega_0}^2. \end{aligned}$$

Similarly, for $c(\psi, \gamma\delta\bar{\beta} \cdot \nabla\psi)$ we have

$$\begin{aligned} (\beta \cdot \nabla\psi, \gamma\delta\bar{\beta} \cdot \nabla\psi) &= \gamma H \|\beta \cdot \nabla\psi\|_{\Omega_0}^2 + \gamma\delta(\beta \cdot \nabla\psi, (\bar{\beta} - \beta) \cdot \nabla\psi) \\ &\geq \gamma H \|\beta \cdot \nabla\psi\|_{\Omega_0}^2 - C\gamma H^2 \|\beta\|_{1,\infty} \|\beta \cdot \nabla\psi\|_{\Omega_0}^2 \\ &\geq \gamma H \left(1 - CH \|\beta\|_{1,\infty}\right) \|\beta \cdot \nabla\psi\|_{\Omega_0}^2, \end{aligned}$$

$$\begin{aligned} (\psi, c\gamma\delta\bar{\beta} \cdot \nabla\psi) &\geq -\frac{1}{4} \|\sqrt{c}\psi\|_{\Omega_0}^2 - C\gamma^2 H^2 \|\bar{\beta} \cdot \nabla\psi\|_{\Omega_0}^2 \\ &\geq -\frac{1}{4} \|\sqrt{c}\psi\|_{\Omega_0}^2 - C\gamma^2 H^2 \left[\|\beta \cdot \nabla\psi\|_{\Omega_0}^2 + \|(\bar{\beta} - \beta) \cdot \nabla\psi\|_{\Omega_0}^2\right] \\ &\geq -\frac{1}{4} \|\sqrt{c}\psi\|_{\Omega_0}^2 - C\gamma^2 H^2 \left(1 + CH^2 \|\beta\|_{1,\infty}^2\right) \|\beta \cdot \nabla\psi\|_{\Omega_0}^2, \end{aligned}$$

and by trace inequality,

$$\begin{aligned} &\langle C_{11}[\psi], [\gamma\delta\bar{\beta} \cdot \nabla\psi] \rangle_{\mathcal{E}} - \langle \beta[\psi], (\gamma\delta\bar{\beta} \cdot \nabla\psi)^+ \rangle_{\mathcal{E} \setminus \mathcal{E}_+^B} \\ &\geq -\sqrt{\max_{e \in \mathcal{E}} C_{11}} \left\| \sqrt{C_{11}}[\psi] \right\|_{\mathcal{E} \cap \Omega_0} \cdot \|\gamma\delta[\bar{\beta} \cdot \nabla\psi]\|_{\mathcal{E} \cap \Omega_0} \\ &\quad - \sqrt{\|\beta\|_{\infty}} \left\| \sqrt{|\beta \cdot \mathbf{n}|}[\psi] \right\|_{\mathcal{E} \cap \Omega_0} \cdot \|(\gamma\delta\bar{\beta} \cdot \nabla\psi)^+\|_{\mathcal{E} \cap \Omega_0} \\ &\geq -\frac{1}{4} \left\| (C_{11} + \beta \cdot \mathbf{n})^{\frac{1}{2}}[\psi] \right\|_{\mathcal{E} \cap \Omega_0}^2 - C\gamma^2 H \|\bar{\beta} \cdot \nabla\psi\|_{\Omega_0}^2 \\ &\geq -\frac{1}{4} \left\| (C_{11} + \beta \cdot \mathbf{n})^{\frac{1}{2}}[\psi] \right\|_{\mathcal{E} \cap \Omega_0}^2 - C\gamma^2 H \left(1 + CH^2 \|\beta\|_{1,\infty}^2\right) \|\beta \cdot \nabla\psi\|_{\Omega_0}^2. \end{aligned}$$

Substituting all these lower bounds into (4.26) yields

$$\begin{aligned} \mathcal{A}(\phi, \psi; 0, \gamma\delta\bar{\beta} \cdot \nabla\psi) &\geq -\frac{1}{4\epsilon} \|\phi\|_{\Omega_0}^2 + \gamma H(1 - C\gamma) \|\beta \cdot \nabla\psi\|_{\Omega_0}^2 \\ &\quad - \frac{1}{4} \|\sqrt{c}\psi\|_{\Omega_0}^2 - \frac{1}{4} \left\| (C_{11} + |\beta \cdot \mathbf{n}|)^{1/2}[\psi] \right\|_{\mathcal{E} \cap \Omega_0}^2 \\ &\geq -\frac{1}{4\epsilon} \|\phi\|_{\Omega_0}^2 + \gamma H(1 - C\gamma) \|\beta \cdot \nabla\psi\|_{\Omega_0}^2 - \frac{1}{4} c(\psi, \psi), \end{aligned}$$

Taking sufficiently small constant γ such that $0 < C_d \leq \gamma(1 - C\gamma) \leq C$, we obtain

$$(4.27) \quad \mathcal{A}(\phi, \psi; 0, \gamma\delta\bar{\beta} \cdot \nabla\psi) \geq C_d H \|\beta \cdot \nabla\psi\|_{\Omega_0}^2 - \frac{1}{4\epsilon} \|\phi\|_{\Omega_0}^2 - \frac{1}{4} c(\psi, \psi).$$

Combining (4.23), (4.25) and (4.27) yields

$$\begin{aligned} &\mathcal{A}(\phi, \psi; \mathbf{r}, v) \\ &\geq \frac{1}{2\epsilon} \|\phi\|^2 + C_d \left(H \|\beta \cdot \nabla\psi\|_{\Omega_0}^2 + h \|\nabla\psi\|_{\mathcal{T}_N \setminus \Omega_0}^2 \right) + \frac{1}{2} c(\psi, \psi) \\ (4.28) \quad &\geq \min\{C_d, \frac{1}{2}\} \|(\phi, \psi)\|_{\mathcal{A}}^2 \\ &\geq C \min\{C_d, \frac{1}{2}\} \|(\phi, \psi)\|_{\mathcal{A}} \|(\mathbf{r}, v)\|_{\mathcal{A}} \end{aligned}$$

for $(\mathbf{r}, v) = (\phi - \gamma h \nabla\psi, \psi + \gamma h \epsilon^{-1} \psi + \gamma\delta\bar{\beta} \cdot \nabla\psi)$. When proving (4.28), we used (4.1), (4.2) and the fact that $\|(\mathbf{r}, v)\|_{\mathcal{A}}$ is bounded by $\|(\phi, \psi)\|_{\mathcal{A}}$. Let $C_D = C \min\{C_d, \frac{1}{2}\}$. The result (4.22) now directly follows from (4.28). \square

