UNIFIED ANALYSIS OF THE SEMIDISCRETE DISCONTINUOUS GALERKIN METHODS FOR 2-D HYPERBOLIC EQUATIONS ON CARTESIAN MESHES USING $\mathbb{P}^K$ ELEMENTS: OPTIMAL ERROR ESTIMATES AND SUPERCONVERGENCE*

WAIXIANG CAO *, CHI-WANG SHU†, AND ZHIMIN ZHANG ‡

Abstract. This paper is concerned with optimal error estimates and superconvergence properties of the discontinuous Galerkin (DG) method for two-dimensional hyperbolic conservation laws on uniform Cartesian meshes when piecewise polynomials of degree not more than k (i.e., $\mathbb{P}^k$ elements) are used. Superconvergence result of the DG approximation errors for the cell average and for the downwind edge average with an order of $\mathcal{O}(h^{k+2})$ is derived, under the condition that the wind direction does not change sign on the whole domain. As a byproduct, an optimal convergence rate of order $\mathcal{O}(h^{k+1})$ in the standard L^2 norm is obtained for both linear and nonlinear hyperbolic equations. Numerical experiments are presented to verify our theoretical findings.

Key words. Discontinuous Galerkin method, superconvergence, $\mathbb{P}^k$ elements, cell average, downwind edge average.

AMS subject classifications. 65M15, 65M60, 65N30

1. Introduction. The DG method is a finite element method using completely discontinuous piecewise polynomial space. It was first introduced by Reed and Hill in 1973 for the neutron linear transport [24], and later developed into the Runge-Kutta DG (RKDG) method by Cockburn and Shu for hyperbolic equations [9, 11, 12, 13]. Due to its remarkable advantage such as flexibility for arbitrarily unstructured meshes, the efficiency in parallel implementation, and the ability to easily handle complex geometries or interfaces and accommodate arbitrary $h-p$ adaptivity, the DG method has found wide applications in solving various differential equations. We refer to [10, 26] and their references for the development and survey of DG methods.

The mathematical study for DG methods can be traced back as early as 1974 by LeSaint and Raviart in [19], where a convergence rate of $\mathcal{O}(h^k)$ for $\mathbb{P}^k$ elements on general triangulations and of $\mathcal{O}(h^{k+1})$ for $\mathbb{Q}^k$ elements (i.e., tensor product bi-k polynomial spaces) on Cartesian grids were proved in the standard L^2 norm. Later, Johnson and Pitkaranta [18] proved a rate of $\mathcal{O}(h^{k+\frac{1}{2}})$ in a mesh-dependent norm for $\mathbb{P}^k$ elements on general triangulations, which was confirmed to be optimal by Peterson in [23]. Richter [25] obtained the optimal rate of convergence of $\mathcal{O}(h^{k+1})$ for some specially structured two-dimensional non-Cartesian grids under the assumption that all element edges are bounded away from the characteristic direction. Falk and Richter [15] investigated the DG method for Friedrich systems and proved a convergence rate of $\mathcal{O}(h^{k+\frac{1}{2}})$ on general triangulations for the DG approximation. A hp-version DG method has been studied in [17] and the exponential convergence was derived for piecewise analytic solution. We also refer to [20, 30] for the discussion on the interrelation between the mesh and the order of convergence for $k=1$ and $k=0$. As for the one-dimensional and some multidimensional problems using $\mathbb{Q}^k$ elements, optimal a priori error estimates of order $\mathcal{O}(h^{k+1})$ was proved for DG methods by

* This work is supported in part by the National Natural Science Foundation of China under grants 12271049, 12131005, and the NSF grant DMS-2309249.

*School of Mathematical Science, Beijing Normal University, Beijing, 100875, China.
†Division of Applied Mathematics, Brown University, Providence, RI 02912, USA.
‡Department of Mathematics, Wayne State University, Detroit, MI 48202, USA.

using the upwind fluxes [8, 29] and upwind-biased numerical fluxes [22].

Although it is well-known that the convergence rate of $\mathcal{O}(h^{k+\frac{1}{2}})$ for the DG method using $\mathbb{P}^k$ elements can not be improved on general triangular meshes, a large amount of numerical experiments show that the convergence rate can be improved from $\mathcal{O}(h^{k+\frac{1}{2}})$ to $\mathcal{O}(h^{k+1})$ by $\frac{1}{2}$ on Cartesian grids. It has been an open question whether the optimal convergence order $\mathcal{O}(h^{k+1})$ for $\mathbb{P}^k$ element DG method holds true on Cartesian meshes. Only recently, Liu et al [21] proved that, for upwind DG method using $\mathbb{P}^k$ elements on uniform Cartesian meshes, the error in the L^2 norm achieved optimal $(k+1)$-th order convergence for linear constant hyperbolic equations. For linear variable coefficient and nonlinear cases, only lower order DG schemes (i.e., $0 \le k \le 3$ for variable coefficient case and $k = 2, 3$ for nonlinear case) were proved to be optimal in L^2 error estimates. However, the theoretical analysis for high-order DG method is still in vacancy when $\mathbb{P}^k$ elements are used to solve the linear variable coefficient and nonlinear hyperbolic problems.

The main purpose of current paper is to establish a unified analysis for $\mathbb{P}^k$ element DG method for 2-D linear and nonlinear hyperbolic equations on uniform Cartesian meshes, where both optimal error estimates and superconvergence properties for the DG solution are investigated. One contribution of this study is to provide a firm answer, with a rigorous mathematical proof, that the optimal convergence order $\mathcal{O}(h^{k+1})$ for DG method solving both linear and nonlinear equations holds true when $\mathbb{P}^k, k \ge 0$ elements and uniform Cartesian meshes are used. Another contribution is the discovery of some new superconvergence phenomena for the $\mathbb{P}^k$ element DG method. Albeit with considerable interest in analyzing superconvergence properties of DG methods (see, e.g., [1, 2, 3, 5, 6, 4, 7, 16, 27, 28]), all the studies are based on one-dimensional problems and multi-dimensional $\mathbb{Q}^k$ elements. To the best of our knowledges, no superconvergence result for the $\mathbb{P}^k$ element DG method has been reported yet in the literature. In this paper, superconvergence of errors for the cell average and for the downwind edge average is established for the first time, with an order of $\mathcal{O}(h^{k+2})$.

To end with this introduction, we would like to pointed out that the theoretical analysis for the $\mathbb{P}^k$ elements is much more difficult and sophisticated than the counterpart $\mathbb{Q}^k$ elements, whose degree of freedom (i.e., $(k+1)^2$) almost doubles that of the $\mathbb{P}^k$ elements $(k+1)(k+2)/2$. The deficiency of degree of freedom makes the error analysis (i.e., construction of the projection) for $\mathbb{P}^k$ elements elusive. To cope with this problem, we first construct a specially designed projection of the exact solution and then use the idea of correction function to correct the error between the DG solution and the special projection. The construction of correction functions finally yields our desired optimal error estimates and superconvergence results for the $\mathbb{P}^k$ element DG method.

The remainder of the paper is organized as follows. In Section 2, we study the optimal error estimates and superconvergence property for the semi-discrete DG scheme solving linear variable coefficient hyperbolic equations. In Section 3, we provide the proof of the optimal error estimates and superconvergence of DG method for nonlinear hyperbolic equations. Some numerical examples are provided in Section 4. Finally, we conclude and give a few perspectives for future work in Section 5.

2. DG method for linear hyperbolic equations. In this section, we present and analyze the DG method for the two-dimensional linear hyperbolic conservation laws

$$\begin{aligned} &u_t + (\alpha u)_x + (\beta u)_y = 0, \qquad (x,y) \in \Omega, t \in (0,T], \\ &u(x,y,0) = u_0(x,y), \end{aligned} \tag{2.1}$$

where $\alpha = \alpha(x,y), \beta = \beta(x,y)$ are smooth function. For simplicity, we assume $\Omega = [0, 2\pi]^2$ and the periodic boundary condition is satisfied. The assumption of the boundary condition is not essential since the analysis can be applied to other boundary conditions such as the Dirichlet boundary condition.