4.5. **Error estimate of $|(\eta_{\mathbf{q}} <, \eta_u)|$.** It follows from the orthogonality property (2.6) of the compact form that

$$\mathcal{A}(\eta_{\mathbf{q}}, \eta_u; \mathbf{r}, v) = \mathcal{A}(\mathbf{q} - \mathbf{q}_N, u - u_N; \mathbf{r}, v) - \mathcal{A}(\xi_{\mathbf{q}}, \xi_u; \mathbf{r}, v) = -\mathcal{A}(\xi_{\mathbf{q}}, \xi_u; \mathbf{r}, v).$$

By the stability (4.22), one has

$$(4.29) \quad \begin{aligned} |(\eta_{\mathbf{q}}, \eta_u)|_{\mathcal{A}} &\leq \|(\eta_{\mathbf{q}}, \eta_u)\|_{\mathcal{A}} \leq C \sup_{(\mathbf{r}, v) \in \mathbf{V}_N \times M_N} \frac{\mathcal{A}(\eta_{\mathbf{q}}, \eta_u; \mathbf{r}, v)}{\|(\mathbf{r}, v)\|_{\mathcal{A}}} \\ &\leq C \sup_{(\mathbf{r}, v) \in \mathbf{V}_N \times M_N} \frac{|\mathcal{A}(\xi_{\mathbf{q}}, \xi_u; \mathbf{r}, v)|}{\|(\mathbf{r}, v)\|_{\mathcal{A}}}. \end{aligned}$$

Then the bound of $|(\eta_{\mathbf{q}}, \eta_u)|_{\mathcal{A}}$ follows directly from the above inequality (4.29) and the upper bound of $\mathcal{A}(\xi_{\mathbf{q}}, \xi_u; \mathbf{r}, v)$ in (4.31):

$$(4.30) \quad |(\eta_{\mathbf{q}}, \eta_u)|_{\mathcal{A}} \leq C \left(\frac{\ln N}{N} \right)^{k+\frac{1}{2}}.$$

Next, we shall complete the proof of the estimation of $|(\eta_{\mathbf{q}}, \eta_u)|_{\mathcal{A}}$ by proving of the boundedness of the compact form $\mathcal{A}(\xi_{\mathbf{q}}, \xi_u; \mathbf{r}, v)$.

Lemma 4.7 (Boundedness). *Suppose that (\mathbf{q}, u) is the solution of (2.1) satisfying the regularity (2.13) and (2.11), $(\mathbf{q}_N, u_N) \in \mathbf{V}_N \times M_N$ is the LDG solution of (2.3). If $0 \leq C_{11} \leq C$, there exists a positive constant C such that the estimate*

$$(4.31) \quad |\mathcal{A}(\xi_{\mathbf{q}}, \xi_u; \mathbf{r}, v)| \leq C \left(\frac{\ln N}{N} \right)^{k+\frac{1}{2}} \|(\mathbf{r}, v)\|_{\mathcal{A}}$$

holds for all $(\mathbf{r}, v) \in \mathbf{V}_N \times M_N$.

Proof. Recalling the definition of the compact form (2.5), we have

$$\begin{aligned} |\mathcal{A}(\xi_{\mathbf{q}}, \xi_u; \mathbf{r}, v)| &= |a(\xi_{\mathbf{q}}, \mathbf{r})| + |b_1(\xi_u, \mathbf{r})| + |b_2(v, \xi_{\mathbf{q}})| + |c(\xi_u, v)| \\ &= |a(\xi_{\mathbf{q}}, \mathbf{r})| + |b_1(\xi_w, \mathbf{r}) - b_2(\xi_{\bar{u}}, \mathbf{r})| + |b_2(v, \xi_{\mathbf{q}})| + |c(\xi_u, v)| \\ &= S_1 + S_2 + S_3 + S_4. \end{aligned}$$

(1) Consider S_1 . Using the estimate (4.11d) of interpolation error of $\xi_{\mathbf{q}}$, one has

$$(4.32) \quad S_1 \leq \epsilon^{-1} \|\xi_{\mathbf{q}}\| \|\mathbf{r}\| \leq C \left(\frac{\ln N}{N} \right)^{k+1} |(\mathbf{r}, v)|_{\mathcal{A}}.$$

(2) Consider S_2 . We use the technique in [8]. Let the superscript “out” denote the trace taken from outside of K . Then we have $w|_{\gamma_i^{\pm}}^{out} = w|_{\gamma_i^{\pm}}$ and $(\Pi w)|_{\gamma_i^-}^{out} = \pi^+ w|_{\gamma_i^-}$ for the corresponding one-dimensional projection π_i^{\pm} ($i = 1, 2$). Thus, we rewrite S_2 as

$$\begin{aligned} S_2 &\leq |(\xi_w, \nabla \cdot \mathbf{r}) - \langle \xi_w^-, \llbracket \mathbf{r} \rrbracket \rangle_{\mathcal{E}^0}| + |(\nabla \xi_{\bar{u}}, \mathbf{r}) + \langle \mathbf{r}^+, \llbracket \xi_{\bar{u}} \rrbracket \rangle_{\mathcal{E}^0} + \langle \mathbf{r} \cdot \mathbf{n}, \xi_{\bar{u}} \rangle_{\mathcal{E}^B}| \\ &\leq \left| \sum_{K \in \mathcal{T}_N} Z_K(\mathbf{r}, w) \right| + |(\nabla \xi_{\bar{u}}, \mathbf{r}) + \langle \mathbf{r} \cdot \mathbf{n}, \xi_{\bar{u}} \rangle_{\mathcal{E}^B}|, \end{aligned}$$

where

$$Z_K(\mathbf{r}, w) = \int_K \xi_w \nabla \cdot \mathbf{r} dx dy - \int_{\gamma_1^+ \cup \gamma_2^+} \xi_w (\mathbf{r} \cdot \mathbf{n}_K) ds - \int_{\gamma_1^- \cup \gamma_2^-} \xi_w^{out} (\mathbf{r} \cdot \mathbf{n}_K) ds.$$