2.1. DG schemes. Let $0 = x_{\frac{1}{2}} < x_{\frac{3}{2}} < \cdots < x_{m+\frac{1}{2}} = 2\pi$ and $0 = y_{\frac{1}{2}} < y_{\frac{3}{2}} < \cdots < y_{n+\frac{1}{2}} = 2\pi$. For any positive integer r, we define $\mathbb{Z}_r = \{1, 2, \ldots, r\}$, and denote by $\mathcal{T}_h$ the rectangular partition of Ω. That is,

$$\mathcal{T}_h = \{\tau_{i,j} = \tau_i^x \times \tau_j^y : \tau_i^x = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}], \tau_j^y = [y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}], (i,j) \in \mathbb{Z}_m \times \mathbb{Z}_n\}.$$

We denote $h_i^x = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}, h_j^y = y_{j+\frac{1}{2}} - y_{j-\frac{1}{2}}$, and $h = \max(h_i^x, h_j^y)$ is the maximal length of all edges, and $x_i = \frac{x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}}{2}, y_j = \frac{y_{j+\frac{1}{2}} - y_{j-\frac{1}{2}}}{2}$ the cell center of τ_i^x, τ_j^y, respectively.

Define the finite element space

$$V_h = \{v : \ v|_\tau \in \mathbb{P}^k(x,y), \tau \in \mathcal{T}_h\},$$

where $\mathbb{P}^k$ denotes the space of polynomials of degree at most k with coefficients as functions of t. The DG solution for (2.1) is to find a $u_h \in V_h$ such that

$$a_\tau(u_h, v) = 0, \quad \forall \tau \in \mathcal{T}_h, v \in V_h, \tag{2.2}$$

where for any $\tau = \tau_{i,j} \in \mathcal{T}_h, (i,j) \in \mathbb{Z}_m \times \mathbb{Z}_n$,

$$\begin{aligned} a_\tau(u_h, v) = & \int_{\tau_{i,j}} \partial_t u_h v dxdy - \int_{\tau_{i,j}} \alpha u_h v_x dxdy - \int_{\tau_{i,j}} \beta u_h v_y dxdy \\ & + \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} \left(\alpha \hat{u}_h(x_{i+\frac{1}{2}}, y) v(x_{i+\frac{1}{2}}^-, y) - \alpha \hat{u}_h(x_{i-\frac{1}{2}}, y) v(x_{i-\frac{1}{2}}^+, y) \right) dy \\ & + \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \left(\beta \tilde{u}_h(x, y_{j+\frac{1}{2}}) v(x, y_{j+\frac{1}{2}}^-) - \beta \tilde{u}_h(x, y_{j-\frac{1}{2}}) v(x, y_{j-\frac{1}{2}}^+) \right) dx. \end{aligned}$$

Here for any function v, $v(x_{i-\frac{1}{2}}^-, \cdot), v(x_{i-\frac{1}{2}}^+, \cdot)$ denote the left and right limits of v across $x_{i-\frac{1}{2}}$, respectively, and $\hat{u}_h, \tilde{u}_h$ are numerical fluxes. In this paper, we consider the upwind numerical fluxes $\hat{u}_h, \tilde{u}_h$, which are defined by

$$\hat{u}_h(x_{i+\frac{1}{2}}, y) = \begin{cases} u_h(x_{i+\frac{1}{2}}^-, y), & \text{if } \alpha(x_{i+\frac{1}{2}}, y) \geq 0, \\ u_h(x_{i+\frac{1}{2}}^+, y), & \text{if } \alpha(x_{i+\frac{1}{2}}, y) < 0, \end{cases}$$

$$\tilde{u}_h(x, y_{j+\frac{1}{2}}) = \begin{cases} u_h(x, y_{j+\frac{1}{2}}^-), & \text{if } \beta(x, y_{j+\frac{1}{2}}) \geq 0, \\ u_h(x, y_{j+\frac{1}{2}}^+), & \text{if } \beta(x, y_{j+\frac{1}{2}}) < 0. \end{cases}$$

Denoting $v^\pm|_{i+\frac{1}{2},y} = v(x_{i+\frac{1}{2}}^\pm, y), v^\pm|_{x,j+\frac{1}{2}} = v(x, y_{j+\frac{1}{2}}^\pm)$, $\{v\}$ and $[v]$ the average and jump of v, respectively. That is,

$$\begin{aligned} &\{v\}_{i+\frac{1}{2},y} = \frac{1}{2}(v^+ + v^-)|_{i+\frac{1}{2},y}, \quad \{v\}_{x,j+\frac{1}{2}} = \frac{1}{2}(v^+ + v^-)|_{x,j+\frac{1}{2}}, \\ &[v]_{i+\frac{1}{2},y} = (v^+ - v^-)|_{i+\frac{1}{2},y}, \qquad [v]_{x,j+\frac{1}{2}} = (v^+ - v^-)|_{x,j+\frac{1}{2}}. \end{aligned}$$

Let

$$a(u,v)=\sum_{\tau\in\mathcal{T}_h}a_\tau(u,v),\quad (u,v)_\tau=\int_\tau uvdxdy,\ (u,v)=\sum_{\tau\in\mathcal{T}_h}(u,v)_\tau.$$

By a direct calculation, we have

$$\begin{aligned}a(v,v)&=(v_t,v)+\frac{1}{2}(\alpha_x+\beta_y,v^2)\\&\quad+\frac{1}{2}\int_0^{2\pi}\left(\sum_{i=1}^{m}\alpha(\{v\}-\hat{v})[v]\big|_{i+\frac{1}{2},y}dy+\sum_{j=1}^{n}\beta(\{v\}-\tilde{v})[v]\big|_{x,j+\frac{1}{2}}dx\right).\end{aligned}$$

Due to the special choice of numerical fluxes, there holds

$$\frac{1}{2}\frac{d}{dt}\|v\|_0^2=(v_t,v)\lesssim a(v,v)+\|v\|_0^2,\quad \forall v\in V_h. \tag{2.3}$$

Especially, the L^2 stability of the upwind DG method follows, by taking $v=u_h$ in the above inequality and using the Gronwall inequality.

2.2. Error analysis. To investigate the optimal error estimates and superconvergence of DG method, our analysis is along this line: we first define a projection $P_hu\in V_h$ of the exact solution, and then construct a specially designed correction function w_h such that $|a(u-P_hu+w_h,v)|$ is of higher order, i.e.,

$$|a(u-P_hu+w_h,v)|\lesssim h^{k+1+l}\|v\|_0$$

for some positive l. Here the notation $A\lesssim B$ indicates $A\le cB$ with c a constant independent of the mesh size h. Finally we adopt the above weak estimate to obtain the desired optimal convergence rate and superconvergence rate for the DG approximation. If no otherwise stated, we always suppose that the mesh is uniform and that both α,β do not change sign over the whole domain Ω, i.e.,

$$\alpha(x,y)\beta(x,y)\neq 0,\quad \forall (x,y)\in\Omega.$$

Without loss of generality, we discuss the case $\alpha>0,\beta>0$. The other three cases (i.e., $\alpha>0,\beta<0$, $\alpha<0,\beta>0$, $\alpha<0,\beta<0$) follow the same arguments.

In the rest of this paper, standard notations for Sobolev spaces are adopted, such as $W^{m,p}(D)$ on subdomain $D\subset\Omega$ equipped with the norm $\|\cdot\|_{m,p,D}$ and semi-norm $|\cdot|_{m,p,D}$. When $D=\Omega$, the index D is omitted. We set $W^{m,2}(D)=H^m(D)$, $\|\cdot\|_{m,2,D}=\|\cdot\|_{m,D}$, and $|\cdot|_{m,2,D}=|\cdot|_{m,D}$.