Note that

$$(4.33) \quad Z_K(\mathbf{r}, w) = Z_{K,1}(r_1, w) + Z_{K,2}(r_2, w),$$

where $\mathbf{r} \in \mathbf{V}_N$ and

$$\begin{aligned} Z_{K,1}(r_1, w) &= \int_K \xi_w \frac{\partial r_1}{\partial x} dx dy - \int_{\gamma_2^+} (u - \pi_2^+ w) r_1 dy - \int_{\gamma_2^-} (u - \pi_2^+ w) r_1 dy, \\ Z_{K,2}(r_2, w) &= \int_K \xi_w \frac{\partial r_2}{\partial y} dx dy - \int_{\gamma_1^+} (u - \pi_1^+ w) r_2 dx - \int_{\gamma_1^-} (u - \pi_1^+ w) r_2 dx. \end{aligned}$$

Using (4.41), we obtain

$$(4.34) \quad \left| \sum_{K \in \mathcal{T}_N} Z_K(\mathbf{r}, w) \right| \leq C \left(\frac{\ln N}{N} \right)^{k+1} |(\mathbf{r}, v)|_{\mathcal{A}}.$$

On the other hand, it follows the Shwarz inequality and (4.11a) that

$$(4.35) \quad |(\nabla \xi_{\bar{u}}, \mathbf{r})| \leq \epsilon^{\frac{1}{2}} \|\nabla \xi_{\bar{u}}\| \cdot \epsilon^{-\frac{1}{2}} \|\mathbf{r}\| \leq C \epsilon^{\frac{1}{2}} N^{-k} |(\mathbf{r}, v)|_{\mathcal{A}} \leq CN^{-(k+\frac{1}{2})} |(\mathbf{r}, v)|_{\mathcal{A}}.$$

We estimate the boundary term by standard approximation theory of 1-D Lagrange interpolation and inverse inequality for \mathbf{r} :

$$(4.36) \quad \begin{aligned} |\langle \mathbf{r} \cdot \mathbf{n}, \xi_{\bar{u}} \rangle_{\mathcal{E}^B}| &\leq |\langle \mathbf{r} \cdot \mathbf{n}, \xi_{\bar{u}} \rangle_{\mathcal{E}^B}| \\ &\leq \|\xi_{\bar{u}}\|_{\mathcal{E}^B} \cdot \|\mathbf{r} \cdot \mathbf{n}\|_{\mathcal{E}^B \cap \bar{\Omega}_{xy}} \\ &\leq CH^{k+1} \cdot C(h^{-1/2} + H^{-1/2}) \|\mathbf{r}\| \\ &\leq CN^{-(k+\frac{1}{2})} |(\mathbf{r}, v)|_{\mathcal{A}}. \end{aligned}$$

The combination of (4.34), (4.35) and (4.36) yields

$$(4.37) \quad S_2 \leq CN^{-(k+\frac{1}{2})} |(\mathbf{r}, v)|_{\mathcal{A}}.$$

(3) Consider S_3 . It follows the definition of S_3 that

$$\begin{aligned} S_3 &\leq |(\nabla v, \xi_{\mathbf{q}})| + |\langle \xi_{\mathbf{q}^+}, \llbracket v \rrbracket \rangle_{\mathcal{E}^0} + \langle \xi_{\mathbf{q}}, v\mathbf{n} \rangle_{\mathcal{E}^B}| \\ &\leq \sum_{e \in \mathcal{E}_+^B} |\langle \xi_{\mathbf{q}}, v\mathbf{n} \rangle_e|, \end{aligned}$$

where we used interpolation properties (4.6) and error estimates on subsets of \mathcal{E} :

$$\begin{aligned} (\nabla v, \xi_{\mathbf{q}})_K &= 0, \quad \forall K \in \mathcal{T}_N, \\ \langle \xi_{\mathbf{q}^+}, \llbracket v \rrbracket \rangle_{\mathcal{E}^0 \cup \mathcal{E}^B} &= \sum_{K \in \mathcal{T}_N} \left[\langle \xi_{\mathbf{q}}, v\mathbf{n} \rangle_{\gamma_1^-} + \langle \xi_{\mathbf{q}}, v\mathbf{n} \rangle_{\gamma_2^-} \right] = 0. \end{aligned}$$

Therefore, by (4.17b) we get that

$$(4.38) \quad S_3 \leq \|\xi_{\mathbf{q}}\|_{\mathcal{E}_+^B} \|v\mathbf{n}\|_{\mathcal{E}_+^B} \leq C \left(\frac{\ln N}{N} \right)^{k+1} |(\mathbf{r}, v)|_{\mathcal{A}}.$$

(4) Consider S_4 . Applying the Schwarz inequality on each $e \in \mathcal{E}$ yields

$$\begin{aligned}
 S_4 &= \left| (\xi_u, \boldsymbol{\beta} \cdot \nabla v) + \langle \xi_{u-}, \boldsymbol{\beta} \cdot \llbracket v \rrbracket \rangle_{\mathcal{E}^0 \cup \mathcal{E}_+^B} + (\xi_u, cv) + \langle C_{11} \llbracket \xi_u \rrbracket, \llbracket v \rrbracket \rangle_{\mathcal{E}} \right| \\
 &\leq \sqrt{\frac{1}{H}} \|\xi_u\|_{\Omega_0} \sqrt{H} \|\boldsymbol{\beta} \cdot \nabla v\|_{\Omega_0} + C \frac{1}{\sqrt{h}} \|\xi_u\|_{\mathcal{T}_N \setminus \Omega_0} \sqrt{h} \|\nabla v\|_{\mathcal{T}_N \setminus \Omega_0} \\
 (4.39) \quad &+ C \|\xi_{u-}\|_{\mathcal{E}^0 \cup \mathcal{E}_+^B} \left\| |\boldsymbol{\beta} \cdot \mathbf{n}|^{1/2} \llbracket v \rrbracket \right\|_{\mathcal{E}} + C \|\xi_u\| \left\| c^{1/2} v \right\| \\
 &+ C \left\| C_{11}^{1/2} \llbracket \xi_u \rrbracket \right\|_{\mathcal{E}} \left\| C_{11}^{1/2} \llbracket v \rrbracket \right\|_{\mathcal{E}} \\
 &\leq C \left(\frac{\ln N}{N} \right)^{k+\frac{1}{2}} \|(\mathbf{r}, v)\|_{\mathcal{A}},
 \end{aligned}$$

where we used the interpolation error estimate (4.11c), (4.14), and (4.17a). $\|c^{1/2}v\|$ can be bounded by $\left\| \left(c - \frac{1}{2} \operatorname{div} \boldsymbol{\beta} \right)^{1/2} v \right\|$ because c is uniformly bounded on $\overline{\Omega}$ and satisfies (1.2).