2.2.1. A special projection of the exact solution. Let $I=[-1,1], L_{-1}=0$, and $L_p, p\ge 0$ the standard Legendre polynomial of degree p in I. Define $L_{i,p}(x), L_{j,p}(y)$ to be the Legendre polynomial of degree p on the interval τ_i^x and τ_j^y, respectively.

For any function v and integer $p\ge 0$, we denote by $\pi_p^x v$ the L^2 projection of v along the x direction onto $\mathbb{P}^p(x)$. To be more precise, $\pi_p^x v|_{\tau_i^x}\in\mathbb{P}^p(x)$ satisfies

$$\int_{\tau_i^x}(v-\pi_p^x v)wdx=0,\quad \forall w\in\mathbb{P}^p(x).$$

The L^2 projection $\pi_p^y v$ of v along the y direction can be defined similarly. By using the standard approximation theory, we have

$$\|v-\pi_p^x v\|_{0,\tau_i^x}+h^{\frac{1}{2}}\|v-\pi_p^x v\|_{0,\infty,\tau_i^x}\lesssim h^l\|v\|_{l,\tau_i^x},\ l\le p+1. \tag{2.4}$$

Furthemore, if

$$v|_{\tau_i^x} = \sum_{l=0}^{\infty} v_l L_{i,l}(x), \quad \text{with } v_l = \frac{2l+1}{h_i^x}\int_{\tau_i^x} v L_{i,l}(x)dx,$$

then

$$\pi_p^x v|_{\tau_i^x} = \sum_{l=0}^{p} v_l L_{i,l}(x), \quad (v-\pi_p^x v)|_{\tau_i^x} = \sum_{l=p+1}^{\infty} L_{i,l}(x)\frac{2l+1}{h_i^x}\int_{\tau_i^x} v L_{i,l}(x)dx. \tag{2.5}$$

In each element $\tau_{i,j}, (i,j) \in \mathbb{Z}_m \times \mathbb{Z}_n$, suppose $u(x,y,t)$ has the following Radau expansion

$$u(x,y,t)|_{\tau_{i,j}} = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty} u_{p,q}^{i,j}(L_{i,p}-L_{i,p-1})(x)(L_{j,q}-L_{j,q-1})(y), \tag{2.6}$$

where (see [4])

$$\begin{aligned} u_{p,q}^{i,j} &= u(x_{i+\frac12}^-, y_{j+\frac12}^-, t) + \sum_{l=0}^{p-1}\sum_{r=0}^{q-1}\frac{(2l+1)(2r+1)}{h_i^x h_j^y}\int_{\tau_{i,j}} u(x,y,t)L_{i,l}(x)L_{j,r}(y)dxdy \\ &\quad - \sum_{l=0}^{p-1}\frac{2l+1}{h_i^x}\int_{\tau_i^x} u(x, y_{j+\frac12}^-, t)L_{i,l}(x)dx - \sum_{r=0}^{q-1}\frac{2r+1}{h_j^y}\int_{\tau_j^y} u(x_{i+\frac12}^-, y, t)L_{j,r}(y)dy. \end{aligned}$$

In light of (2.5), we have

$$\begin{aligned} u_{p,q}^{i,j} &= (u-\pi_{p-1}^x u)(x_{i+\frac12}^-, y_{j+\frac12}^-, t) - \pi_{q-1}^y(u-\pi_{p-1}^x u)(x_{i+\frac12}^-, y_{j+\frac12}^-, t) \\ &= (\mathcal{I}-\pi_{q-1}^y)(\mathcal{I}-\pi_{p-1}^x)u(x_{i+\frac12}^-, y_{j+\frac12}^-, t). \end{aligned}$$

Here $\mathcal{I}$ denotes the identity operator. Then we conclude from (2.4) that

$$|u_{p,q}^{i,j}| \lesssim h^{p+q-1}\|u\|_{p+q,\tau_{i,j}}. \tag{2.7}$$

In addition, we have from (2.5) that

$$u_{p,q}^{i,j} = \sum_{l=p}^{\infty}\sum_{r=q}^{\infty}\frac{(2l+1)(2r+1)}{h_i^x h_j^y}\int_{\tau_{i,j}} u(x,y,t)L_{i,l}(x)L_{j,r}(y)dxdy. \tag{2.8}$$

Denote by $\mathbb{Q}^k = \mathbb{P}^k(x)\times\mathbb{P}^k(y)$ the standard space of bi-k polynomials, and define $P_h^- u \in \mathbb{Q}^k(x,y)$ the truncated Radau expansionof the exact u. That is,

$$P_h^- u\big|_{\tau_{i,j}} = \sum_{p=0}^{k}\sum_{q=0}^{k} u_{p,q}^{i,j}(L_{i,p}-L_{i,p-1})(x)(L_{j,q}-L_{j,q-1})(y). \tag{2.9}$$

We split $P_h^- u$ into three parts

$$\begin{aligned} P_h^- u|_{\tau_{i,j}} &= \left(\sum_{p+q\le k} + \sum_{p+q=k+1} + \sum_{k+2\le p+q\le (k+1)^2}\right) u_{p,q}^{i,j}(L_{i,p}-L_{i,p-1})(x)(L_{j,q}-L_{j,q-1})(y) \\ &:= Q_1 u + Q_2 u + Q_3 u. \end{aligned}$$

Define

$$\zeta_0|_{\tau_{i,j}} = \sum_{p+q=k+1} u_{p,q}^{i,j} L_{i,p}(x) L_{j,q}(y). \tag{2.10}$$

Now we define a projection $P_h u \in V_h$ of u as follows:

$$P_h u|_{\tau_{i,j}} = Q_1 u + Q_2 u - \zeta_0.$$

Given a sequence of coefficients $\{v^{i,j}\}$, define

$$\begin{aligned} &D_1 v^{i,j} = v^{i,j} - v^{i-1,j}, \quad D_2 v^{i,j} = v^{i,j} - v^{i,j-1}, \\ &D_1^l v^{i,j} = D_1(D_1^{l-1} v^{i,j}), \quad D_2^l v^{i,j} = D_2(D_2^{l-1} v^{i,j}), \quad l \geq 2. \end{aligned}$$

We next estimate the coefficients $u_{p,q}^{i,j}$ given in (2.6) and the function ζ_0.

LEMMA 2.1. *Suppose* $u_{p,q}^{i,j}$ *are the coefficients given in* (2.6). *Then for any positive integer* r', l'

$$\sum_{i=1}^{m} \sum_{j=1}^{n} |D_2^{r'} D_1^{l'} u_{p,q}^{i,j}|^2 \lesssim h^{2(p+q+r'+l'-1)} \|\partial_x^{p+l'} \partial_y^{q+r'} u\|_0^2. \tag{2.11}$$

Proof. First, given any fixed positive integer r, l, we denote

$$v^{i,j} := v_{l,r}^{i,j} = \frac{4}{h_i^x h_j^y} \int_{\tau_i^x} \int_{\tau_j^y} u(x,y,t) L_{i,l}(x) L_{j,r}(y) dx dy.$$

Let $\bar{h} = h/2$. Since the mesh is uniform, we have, by a scaling from τ_i^x, τ_j^y to $[-1, 1]$,

$$\begin{aligned} D_2 v^{i,j} &= \int_{-1}^{1} \int_{-1}^{1} u(x_i + \bar{h}\xi, y_j + \bar{h}s) - u(x_i + \bar{h}\xi, y_{j-1} + \bar{h}s)) L_l(\xi) L_r(s) d\xi ds \\ &= c \int_{-1}^{1} \int_{-1}^{1} \partial_\xi^p \partial_s^q (u(x_i + \bar{h}\xi, y_j + \bar{h}s) - u(x_i + \bar{h}\xi, y_{j-1} + \bar{h}s)) \frac{d^{l-p}(\xi^2-1)^l}{d\xi^{l-p}} \frac{d^{r-q}(s^2-1)^r}{ds^{r-q}} ds d\xi. \end{aligned}$$