The inequality (4.31) now follows the combination of (4.32), (4.37), (4.38), and (4.39). □

In the estimation of S_2 (see the proof of Lemma 4.7, Part 2), we need a super-convergence result, essentially due to LeSaint and Raviart [12], which is similar to Lemma 3.6 in [8].

Lemma 4.8. *Let the bilinear form $Z_K(\mathbf{r}, \varphi)$ be defined by (4.33). If $\varphi \in H^{k+2}(\Omega)$ and $(\mathbf{r}, v) \in \mathbf{V}_N \times M_N$, then we have for $s \geq 0$,*

$$(4.40a) \quad |Z_{K,1}(r_1, \varphi)| \leq Ch_x^{-1} \sum_{|\alpha|=1} \tilde{h}^{(k+2)\alpha} \left\| D^{(k+2)\alpha} \varphi \right\|_K \|r_1\|_K,$$

$$(4.40b) \quad |Z_{K,2}(r_2, \varphi)| \leq Ch_y^{-1} \sum_{|\alpha|=1} \tilde{h}^{(k+2)\alpha} \left\| D^{(k+2)\alpha} \varphi \right\|_K \|r_2\|_K.$$

The constant C depends only on s and k . Here we use the notation $\mathbf{r} = (r_1, r_2)$, $\tilde{h} = (h_x, h_y)$, and multi-index notation $\alpha = (\alpha_1, \alpha_2) \in \{(1, 0), (0, 1)\}$.

Proof. We only provide the proof of (4.40a). The other inequality follows the same argument.

For fixed $\hat{r}_1 \in \mathcal{Q}_k(\hat{K})$, the linear functional $\hat{\varphi} \mapsto Z_{\hat{K},1}(\hat{r}_1, \hat{\varphi})$ is continuous on $H^{k+2}(\hat{K})$ with norm bounded by $C \|\hat{r}_1\|_{\hat{K}}$. Since on the reference element $\hat{K} = (-1, 1)^2$, it was proved in Lemma 3.6, [8] that

$$Z_{\hat{K},1}(\hat{r}_1, \hat{\varphi}) = 0, \quad \forall u \in \mathcal{P}_{k+1}(\hat{K}), \quad r_1 \in \mathcal{Q}_k(\hat{K}).$$

By applying the Bramble-Hilbert lemma (see Lemma 6 [6], for instance) and Lemma 2.12 [1] we obtain

$$Z_{\hat{K},1}(\hat{r}_1, \hat{\varphi}) \leq C \|\hat{r}_1\|_{\hat{K}} [\varphi]_{H^{k+2}(\hat{K})},$$

where $[\varphi]_{H^{k+2}(\hat{K})}$ is defined in (4.9). The inequality (4.40a) follows directly from the above inequality and the scaling argument. □

The following lemma is a corollary of Lemma 4.8.

Lemma 4.9. *Let $Z_K(\mathbf{r}, w)$ be defined by (4.33). There exists a constant C which depends only on s, k such that*

$$(4.41) \quad \sum_{K \in \mathcal{T}_N} |Z_K(\mathbf{r}, w)| \leq C \left(\frac{\ln N}{N} \right)^{k+1} \cdot \epsilon^{-\frac{1}{2}} \|\mathbf{r}\|.$$

Proof. We use the same technique as in the proof of Lemma 4.4 and Lemma 4.5.

First, we use the inequalities (4.40a)-(4.40b) in Lemma 4.8 for the estimation of $|Z_K(\mathbf{r}, w)|$ on subdomain Ω_{xy} . From the regularity results (2.11), we have

$$(4.42) \quad \begin{aligned} \sum_{K \in \Omega_{xy}} |Z_K(\mathbf{r}, w)| &\leq Ch^{-1}h^{k+2}[w]_{H^{k+2}(\Omega_{xy})} \|\mathbf{r}\|_{\Omega_{xy}} \\ &\leq Ch^{k+1}\epsilon^{-(k+2)}\epsilon \|\mathbf{r}\|_{\Omega_{xy}} \\ &\leq C\sqrt{\epsilon}(\ln N/N)^{k+1} \cdot \epsilon^{-\frac{1}{2}} \|\mathbf{r}\|_{\Omega_{xy}}. \end{aligned}$$

Second, consider the error estimate on subdomain Ω_0 . Instead of using the error estimates of ξ_u in Lemma 4.8, we analyze u and Π^+w on Ω_0 , and π^+w on the interior boundary in $\mathcal{E} \cap \overline{\Omega_0}$, individually. Using the same technique as in Lemma 4.5, we derive the error bound of $\sum_{K \in \Omega_0} |Z_K(\mathbf{r}, w)|$ as follows:

$$(4.43) \quad \sum_{K \in \Omega_0} |Z_K(\mathbf{r}, w)| \leq CN^{-(k+1)} \cdot \epsilon^{-\frac{1}{2}} \|\mathbf{r}\|_{\Omega_0}.$$

Third, consider Ω_x . We estimate $\sum_{K \in \Omega_x} |Z_{K,1}(r_1, \varphi)|$ for two groups of functions $\{w_1\}$ and $\{w_2, w_0\}$ separately. For any $\varphi \in \{w_1\}$, applying the inequality (4.40a) yields

$$(4.44) \quad \sum_{K \in \Omega_x} |Z_{K,1}(r_1, \varphi)| \leq C(\ln N/N)^{k+1} \cdot \epsilon^{-\frac{1}{2}} \|r_1\|_{\Omega_x}.$$

For any $\varphi \in \{w_2, w_0\}$, we analyze u and Π^+u on Ω_x , and π^+u on the interior boundary in $\mathcal{E} \cap \Omega_x$, separately. The same error bound can be obtained by the same argument as proving (4.44).

Next, we move onto the estimation of $\sum_{K \in \Omega_x} |Z_{K,2}(r_2, \varphi)|$. If $\varphi \in \{w_1\}$, they have the same error bounds as in (4.44) by (4.40b). Otherwise, if $\varphi \in \{w_2, w_0\}$, the combination of the interpolation error estimate (4.10c), (4.10f), and the fact that w_2 and w_0 could be arbitrarily small on Ω_x , gives rise to the same error bounds as in (4.44).

Error estimates on Ω_y will be the same as that of Ω_x because of the symmetry of their expressions. Summarizing (4.42)–(4.44), we complete the proof of Lemma 4.9. □

5. NUMERICAL EXPERIMENTS

We use $N = 2^n$ for $n = 2, 3, 4, 5, 6$, and $C_{11} = 0$ on \mathcal{E} . Let $E_k^j(N)$ stand for different errors, where the superscript j indicates two values of the parameter $\epsilon = 10^{-4}, 10^{-6}$ and k denotes the degree of polynomials. The rates of convergence are computed by the formula

$$R_2 := \log_2 \frac{E_k^j(N)}{E_k^j(2N)}, \quad \text{or} \quad R_p := \log_p \frac{E_k^j(N)}{E_k^j(2N)}, \quad j = 1, 2, 3.$$

Here $p := 2(\ln N/\ln 2N) = 2n/(n + 1)$. We will use the formula R_p in computing rates of convergence for all errors except $\|\xi_u\|_{L^2}$.

For each example, we present the numerical results in the ϵ -weighted semi-norm $|(\mathbf{r}, v)|_{\mathcal{A}}$, L^2 -norm, and in L^∞ -norm. We set $\kappa = k + 1$ in the numerical experiments. Let $(\beta_1, \beta_2) = 0.99 \cdot \boldsymbol{\beta}$ so that the assumption $\boldsymbol{\beta} > (\beta_1, \beta_2) > (0, 0)$ is satisfied.

For each example, we test the following three different choices of C_{11} :

- (1) $C_{11} = 0$ on \mathcal{E} ,
- (2) $C_{11} = 1$ on \mathcal{E} ,
- (3) $C_{11} = 0$ on $\mathcal{E} \setminus \mathcal{E}_+^B$; and $C_{11} = H^{-1}$ on \mathcal{E}_+^B .

Numerical experiments show that the first type of parameter C_{11} always yields the smallest errors. So, we only list the numerical results when $C_{11} = 0$ on \mathcal{E} .

Example 1: First, we consider Problem (1.1) with $\boldsymbol{\beta} = (1, 1)$, $c = 0$. The right side term f is properly chosen such that the real solution

$$u = xy \left(1 - e^{-2(1-x)/\epsilon}\right) \left(1 - e^{-2(1-y)/\epsilon}\right).$$

In Table 1 and Table 2, we report the error $|(e_q, e_u)|_{\mathcal{A}}$ and its convergence rates that are of $O((\ln N/N)^{k+1/2})$, in full agreement with Theorem 3.1. The rates of convergence in the L^2 -norm and in the L^∞ -norm of the error in \mathbf{q} appear to be $O((\ln N/N)^{k+1})$. Noticing that for this example, $|(e_q, e_u)|_{\mathcal{A}}$ comprises $\epsilon^{-\frac{1}{2}} \|e_q\|$ and boundary term $\|[(e_q)]\|$, we realized the error is being dominated by boundary terms.

The results displayed in Table 3 are the L^2 error in u and its convergence rates. Instead of formula R_p , we use formula R_2 to compute the convergence rates, which appear to be $O(N^{-(k+\frac{1}{2})})$. The reason may be that the L^2 error in u is dominated by its error on the regular part of the domain Ω_0 , where the rate of convergence is of $O(N^{-r})$ for some positive constant r . Furthermore, the rates of convergence of $\|\xi_u\|_{L^\infty}$ seems close to $O((\ln N/N)^{k+1})$.

Example 2: We consider Problem (1.1) with $\boldsymbol{\beta} = (2, 2)$, $c = 1$. The right side term f is properly chosen such that the real solution

$$u = \sin(xy) \left(1 - e^{-2(1-x)/\epsilon}\right) \left(1 - e^{-2(1-y)/\epsilon}\right).$$

The convergence behaviors of the LDG solution of Example 2 are similar to Example 1. Numerical results presented in Table 4 and Table 5 show that the convergence rate of $|(e_q, e_u)|_{\mathcal{A}}$ is of $O((\ln N/N)^{k+1/2})$. The errors of e_q converges at a rate of $O((\ln N/N)^{k+1})$ when $k = 1, 2$, while a slight degradation appears when $k = 3$ and $\epsilon = 10^{-4}$.

We also see from Table 6 that e_u in L^2 -norm converges at a rate of $O(N^{-(k+\frac{1}{2})})$. The rate of convergence of e_u in L^∞ -norm is not clear.

Remark 5.1. We test the effect of the choice of the parameter C_{11} on the order of the convergence. Comparing with the LDG method for elliptic problem in [8], we can conclude that the convergence of LDG for Problem (1.1) is not very sensitive to the choice of the parameter C_{11} .

We observed the same convergence rate for $|(\eta_{\mathbf{q}}, \eta_u)|_{\mathcal{A}}$ as proved in Theorem 3.1 when these three types of parameter C_{11} are used. However, the first type of parameter C_{11} always yields smaller errors. Furthermore, we observed generally no notable change in the numerical results when the first type and the third type of parameter C_{11} . We also observed that taking $C_{11} \sim O(1)$ on \mathcal{E} results in a slight degradation of orders of convergence in the L^2 norm and L^∞ norm of \mathbf{q} and u , comparing with the other two choices of C_{11} .