Here $c = \frac{(-1)^{p+q}}{2^{l+r} l! r!}$, $q \leq r, p \leq l$, and we have used the integration by parts in the last step. Noticing that $\partial_\xi^p \partial_s^q u = O(h^{p+q}) \partial_x^p \partial_y^q u$, we get

$$|D_2 v^{i,j}| \lesssim \frac{h^{p+q-1}}{2^{l+r} l! r!} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \int_{y_{j-\frac{3}{2}}}^{y_{j+\frac{1}{2}}} |\partial_x^p \partial_y^{q+1} u| dx dy \lesssim \frac{h^{p+q}}{2^{l+r} l! r!} \|\partial_x^p \partial_y^{q+1} u\|_{0, \tau_{i,j} \cup \tau_{i,j-1}},$$

which yields, together with (2.8),

$$|D_2 u_{p,q}^{i,j}| \lesssim \sum_{l=p}^{\infty} \sum_{r=q}^{\infty} |D_2 v^{i,j}| \lesssim h^{p+q} \|\partial_x^p \partial_y^{q+1} u\|_{0, \tau_{i,j} \cup \tau_{i,j-1}}.$$

Following the same argument, we have

$$\begin{aligned} &|D_1 u_{p,q}^{i,j}| \lesssim h^{p+q} \|\partial_x^{p+1} \partial_y^q u\|_{0, \tau_{i,j} \cup \tau_{i-1,j}}, \\ &|D_2 D_1 u_{p,q}^{i,j}| \lesssim h^{p+q+1} \sum_{i'=i,i-1} \sum_{j'=j,j-1} \|\partial_x^{p+1} \partial_y^{q+1} u\|_{0, \tau_{i',j'}}. \end{aligned}$$

Summing up all i,j, we conclude that (2.11) holds true for $(r',l')=(0,1),(1,0),(1,1)$. By the induction method, we can prove (2.11) is also valid for $r',l'\geq 2$ and we omit it here for simplicity. $\square$

Given any function w, define the difference operator of w along the x direction and y direction by

$$D_x w_i^{\pm} = w(x^{\pm}_{i+\frac12},y) - w(x^{\pm}_{i-\frac12},y),\quad D_y w_j^{\pm} = u(x,y^{\pm}_{j+\frac12}) - w(x,y^{\pm}_{j-\frac12}),$$
$$D_x^{l+1} w_i^{\pm} = D_x(D_x^l w_i^{\pm}),\quad D_y^{l+1} w_j^{\pm} = D_y(D_y^l w_j^{\pm}),\quad l\geq 1.$$

We have the following results for ζ_0.

Lemma 2.2. *Given any fixed positive integer $r\geq 1$, there hold*

$$\sum_{\tau_{i,j}} \Big|\int_{\tau_i^x}\beta\zeta_0^-|_{x,j+\frac12}L_{i,r}(x)dx - \int_{\tau_{i-1}^x}\beta\zeta_0^-|_{x,j+\frac12}L_{i-1,r}(x)dx\Big|^2 \lesssim h^{2(k+2)}\|u\|^2_{k+2}, \tag{2.12}$$

$$\sum_{\tau_{i,j}} \Big|\int_{\tau_i^x}D_y(\beta\zeta_0)_j^-L_{i,r}(x)dx - \int_{\tau_{i-1}^x}D_y(\beta\zeta_0)_j^-L_{i-1,r}(x)dx\Big|^2 \lesssim h^{2(k+3)}\|u\|^2_{k+3}, \tag{2.13}$$

$$\sum_{\tau_{i,j}} \Big|\int_{\tau_j^y}D_x(\alpha\zeta_0)_i^-L_{j,r}(y)dy\Big|^2 + \Big|\int_{\tau_j^y}\alpha\zeta_0^-|_{i+\frac12,y}dy\Big|^2 \lesssim h^{2(k+2)}\|u\|^2_{k+2}, \tag{2.14}$$

$$\sum_{\tau_{i,j}} \Big|\int_{\tau_j^y}D_x(\alpha\zeta_0)_i^-dy\Big|^2 \lesssim h^{2(k+3)}\|u\|^2_{k+2}. \tag{2.15}$$

Proof. For any fixed p,r, we denote

$$\beta^{i,j} = \int_{\tau_i^x}\beta|_{x,j+\frac12}L_{i,p}(x)L_{i,r}(x)dx = \frac{h}{2}\int_{-1}^1\beta(x_i+\frac{hs}{2},y_{j+\frac12})L_p(s)L_r(s)ds.$$

For smooth function β, we have

$$|\beta^{i,j}|\lesssim h,\quad |D_1\beta^{i,j}|+|D_2\beta^{i,j}|\lesssim h^2,\quad |D_1D_2\beta^{i,j}|\lesssim h^3. \tag{2.16}$$

Recalling the definition of ζ_0 in (2.10), we have, from a direct calculation

$$\int_{\tau_i^x}\beta\zeta_0^-|_{x,j+\frac12}L_{i,r}(x)dx - \int_{\tau_{i-1}^x}\beta\zeta_0^-|_{x,j+\frac12}L_{i-1,r}(x)dx$$
$$=\sum_{p+q=k+1}(u_{p,q}^{i,j}\beta^{i,j} - u_{p,q}^{i-1,j}\beta^{i-1,j}) = \sum_{p+q=k+1}\beta^{i,j}D_1u_{p,q}^{i,j} + u_{p,q}^{i-1,j}D_1\beta^{i,j}, \tag{2.17}$$

and thus

$$\Big|\int_{\tau_i^x}\beta\zeta_0^-|_{x,j+\frac12}L_{i,r}(x)dx - \int_{\tau_{i-1}^x}\beta\zeta_0^-|_{x,j+\frac12}L_{i-1,r}(x)dx\Big| \lesssim \sum_{p+q=k+1}(h|D_1u_{p,q}^{i,j}|+h^2|u_{p,q}^{i-1,j}|).$$

Then (2.12) follows by summing up all $\tau_{i,j}$ and using the estimates in (2.11). Following the same argument, there holds

$$\int_{\tau_i^x}D_y(\beta\zeta_0)_j^-L_{i,r}(x)dx - \int_{\tau_{i-1}^x}D_y(\beta\zeta_0)_j^-L_{i-1,r}(x)dx$$
$$=\sum_{p+q=k+1}D_2(\beta^{i,j}u_{p,q}^{i,j}) - D_2(\beta^{i-1,j}u_{p,q}^{i-1,j})$$
$$=\sum_{p+q=k+1}\beta^{i,j}D_1D_2u_{p,q}^{i,j} + D_1\beta^{i,j}D_2u_{p,q}^{i-1,j} + D_1u_{p,q}^{i,j-1}D_2\beta^{i,j} + u_{p,q}^{i-1,j-1}D_1D_2\beta^{i,j}.$$

Summing up all $\tau_{i,j}$ and using (2.11) and (2.16) again, we get (2.13) directly.

Similarly, for any fixed $q \geq 1$, we denote

$$\alpha^{i,j} = \int_{\tau_j^y} \alpha(x_{i+\frac{1}{2}}, y) L_{j,q}(y) L_{j,r}(y) dy.$$

Then

$$\int_{\tau_j^y} (\alpha\zeta_0|^-_{i+\frac{1}{2},y} - \alpha\zeta_0|^-_{i-\frac{1}{2},y}) L_{j,r}(y) dy = \sum_{p+q=k+1} \alpha^{i,j} D_1 u^{i,j}_{p,q} + u^{i-1,j}_{p,q} D_1 \alpha^{i,j}.$$

Noticing that for $r \geq 1$, we have

$$|\alpha^{i,j}| \lesssim h, \quad D_1 \alpha^{i,j} \lesssim h^2.$$

While for $r = 0$, we use the orthogonality of Legendre polynomials to get

$$\alpha^{i,j} = \int_{\tau_i^x} (\alpha(x_{i+\frac{1}{2}}, y) - \alpha(x_{i+\frac{1}{2}}, y_{j+\frac{1}{2}})) L_{j,q}(y) dx \lesssim h^2, \ D_1 \alpha^{i,j} \lesssim h^3.$$

Summing up all $\tau_{i,j}$ and using (2.11) again yields the desired results (2.14)-(2.15). $\square$

2.2.2. Construction of the correction function. Define

$$V_h^0 = \{v \in V_h : \int_\tau v = 0, \quad \forall \tau \in \mathcal{T}_h\}, \quad \bar{V}_h = \{v : v|_\tau \in \mathbb{P}^0(x, y), \quad \forall \tau \in \mathcal{T}_h\}.$$

Denote by $\bar{\alpha}, \bar{\beta} \in \bar{V}_h$ the cell average of α, β, respectively. That is,

$$\bar{\alpha}|_\tau = \frac{1}{|\tau|} \int_\tau \alpha dx dy, \quad \bar{\beta}|_\tau = \frac{1}{|\tau|} \int_\tau \beta dx dy.$$

For all $\tau = \tau_{i,j}$, define

$$\begin{aligned} b_\tau(u, v; \alpha, \beta) = &- \int_{\tau_{i,j}} \alpha u v_x dx dy + \int_{\tau_j^y} \alpha u(x^-_{i+\frac{1}{2}}, y) \left(v^-|_{i+\frac{1}{2},y} - v^+|_{i-\frac{1}{2},y}\right) dy \\ &- \int_{\tau_{i,j}} \beta u v_y dx dy + \int_{\tau_i^x} \beta u(x, y^-_{j+\frac{1}{2}}) \left(v^-|_{x,j+\frac{1}{2}} - v^+|_{x,j-\frac{1}{2}}\right) dx, \end{aligned}$$

and

$$\begin{aligned} \mathcal{H}_\tau(u, v; \alpha, \beta) &= \int_{\tau_j^y} (\alpha u^- v^-|_{i+\frac{1}{2},y} - \alpha u^- v^+|_{i-\frac{1}{2},y}) dy + \int_{\tau_i^x} (\beta u^- v^-|_{x,j+\frac{1}{2}} - \beta u^- v^+|_{x,j-\frac{1}{2}}) dx, \\ \mathcal{H}^1_\tau(u, v; \alpha, \beta) &= (u_t, v)_\tau + b_\tau(u, v; \alpha - \bar{\alpha}, \beta - \bar{\beta}) \\ &\quad + \int_{\tau_j^y} D_x(\alpha u)^-_i v^+|_{i-\frac{1}{2},y} dy + \int_{\tau_i^x} D_y(\beta u)^-_j v^+|_{x,j-\frac{1}{2}} dx. \end{aligned}$$

Given any function ψ, let $R_h\psi \in V_h^0$ and $R_h^1\psi \in V_h^0$ be two special projections of ψ such that for all $v \in V_h^0$

$$b_\tau(R_h\psi, v; \bar{\alpha}, \bar{\beta}) = -\mathcal{H}_\tau(\psi, v; \alpha, \beta), \tag{2.18}$$
$$b_\tau(R_h^1\psi, v; \bar{\alpha}, \bar{\beta}) = -\mathcal{H}^1_\tau(\psi, v; \alpha, \beta). \tag{2.19}$$

LEMMA 2.3. *The operators R_h and R_h^1 are well-defined.*

Proof. Noticing that the only difference between R_h and R_h^1 lies in the right hand side, we only prove the uniqueness of R_h since the similar argument can be applied to R_h^1. Towards this end, we need to show that the zero right hand side $\psi = 0$ will yields a zero solution $R_h\psi$=0.

Denoting $w = R_h\psi$. By choosing $v = w$ in (2.18), we easily get

$$\bar{\alpha}\int_{\tau_j^y}(w^-|_{i+\frac{1}{2},y} - w^+|_{i-\frac{1}{2},y})^2dy + \bar{\beta}\int_{\tau_i^x}(w^-|_{x,j+\frac{1}{2}} - w^+|_{x,j-\frac{1}{2}})^2dx = 0,$$

which yields

$$w^-|_{i+\frac{1}{2},y} = w^+|_{i-\frac{1}{2},y}, \quad w^-|_{x,j+\frac{1}{2}} = w^+|_{x,j-\frac{1}{2}}, \quad \forall(x,y)\in\tau.$$

Consequently,

$$\int_\tau w_x dxdy = \int_\tau w_y dxdy = 0.$$

Consequently, $w_x, w_y \in V_h^0$. On the other hand, we use the integration by parts to derive that

$$b_\tau(w,v) = (\bar{\alpha}w_x + \bar{\beta}w_y, v)_\tau.$$

By choosing $v = \bar{\alpha}w_x + \bar{\beta}w_y \in V_h^0$ in the above inequality yields

$$\bar{\alpha}w_x + \bar{\beta}w_y = 0,$$

and thus w is a constant in each element τ. Since $w \in V_h^0$, we have $w \equiv 0$. Then $w = R_h\psi$ is uniquely defined. This finishes our proof. ☐

Now we define two correction functions by

$$\zeta_1 = R_h\zeta_0, \quad \zeta_2 = R_h^1\zeta_1. \tag{2.20}$$

We have the following properties for ζ_1 and ζ_2.

LEMMA 2.4. *There hold for $l = 1, 2$ and $r = 0, 1$ that*

$$\|\zeta_l\|_0 \lesssim h^{k+l}\|u\|_{k+l}, \tag{2.21}$$

$$\Big(\sum_{\tau_{i,j}}\int_{\tau_j^y}|D_x(\alpha\zeta_l)_i^-|^2dy + \int_{\tau_i^x}|D_y(\beta\zeta_l)_j^-|^2dx\Big)^{\frac{1}{2}} \lesssim h^{k+l+\frac{1}{2}}\|u\|_{k+1+l}, \tag{2.22}$$

$$\Big(\sum_{\tau_{i,j}}|\int_{\tau_j^y}D_x^r(\alpha\zeta_l)_i^-dy|^2 + |\int_{\tau_i^x}D_y^r(\beta\zeta_l)_j^-dx|^2\Big)^{\frac{1}{2}} \lesssim h^{k+2+r}\|u\|_{k+2+r}. \tag{2.23}$$

Here for any function v, $D_x^0v_i^- = v^-|_{i+\frac{1}{2},y}, D_y^0v_j^- = v^-|_{x,j+\frac{1}{2}}$.

The proof of Lemma 2.4 is given in the appendix.

2.3. Weak estimates of the projection. With the correction function defined in (2.20), we design a special projection of u by

$$u_I^l = P_hu - w_h^l, \quad \text{with } w_h^l = \sum_{r=1}^{l}\zeta_r, \quad 1 \le l \le 2. \tag{2.24}$$

THEOREM 2.5. *Let $u \in H^{k+1+l}(\Omega), 1 \le l \le 2$ be the solution of* (2.1)*, and $u_h \in V_h$ be solution of* (2.2)*. Suppose $u_I^l \in V_h$ is defined by* (2.24)*. Then*

$$|a(u-u_I^l, v)| \lesssim h^{k+l}\|u\|_{k+l+1}\|v\|_0, \quad \forall v \in V_h. \tag{2.25}$$

Proof. Recalling the definition of $a(\cdot,\cdot)$ and using the orthogonality of $Q_3 u$, we have

$$a_\tau(P_h^- u - P_h u, v) = a_\tau(Q_3 u + \zeta_0, v) = -(\zeta_0 + Q_3 u, \alpha v_x + \beta v_y)_\tau + \mathcal{H}_\tau(\zeta_0, v; \alpha, \beta).$$