TABLE 1. History of convergence when $\epsilon = 10^{-4}$ (Example 1).

n	k	$ (e_q, e_u) _{\mathcal{A}}$		$\epsilon^{-\frac{1}{2}} \ e_q\ _{L^2}$		$\ e_q\ _{L^\infty}$	
		error	order	error	order	error	order
2	1	7.13e-2	–	2.28e-2	–	8.56e-2	–
3		3.59e-2	2.38	9.07e-3	3.21	4.05e-2	2.60
4		1.89e-2	1.58	3.94e-3	2.06	1.80e-2	2.01
5		9.38e-3	1.49	1.56e-3	1.97	7.24e-3	1.93
6		4.39e-3	1.49	5.70e-4	1.97	2.69e-3	1.94
2		2	1.60e-2	–	4.47e-3	–	2.29e-2
3	4.96e-3		4.08	1.21e-3	4.54	7.55e-3	3.85
4	1.76e-3		2.55	3.61e-4	2.98	2.36e-3	2.87
5	5.52e-4		2.47	9.04e-5	2.95	5.85e-4	2.97
6	1.55e-4		2.48	1.98e-5	2.97	1.26e-4	3.01
2	3		3.77e-3	–	9.47e-4	–	4.08e-3
3		7.17e-4	5.77	1.74e-4	5.89	7.76e-4	5.77
4		1.74e-4	3.49	3.56e-5	3.92	1.80e-4	3.60
5		3.45e-5	3.45	5.62e-6	3.93	3.17e-5	3.70
6		5.86e-6	3.47	7.43e-7	3.96	4.51e-6	3.82

TABLE 2. History of convergence when $\epsilon = 10^{-6}$ (Example 1).

n	k	$ (e_q, e_u) _{\mathcal{A}}$		$\epsilon^{-\frac{1}{2}} \ e_q\ _{L^2}$		$\ e_q\ _{L^\infty}$	
		error	order	error	order	error	order
2	1	7.13e-2	–	2.28e-2	–	8.56e-2	–
3		3.59e-2	2.38	9.04e-3	3.21	4.05e-2	2.60
4		1.89e-2	1.58	3.93e-3	2.05	1.80e-2	2.01
5		9.38e-3	1.49	1.56e-3	1.97	7.24e-3	1.93
6		4.39e-3	1.49	5.69e-4	1.97	2.69e-3	1.94
2		2	1.61e-2	–	4.41e-3	–	2.29e-2
3	4.97e-3		4.08	1.20e-3	4.51	7.55e-3	3.86
4	1.76e-3		2.55	3.60e-4	2.97	2.36e-3	2.86
5	5.52e-4		2.47	9.03e-5	2.94	5.85e-4	2.97
6	1.55e-4		2.48	1.98e-5	2.97	1.26e-4	3.01
2	3		3.78e-3	–	9.13e-4	–	4.11e-3
3		7.17e-4	5.77	1.72e-4	5.78	7.77e-4	5.79
4		1.74e-4	3.49	3.55e-5	3.90	1.80e-4	3.61
5		3.45e-5	3.44	5.61e-6	3.92	3.16e-5	3.70
6		5.86e-6	3.47	7.42e-7	3.96	4.55e-6	3.79

TABLE 3. History of convergence of e_u (Example 1).

n	k	$\epsilon = 10^{-4}$				$\epsilon = 10^{-6}$			
		$\ e_u\ _{L^2}$		$\ e_u\ _{L^\infty}$		$\ e_u\ _{L^2}$		$\ e_u\ _{L^\infty}$	
		error	order	error	order	error	order	error	order
2	1	1.41e-2	–	8.85e-2	–	1.41e-2	–	8.86e-2	–
3		2.61e-3	2.43	3.50e-2	3.22	2.62e-3	2.43	3.50e-2	3.22
4		4.78e-4	2.45	1.61e-2	1.92	4.78e-4	2.45	1.61e-2	1.92
5		8.76e-5	2.45	6.68e-3	1.87	8.65e-5	2.46	6.68e-3	1.87
6		1.64e-5	2.41	2.54e-3	1.90	1.56e-5	2.47	2.54e-3	1.90
2	2	2.86e-3	–	2.25e-2	–	2.87e-3	–	2.26e-2	–
3		2.63e-4	3.45	5.38e-3	4.96	2.64e-4	3.44	5.39e-3	4.97
4		2.41e-5	3.45	1.68e-3	2.87	2.40e-5	3.46	1.68e-3	2.87
5		2.34e-6	3.37	4.59e-4	2.77	2.18e-6	3.46	4.59e-4	2.77
6		2.77e-7	3.07	1.07e-4	2.85	1.98e-7	3.46	1.07e-4	2.85
2	3	6.18e-4	–	5.88e-3	–	6.21e-4	–	5.91e-3	–
3		2.84e-5	4.44	7.58e-4	7.12	2.86e-5	4.44	7.58e-4	7.13
4		1.33e-6	4.41	1.80e-4	3.55	1.31e-6	4.45	1.80e-4	3.55
5		8.10e-8	4.04	3.16e-5	3.70	5.98e-8	4.45	3.16e-5	3.70
6		7.92e-9	3.35	4.51e-6	3.81	2.83e-9	4.40	4.51e-6	3.81

TABLE 4. History of convergence when $\epsilon = 10^{-4}$ (Example 2).

n	k	$ (e_q, e_u) _A$		$\epsilon^{-\frac{1}{2}} \ e_q\ _{L^2}$		$\ e_q\ _{L^\infty}$	
		error	order	error	order	error	order
2	1	6.84e-2	–	2.86e-2	–	1.48e-1	–
3		3.44e-2	2.40	1.19e-2	3.06	7.02e-2	2.59
4		1.80e-2	1.60	5.22e-3	2.03	3.12e-2	2.00
5		8.86e-3	1.51	2.07e-3	1.96	1.26e-2	1.93
6		4.12e-3	1.50	7.58e-4	1.97	4.67e-3	1.94
2	2	1.44e-2	–	5.29e-3	–	3.83e-2	–
3		4.77e-3	3.84	1.61e-3	4.14	1.10e-2	4.35
4		1.71e-3	2.54	4.88e-4	2.94	3.62e-3	2.73
5		5.31e-4	2.48	1.23e-4	2.94	8.48e-4	3.08
6		1.49e-4	2.49	2.72e-5	2.94	3.62e-4	1.67
2	3	3.25e-3	–	1.08e-3	–	6.17e-3	–
3		7.02e-4	5.33	2.36e-4	5.28	1.41e-3	5.14
4		1.72e-4	3.46	4.91e-5	3.87	3.25e-4	3.61
5		3.43e-5	3.43	9.26e-6	3.55	3.55e-4	-.17
6		9.12e-6	2.59	7.17e-6	0.50	4.58e-4	-.50