Noticing that

$$a_\tau(\zeta_r, v) = b_\tau(\zeta_r, v; \bar\alpha, \bar\beta) + \mathcal{H}_\tau^1(\zeta_r, v; \alpha, \beta),\ 1 \le r \le 2,$$

we have

$$a_\tau(P_h^- u - u_I^l, v) = b_\tau(w_h^l, v; \bar\alpha, \bar\beta) + \mathcal{H}_\tau^1(w_h^l, v; \alpha, \beta) + \mathcal{H}_\tau(\zeta_0, v; \alpha, \beta) - J_\tau.$$

where

$$J_\tau = (\zeta_0 + Q_3 u, \alpha v_x + \beta v_y)_\tau.$$

In light of (2.18)-(2.19), we have for all $v \in V_h^0$ that

$$a_\tau(P_h^- u - u_I^l, v) = \mathcal{H}_\tau^1(\zeta_l, v; \alpha, \beta) - J_\tau.$$

By using the Cauchy-Schwartz inequality and the inverse inequality,

$$\begin{aligned}|\mathcal{H}_\tau^1(\zeta_l, v; \alpha, \beta)| &\lesssim \|\partial_t \zeta_l\|_{0,\tau}\|v\|_{0,\tau} + h\|\zeta_l\|_{0,\tau}\|v\|_{1,\tau} + \|\zeta_l\|_{0,\tau}\|v\|_{0,\tau} \\ &\quad + h^{-\frac12}\|v\|_{0,\tau}\left(\int_{\tau_j^y}|D_x(\alpha\zeta_l)_i^-|^2 dy + \int_{\tau_i^x}|D_y(\beta\zeta_l)_j^-|^2 dx\right)^{\frac12}.\end{aligned}$$

On the other hand, using the orthogonality of ζ_0 and $Q_3 u$ again yields

$$|J_\tau| = |((\zeta_0 + Q_3 u), (\alpha - \pi_1^x\pi_1^y\alpha)v_x + (\beta - \pi_1^x\pi_1^y\beta)v_y)_\tau| \lesssim h^2\|\zeta_0 + Q_3 u\|_{0,\tau}\|v\|_{1,\tau}.$$

Consequently, for all $v \in V_h^0$,

$$\begin{aligned}|a_\tau(P_h^- u - u_I^l, v)| &\lesssim (\|\zeta_l\|_{0,\tau} + \|\partial_t\zeta_l\|_{0,\tau} + h\|\zeta_0 + Q_3 u\|_{0,\tau})\|v\|_{0,\tau} \\ &\quad + h^{-\frac12}\|v\|_{0,\tau}\left(\int_{\tau_j^y}|D_x(\alpha\zeta_l)_i^-|^2 dy + \int_{\tau_i^x}|D_y(\beta\zeta_l)_j^-|^2 dx\right)^{\frac12}.\end{aligned}$$

Summing up all element τ, and using (2.21)-(2.22) and the estimates of $\zeta_0, Q_3 u$, we have for all $v \in V_h^0$

$$|a(P_h^- u - u_I^l, v)| \lesssim h^{k+l}(\|u_t\|_{k+l} + \|u\|_{k+l+1})\|v\|_0 \lesssim h^{k+l}\|u\|_{k+1+l}\|v\|_0.$$

For all $v_0 \in \bar V_h$, a direct calculation yields

$$a(P_h^- u - u_I^l, v_0) = \sum_{\tau_{i,j}}\sum_{r=0}^{l} v_0\big(\int_{\tau_j^y} D_x(\alpha\zeta_r)_i^- dy + \int_{\tau_i^x} D_y(\beta\zeta_r)_j^- dx\big).$$

If $l=1$, we use (2.15), (2.22) and the Cauchy-Schwarz inequality to derive that

$$
\begin{aligned}
|a(P_h^- u - u_I^l, v_0)| &\lesssim \big(\sum_{\tau_{i,j}} h v_0^2\big)^{\frac{1}{2}} \left(\sum_{r=0}^{l} \sum_{\tau_{i,j}} \big(\int_{\tau_j^y} |D_x(\alpha\zeta_r)_i^-|^2 dy + \int_{\tau_i^x} |D_y(\beta\zeta_r)_j^-|^2 dx \big) \right)^{\frac{1}{2}} \\
&\lesssim h^{k+1} \|u\|_{k+2} \|v_0\|_0.
\end{aligned}
$$

By using (2.15), (2.23) and the Cauchy-Schwarz inequality again, we have for $l=2$,

$$
\begin{aligned}
|a(P_h^- u - u_I^l, v_0)| &\lesssim \big(\sum_{\tau_{i,j}} v_0^2\big)^{\frac{1}{2}} \left(\sum_{r=0}^{l} \sum_{\tau_{i,j}} | \int_{\tau_j^y} D_x(\alpha\zeta_r)_i^- dy|^2 + | \int_{\tau_i^x} D_y(\beta\zeta_r)_j^- dx|^2 \right)^{\frac{1}{2}} \\
&\lesssim h^{k+2} \|u\|_{k+3} \|v_0\|_0.
\end{aligned}
$$

Consequently,

$$
|a(P_h^- u - u_I^l, v_0)| \lesssim h^{k+l} \|u\|_{k+1+l} \|v_0\|_0, \quad \forall v_0 \in \bar{V}_h.
$$

Since any function $v \in V_h$ can be decomposed into $v = v_1 + v_0$ with $v_1 \in V_h^0$ and $v_0 \in \bar{V}_h$, then we conclude from the last two inequalities that

$$
|a(P_h^- u - u_I^l, v)| = |a(P_h^- u - u_I^l, v_0) + a(P_h^- u - u_I^l, v_1)| \lesssim h^{k+l} \|u\|_{k+l+1} \|v\|_0.
$$

Note that (see [4])

$$
|a(u - P_h^- u, v)| \lesssim h^{k+1+r} \|v\|_0 \|u\|_{k+r+2}, \quad \forall v \in V_h, \ 1 \le r \le k.
$$

Consequently,

$$
|a(u - u_I^l, v)| = |a(u - P_h^- u, v) + a(P_h^- u - u_I^l, v)| \lesssim h^{k+l} \|u\|_{k+l+1}, \quad \forall v \in V_h, 1 \le l \le 2.
$$

This finishes the proof of (2.25). ∎

2.4. Optimal error estimates and superconvergence. Define the cell average error and downwind edge average error as follows:

$$
e_c := \left(\frac{1}{nm} \sum_{\tau \in \mathcal{T}_h} \big(\frac{1}{|\tau|} \int_\tau (u - u_h) dx dy \big)^2 \right)^{\frac{1}{2}},
$$

$$
e_d := \left(\frac{1}{nm} \sum_{\tau_{i,j} \in \mathcal{T}_h} \big(\frac{1}{h_j^y} \int_{\tau_j^y} (u - u_h)^-_{i+\frac{1}{2},y} dy \big)^2 + \big(\frac{1}{h_i^x} \int_{\tau_i^x} (u - u_h)^-_{x,j+\frac{1}{2}} dx \big)^2 \right)^{\frac{1}{2}}.
$$

Now we are ready to present the optimal error estimates in L^2 norm and superconvergence for the cell error and downwind edge average error.

THEOREM 2.6. *Let $u \in H^{k+2}(\Omega)$ be the solution of* (2.1) *and $u_I^l \in V_h$ be the special projection of u defined by* (2.24). *Assume that $u_h \in V_h$ is the solution of* (2.2) *using uniform meshes with the initial value chosen as $u_h(x,y,0) = u_I^l(x,y,0)$. Then*

$$
\|(u - u_h)(\cdot, t)\|_0 \lesssim h^{k+1} \|u\|_{k+2}. \tag{2.26}
$$

Furthermore, if $u \in H^{k+3}(\Omega)$, then for $k \geq 1$,

$$e_{u,c} \lesssim h^{k+2}\|u\|_{k+3}, \quad e_d \lesssim h^{k+2}\|u\|_{k+3}. \tag{2.27}$$

Proof. Choosing $v = u_h - u_I^l$ in (2.3) and using the orthogonality $a(u-u_h, v) = 0$ for all $v \in V_h$, we have

$$\begin{aligned}\|u_I^l - u_h\|_0 \frac{d}{dt}\|u_I - u_h\|_0 &\leq \left|a(u_h - u_I^l, u_I^l - u_h)\right| + \|u_I^l - u_h\|_0^2 \\ &= \left|a(u - u_I^l, u_I - u_h)\right| + \|u_I^l - u_h\|_0^2.\end{aligned}$$

Due to the special choice of initial values and the estimates in (2.25), we have

$$\|u_I^l - u_h\|_0 \lesssim h^{k+l}\|u\|_{k+l+1}, \quad 1 \leq l \leq 2. \tag{2.28}$$

Consequently, if $u \in H^{k+2}$, then

$$\|u - u_h\|_0 \leq \|u_I^1 - u\|_0 + \|u_I^1 - u_h\|_0 \lesssim h^{k+1}\|u\|_{k+2}.$$

This finishes the proof of (2.26).

Noticing that

$$\int_\tau (u - u_I^l)dxdy = \int_\tau (u - P_h u + w_h^l)dxdy = 0,$$

we have

$$e_c = \left(\frac{1}{nm}\sum_{\tau\in\mathcal{T}_h}\Big(\frac{1}{|\tau|}\int_\tau (u_I^l - u_h)dxdy\Big)^2\right)^{\frac{1}{2}} \lesssim \|u_I^l - u_h\|_0,$$

which yields (together with (2.28)) the first inequality of (2.27). Similarly, using the orthogonality of Legendre polynomials,

$$\begin{aligned}&\int_{\tau_j^y}(u - u_I^l)(x_{i+\frac{1}{2}}^-, y)dy = \int_{\tau_j^y}(P_h^- u - u_I^l)(x_{i+\frac{1}{2}}^-, y)dy = \int_{\tau_j^y} w_h^l(x_{i+\frac{1}{2}}^-, y)dy, \\ &\int_{\tau_i^x}(u - u_I^l)(x, y_{j+\frac{1}{2}}^-)dx = \int_{\tau_j^y} w_h^l(x, y_{j+\frac{1}{2}}^-)dx.\end{aligned}$$

Consequently,

$$\begin{aligned}|e_d|^2 &= \frac{1}{nm}\sum_{\tau_{i,j}}\Big|\frac{1}{h}\int_{\tau_j^y}(w_h^l + u_I^l - u_h)(x_{i+\frac{1}{2}}^-, y)dy\Big|^2 + \Big|\frac{1}{h}\int_{\tau_i^x}(w_h^l + u_I^l - u_h)(x, y_{j+\frac{1}{2}}^-)dx\Big|^2 \\ &\lesssim \|u_I^l - u_h\|_0^2 + \sum_{\tau_{i,j}}\Big|\int_{\tau_j^y} w_h^l(x_{i+\frac{1}{2}}^-, y)dy\Big|^2 + \Big|\int_{\tau_i^x} w_h^l(x, y_{j+\frac{1}{2}}^-)dx\Big|^2.\end{aligned}$$

Then the second inequality of (2.27) follows by choosing $r = 0$ in (2.23) and $l = 2$ in (2.28). ∎

3. DG method for nonlinear hyperbolic equations. In this section, we consider the DG method for the two-dimensional nonlinear hyperbolic conservation laws with periodic boundary condition

$$
\begin{aligned}
&u_t + f(u)_x + g(u)_y = 0, \qquad (x,y) \in \Omega,\ t \in (0,T],\\
&u(x,y,0) = u_0(x,y),
\end{aligned} \tag{3.1}
$$

where $f(u), g(u)$ are smooth functions. In this paper, we suppose that $f'(u)$ and $g'(u)$ do not change sign.

The DG solution for (3.1) is to find $u_h \in V_h$ such that

$$
\begin{aligned}
(\partial_t u_h, v)_{\tau_{i,j}} = (f(u_h), v_x)_{\tau_{i,j}} - \int_{\tau_j^y} \left(\hat{f}(u_h) v^-|_{i+\frac{1}{2},y} - \hat{f}(u_h) v^-|_{i-\frac{1}{2},y} \right) dy\\
+ (g(u_h), v_y)_{\tau_{i,j}} - \int_{\tau_i^x} \left(\tilde{g}(u_h) v^-|_{x,j+\frac{1}{2}} - \tilde{g}(u_h) v^+|_{x,j+\frac{1}{2}} \right) dx,
\end{aligned} \tag{3.2}
$$

where $\hat{f}(u_h), \tilde{g}(u_h)$ denote the numerical fluxes, which are single-valued functions defined at each cell interface and in general depends on the values of the numerical solution u_h from both sides of the interface. Here we still choose the upwind monotone numerical fluxes, i.e.,

$$
\hat{f}(u_h) = \begin{cases} f(u_h^-), & \text{if } f'(u) \ge 0,\\ f(u_h^+), & \text{if } f'(u) < 0, \end{cases} \qquad \tilde{g}(u_h) = \begin{cases} g(u_h^-), & \text{if } g'(u) \ge 0,\\ g(u_h^+), & \text{if } g'(u) < 0, \end{cases}
$$

Without loss of generality, we assume that $f'(u) > 0, g'(u) > 0$. To deal with the nonlinearity, we first adopt the Taylor expansion for $f(u)$ and $g(u)$,

$$
\begin{aligned}
f(u) &= f(u_h) + f'(u)(u - u_h) - \frac{1}{2}\bar{f}''_u (u - u_h)^2,\\
g(u) &= g(u_h) + g'(u)(u - u_h) - \frac{1}{2}\bar{g}''_u (u - u_h)^2,
\end{aligned}
$$

where

$$
\bar{f}''_u = f''(\theta_1 u + (1-\theta_1)u_h), \quad \bar{g}''_u = g''(\theta_1 u + (1-\theta_1)u_h),\ 0 \le \theta_1, \theta_2 \le 1.
$$

Second, we need a priori assumption for the error $u - u_h$, i.e.,

$$
\|u - u_h\|_{0,\infty} \lesssim h. \tag{3.3}
$$

This assumption is frequently used in the DG error analysis for nonlinear problems (see, e.g., [5, 22]). We will justify this assumption for $k \ge 1$.

Third, we slightly modified the correction function ζ_1, ζ_2.

$$
\begin{aligned}
b_\tau(\zeta_1, v; \bar{f}'(u), \bar{g}'(u)) = -\mathcal{H}_\tau(\zeta_0, v; f'(u), g'(u)),\\
b_\tau(\zeta_2, v; \bar{f}'(u), \bar{g}'(u)) = -\mathcal{H}_\tau^1(\zeta_1, v; f'(u), g'(u)).
\end{aligned}
$$

Let $u_I^l = P_h u - w_h^l$ with $w_h^l = \sum_{r=1}^l \zeta_r$, and denote

$$
e = u - u_h, \quad \eta = u - u_I^l, \quad \xi = u_I^l - u_h.
$$

Following the same argument as what we did for the linear problems, we have

$$
|A(u - u_I^l, v)| \lesssim h^{k+l} \|u\|_{k+l+1} \|v\|_0, \quad 1 \le l \le 2, \tag{3.4}
$$

where

$$A(w,v) = (\partial_t w, v) - (f'(u)w, v_x) + \sum_{\tau_{i,j}} \int_{\tau_j^y} \left(f'(u)\hat{w}v^-|_{i+\frac{1}{2},y} - f'(u)\hat{w}v^-|_{i-\frac{1}{2},y} \right) dy$$
$$- (g'(u)w, v_y) + \sum_{\tau_{i,j}} \int_{\tau_i^x} \left(g'(u)\tilde{w}v^-|_{x,j+\frac{1}{2}} - g'(u)\tilde{w}v^+|_{x,j+\frac{1}{2}} \right) dx.$$

THEOREM 3.1. *Let $u \in H^{k+2}(\Omega)$ be the solution of (3.1) and $u_I^l \in V_h$ be the special projection of u. Assume that $u_h \in V_h$ is the solution of (3.2) using uniform meshes with the initial value chosen as $u_h(x,y,0) = u_I^l(x,y,0)$. Then*

$$\|(u - u_h)(\cdot, t)\|_0 \lesssim h^{k+1}\|u\|_{k+2}. \tag{3.5}$$

Furthermore, if $u \in H^{k+3}(\Omega)$, there hold for $k \geq 1$

$$e_c \lesssim h^{\min(2k,k+2)}\|u\|_{k+3}, \quad e_d \lesssim h^{\min(2k,k+2)}\|u\|_{k+3}. \tag{3.6}$$