TABLE 5. History of convergence when $\epsilon = 10^{-6}$ (Example 2).

n	k	$ (e_q, e_u) _{\mathcal{A}}$		$\epsilon^{-\frac{1}{2}} \ e_q\ _{L^2}$		$\ e_q\ _{L^\infty}$	
		error	order	error	order	error	order
2	1	6.85e-2	-	2.86e-2	-	1.48e-1	-
3		3.44e-2	2.40	1.19e-2	3.06	7.02e-2	2.59
4		1.80e-2	1.60	5.21e-3	2.03	3.12e-2	2.00
5		8.86e-3	1.51	2.07e-3	1.96	1.26e-2	1.93
6		4.12e-3	1.50	7.57e-4	1.97	4.67e-3	1.94
2		2	1.44e-2	-	5.26e-3	-	3.83e-2
3	4.77e-3		3.84	1.61e-3	4.12	1.10e-2	4.35
4	1.71e-3		2.54	4.88e-4	2.94	3.71e-3	2.67
5	5.31e-4		2.48	1.22e-4	2.94	9.80e-4	2.83
6	1.49e-4		2.49	2.69e-5	2.97	2.17e-4	2.95
2	3		3.25e-3	-	1.06e-3	-	6.19e-3
3		7.02e-4	5.33	2.36e-4	5.23	1.41e-3	5.15
4		1.72e-4	3.46	4.91e-5	3.87	3.26e-4	3.61
5		3.39e-5	3.46	7.78e-6	3.92	5.82e-5	3.66
6		6.02e-6	3.38	2.13e-6	2.54	1.30e-5	2.93

TABLE 6. History of convergence of e_u (Example 2).

n	k	$\epsilon = 10^{-4}$				$\epsilon = 10^{-6}$			
		$\ e_u\ _{L^2}$		$\ e_u\ _{L^\infty}$		$\ e_u\ _{L^2}$		$\ e_u\ _{L^\infty}$	
		error	order	error	order	error	order	error	order
2	1	1.25e-2	-	8.09e-2	-	1.25e-2	-	8.09e-2	-
3		2.28e-3	2.45	2.73e-2	3.77	2.28e-3	2.45	2.73e-2	3.77
4		4.10e-4	2.47	1.37e-2	1.71	4.09e-4	2.48	1.37e-2	1.71
5		7.38e-5	2.47	5.81e-3	1.82	7.32e-5	2.48	5.81e-3	1.82
6		1.37e-5	2.43	2.22e-3	1.89	1.32e-5	2.47	2.22e-3	1.89
2		2	2.38e-3	-	1.88e-2	-	2.38e-3	-	1.88e-2
3	2.10e-4		3.50	4.79e-3	5.08	2.10e-4	3.50	4.35e-3	5.09
4	1.85e-5		3.50	1.49e-3	2.63	1.84e-5	3.51	1.49e-3	2.63
5	1.71e-6		3.43	4.08e-4	2.76	1.60e-6	3.52	4.08e-4	2.76
6	1.98e-7		3.11	1.41e-4	2.08	1.36e-7	3.56	9.53e-5	2.85
2	3		4.98e-4	-	4.75e-3	-	4.99e-4	-	4.76e-3
3		2.17e-5	4.52	6.84e-4	6.73	2.18e-5	4.52	6.84e-4	6.75
4		9.69e-7	4.49	1.63e-4	3.54	9.43e-7	4.53	1.63e-4	3.54
5		8.03e-8	3.59	1.65e-4	-0.2	4.16e-8	4.50	2.87e-5	3.69
6		9.29e-8	-2.1	2.11e-4	-4.8	6.13e-9	2.76	5.28e-6	3.31

6. CONCLUDING REMARKS

In this paper, we have shown a uniform convergence of the LDG method on Shishkin mesh for a singularly perturbed convection–diffusion problem. The proof is based on suitably defined projection introduced in [8], which exhibit a tensor product structure. We showed that the error of the LDG approximations to the auxiliary variable q and the primary variable u in a DG-norm is uniformly bounded by $C(\ln N/N)^{k+\frac{1}{2}}$, when polynomial spaces \mathcal{Q}_k are used. Numerical experiments were presented to verify that this rate of convergence is sharp.

ACKNOWLEDGMENTS

The authors would like to thank the anonymous referees for their helpful remarks.

REFERENCES

- [1] Thomas Apel, *Anisotropic finite elements: local estimates and applications*, Advances in Numerical Mathematics, B. G. Teubner, Stuttgart, 1999. MR1716824 (2000k:65002)
- [2] S.C. Brenner and L.R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer-Verlag, 1998.
- [3] Paul Castillo, *An optimal estimate for the local discontinuous Galerkin method*, Discontinuous Galerkin methods (Newport, RI, 1999), Lect. Notes Comput. Sci. Eng., vol. 11, Springer, Berlin, 2000, pp. 285–290, DOI 10.1007/978-3-642-59721-3_23. MR1842183 (2002d:65096)
- [4] Paul Castillo, Bernardo Cockburn, Ilaria Perugia, and Dominik Schötzau, *An a priori error analysis of the local discontinuous Galerkin method for elliptic problems*, SIAM J. Numer. Anal. **38** (2000), no. 5, 1676–1706 (electronic), DOI 10.1137/S0036142900371003. MR1813251 (2002k:65175)
- [5] Paul Castillo, Bernardo Cockburn, Dominik Schötzau, and Christoph Schwab, *Optimal a priori error estimates for the hp-version of the local discontinuous Galerkin method for convection-diffusion problems*, Math. Comp. **71** (2002), no. 238, 455–478, DOI 10.1090/S0025-5718-01-01317-5. MR1885610 (2003e:65214)
- [6] P. G. Ciarlet and P.-A. Raviart, *General Lagrange and Hermite interpolation in \mathbf{R}^n with applications to finite element methods*, Arch. Rational Mech. Anal. **46** (1972), 177–199. MR0336957 (49 #1730)
- [7] Bernardo Cockburn and Bo Dong, *An analysis of the minimal dissipation local discontinuous Galerkin method for convection-diffusion problems*, J. Sci. Comput. **32** (2007), no. 2, 233–262, DOI 10.1007/s10915-007-9130-3. MR2320571 (2008e:65345)
- [8] Bernardo Cockburn, Guido Kanschat, Ilaria Perugia, and Dominik Schötzau, *Superconvergence of the local discontinuous Galerkin method for elliptic problems on Cartesian grids*, SIAM J. Numer. Anal. **39** (2001), no. 1, 264–285 (electronic), DOI 10.1137/S0036142900371544. MR1860725 (2002g:65140)
- [9] Paul Houston, Christoph Schwab, and Endre Süli, *Discontinuous hp-finite element methods for advection-diffusion-reaction problems*, SIAM J. Numer. Anal. **39** (2002), no. 6, 2133–2163, DOI 10.1137/S0036142900374111. MR1897953 (2003d:65108)
- [10] C. Johnson, A. H. Schatz, and L. B. Wahlbin, *Crosswind smear and pointwise errors in streamline diffusion finite element methods*, Math. Comp. **49** (1987), no. 179, 25–38, DOI 10.2307/2008248. MR890252 (88i:65130)
- [11] Natalia Kopteva and Eugene O’Riordan, *Shishkin meshes in the numerical solution of singularly perturbed differential equations*, Int. J. Numer. Anal. Model. **7** (2010), no. 3, 393–415. MR2644280 (2011f:65130)
- [12] P. Lesiant and P.A. Raviart, *On a finite element method for solving the neutron transport equation*, In: Mathematical Aspects for Finite Elements in Partial Differential Equations (C. de Boor, Ed.), Academic Press, New York, 1974.
- [13] Torsten Linß and Martin Stynes, *Asymptotic analysis and Shishkin-type decomposition for an elliptic convection-diffusion problem*, J. Math. Anal. Appl. **261** (2001), no. 2, 604–632, DOI 10.1006/jmaa.2001.7550. MR1853059 (2002h:35039)

- [14] J. M. Melenk and C. Schwab, *Analytic regularity for a singularly perturbed problem*, SIAM J. Math. Anal. **30** (1999), no. 2, 379–400, DOI 10.1137/S0036141097317542. MR1664765 (2000b:35010)
- [15] J.J. Miller, E. O’Riordan and G.I. Shishkin, *Fitted numerical methods for singularly perturbed reaction-diffusion problems in two and one dimensions*, World Scientific, Singapore, 1996.
- [16] Hans-Görg Roos, Martin Stynes, and Lutz Tobiska, *Robust numerical methods for singularly perturbed differential equations*, 2nd ed., Springer Series in Computational Mathematics, vol. 24, Springer-Verlag, Berlin, 2008. Convection-diffusion-reaction and flow problems. MR2454024 (2009f:65002)
- [17] Dominik Schötzau and Christoph Schwab, *Time discretization of parabolic problems by the hp-version of the discontinuous Galerkin finite element method*, SIAM J. Numer. Anal. **38** (2000), no. 3, 837–875, DOI 10.1137/S0036142999352394. MR1781206 (2001i:65107)
- [18] Martin Stynes and Eugene O’Riordan, *A uniformly convergent Galerkin method on a Shishkin mesh for a convection-diffusion problem*, J. Math. Anal. Appl. **214** (1997), no. 1, 36–54, DOI 10.1006/jmaa.1997.5581. MR1645503 (99f:65177)
- [19] Martin Stynes and Lutz Tobiska, *Using rectangular Q_p elements in the SDFEM for a convection-diffusion problem with a boundary layer*, Appl. Numer. Math. **58** (2008), no. 12, 1789–1802, DOI 10.1016/j.apnum.2007.11.004. MR2464811 (2009j:65339)
- [20] Ziqing Xie, Zuozheng Zhang, and Zhimin Zhang, *A numerical study of uniform superconvergence of LDG method for solving singularly perturbed problems*, J. Comput. Math. **27** (2009), no. 2-3, 280–298. MR2495061 (2010c:65124)
- [21] Ziqing Xie and Zhimin Zhang, *Uniform superconvergence analysis of the discontinuous Galerkin method for a singularly perturbed problem in 1-D*, Math. Comp. **79** (2010), no. 269, 35–45, DOI 10.1090/S0025-5718-09-02297-2. MR2552216 (2010j:65106)
- [22] Helena Zarin and Hans-Görg Roos, *Interior penalty discontinuous approximations of convection-diffusion problems with parabolic layers*, Numer. Math. **100** (2005), no. 4, 735–759, DOI 10.1007/s00211-005-0598-1. MR2194592 (2006i:76070)
- [23] Zhimin Zhang, *Finite element superconvergence on Shishkin mesh for 2-D convection-diffusion problems*, Math. Comp. **72** (2003), no. 243, 1147–1177 (electronic), DOI 10.1090/S0025-5718-03-01486-8. MR1972731 (2004d:65130)
- [24] H. Zhu, *Discontinuous Galerkin methods for singularly perturbed problems*, Ph.D. thesis, Wayne State University, Detroit, 2009.
- [25] H. Zhu and Z. Zhang, *Convergence analysis of The LDG Method applied to Singularly Perturbed Problems*, Numer. Methods Partial Differential Equations **29** (2013), 396–421. MR3022892
- [26] Huiqing Zhu, Haiyan Tian, and Zhimin Zhang, *Convergence analysis of the LDG method for singularly perturbed two-point boundary value problems*, Commun. Math. Sci. **9** (2011), no. 4, 1013–1032. MR2901814

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF SOUTHERN MISSISSIPPI, HATTIESBURG, MISSISSIPPI 39406

E-mail address: Huiqing.Zhu@usm.edu

DEPARTMENT OF MATHEMATICS, WAYNE STATE UNIVERSITY, DETROIT, MICHIGAN 48202

BEIJING COMPUTATIONAL SCIENCE RESEARCH CENTER, NO. 3 HEQING ROAD, HAIDIAN DISTRICT, BEIJING 100084, CHINA.

E-mail address: zzhang@math.wayne.edu