Proof. Noticing that the exact solution u also satisfy the equation (3.2), we have the following error equation

$$2A(e,v) = (\bar{f}_u'' e^2, v_x) + (\bar{g}_u'' e^2, v_y) + \sum_{\tau_{i,j}} \int_{\tau_j^y} (\bar{f}_u'' e^2)^- [v]|_{i+\frac{1}{2},y} dy + \int_{\tau_i^x} (\bar{g}_u'' e^2)^- [v]|_{x,j+\frac{1}{2}}) dx.$$

By denoting

$$\mathcal{J} = (\bar{f}_u'' e^2, \xi_x) + (\bar{g}_u'' e^2, \xi_y) + \sum_{\tau_{i,j}} \int_{\tau_j^y} (\bar{f}_u'' e^2)^- [\xi]|_{i+\frac{1}{2},y} dy + \int_{\tau_i^x} (\bar{g}_u'' e^2)^- [\xi]|_{x,j+\frac{1}{2}}) dx,$$

we have

$$A(\xi,\xi) = A(\eta,\xi) + \frac{\mathcal{J}}{2}. \tag{3.7}$$

We next estimate the term $\mathcal{J}$. By using the Cauchy-Schwarz inequality,

$$\mathcal{J} \lesssim \|e\|_{0,\infty}(\|e\|_0\|\xi\|_1 + \|\xi\|_{\Gamma_h}^2 + \|\xi\|_{\Gamma_h}\|\eta\|_{\Gamma_h}), \tag{3.8}$$

where

$$\|\xi\|_{\Gamma_h}^2 = \sum_{\tau_{i,j}} \int_{\tau_j^y} (\xi^2(x_{i+\frac{1}{2}}^-, y) + \xi^2(x_{i+\frac{1}{2}}^+, y)) dy + \int_{\tau_i^x} (\xi^2(x, y_{j+\frac{1}{2}}^-) + \xi^2(x, y_{j+\frac{1}{2}}^+)) dx.$$

Using the inverse inequality and the approximation property of u_I^l, we have

$$\|\xi\|_{\Gamma_h}^2 \lesssim h^{-1}\|\xi\|_0^2, \quad \|\eta\|_{\Gamma_h}^2 \lesssim h^{2k+1}\|u\|_{k+1}^2.$$

Substituting the above inequality into (3.8) and using the inverse inequality yields

$$\mathcal{J} \lesssim h^{-1}\|e\|_{0,\infty}(\|e\|_0\|\xi\|_0 + \|\xi\|_0^2 + h^{k+1}\|u\|_{k+1}\|\xi\|_0).$$

Recalling the definition of $A(\cdot,\cdot)$, we obtain from a direct calculation,

$$\begin{aligned} A(\xi,\xi) &= (\xi_t, \xi) + \frac{1}{2}(\partial_x f'(u) + \partial_y g'(u), \xi^2) \\ &\quad + \sum_{\tau_{i,j}} \int_{\tau_j^y} f'(u)[\xi]^2|_{i+\frac{1}{2},y} dy + \int_{\tau_i^x} f'(u)[\xi]^2|_{x,j+\frac{1}{2}} dx. \end{aligned} \tag{3.9}$$

Combining (3.4), (3.7), (3.9) and the estimates of $\mathcal{J}$ together, we get

$$(\xi_t,\xi) \lesssim (1+h^{-1}\|e\|_{0,\infty})\|\xi\|_0^2 + h^{2k}\|e\|_{0,\infty}^2\|u\|_{k+1}^2 + h^{k+1+l}\|u\|_{k+2+l}\|\xi\|_0. \tag{3.10}$$

In light of (3.3), we have

$$\frac{1}{2}\frac{d}{dt}\|\xi\|_0^2 \le (\xi_t,\xi) \lesssim \|\xi\|_0^2 + h^{2k+2}\|u\|_{k+1}^2,$$

and thus

$$\|\xi\|_0 \lesssim h^{k+1}\|u\|_{k+2}.$$

Then (3.5) follows from the triangle inequality.

To derive the superconvergence result, we first adopt the inverse inequality from the last inequality to derive

$$\|\xi\|_{0,\infty} \lesssim h^k\|u\|_{k+2}.$$

Then

$$\|e\|_{0,\infty} \le \|\xi\|_{0,\infty} + \|\eta\|_{0,\infty} \lesssim h^k\|u\|_{k+2}.$$

Substituting the above inequality into (3.10), we have

$$(\xi_t,\xi) \lesssim (1+h^{k-1})\|\xi\|_0^2 + h^{4k}\|u\|_{k+1}^2 + h^{k+2}\|u\|_{k+3}\|\xi\|_0.$$

Consequently, for sufficiently small h, we have from the Gronwall inequality that

$$\|\xi\|_0 \lesssim h^{\min(2k,k+2)}\|u\|_{k+3}.$$

Following the same argument as that in Theorem 2.6, we obtain (3.6). □

REMARK 3.2. *Now we show that the a priori assumption* (3.3) *is reasonable and justify it. Actually, we have from the optimal error estimate* (3.5) *that*

$$\|u_I - u_h\|_0 \lesssim h^{k+1}.$$

Using the inverse inequality and triangle inequality, there holds

$$\begin{aligned}\|u-u_h\|_{0,\infty} &\le \|u-u_I\|_{0,\infty} + \|u_I-u_h\|_{0,\infty}\\ &\lesssim \|u-P_hu\|_{0,\infty} + h^{-1}\|u_I-u_h\|_{\infty} + h^{-1}\|\zeta_1+\zeta_2\|_0\\ &\lesssim h^k.\end{aligned}$$

Consequently, the assumption (3.3) *holds true for* $k \ge 1$*.*

4. Numerical results. In this section, we present some numerical experiments to verify our theoretical findings. In our numerical experiments, we adopt the upwind DG scheme using $\mathbb{P}^k$ elements with $0 \le k \le 3$ for solving the linear constant coefficient equation, the linear variable coefficient equation, and the nonlinear equation, and we test the standard L^2 error $\|e\|_0$, the error for the cell average e_c, and the error for the downwind edge average e_d on both uniform and nonuniform meshes. The uniform mesh is obtained by equally dividing the computational domain $[0,2\pi]\times[0,2\pi]$ into $N\times N$ rectangles. Nonuniform meshes of $N\times N$ rectangles are obtained by randomly and independently perturbing each node in the x and y axes of a uniform mesh by up to some percentage. That is,

$$x_i = \frac{2\pi i}{N} + \frac{\pi}{2N} randn(), \quad y_j = \frac{2\pi j}{N} + \frac{\pi}{2N} randn(), \ 1 \le i, j \le N-1.$$

Here $randn$ denotes the random number in the interval $[-1, 1]$. To diminish the time discretization error, forth-order Runge-Kutta method is used with the time step size $\Delta t = 0.05 h_{min}^2$ with $h_{min} = \min(h_i^x, h_j^y)$.

Example 1. We first consider the following linear constant hyperbolic equation

$$\begin{cases} u_t + u_x - 2u_y = f, & (x, y, t) \in [0, 2\pi] \times [0, 2\pi] \times (0, 1), \\ u(x, y, 0) = \sin(x + y) \end{cases}$$

with the periodic boundary condition

$$u(0, y, t) = u(2\pi, y, t) \quad \text{and} \quad u(x, 0, t) = u(x, 2\pi, t).$$

The right-hand side function f is chosen such that the exact solution is

$$u(x, y, t) = \sin(x + y - 2t).$$

Listed in Tables 4.1-4.2 are errors and the corresponding convergence rates calculated from the DG method for $k = 0, 1, 2, 3$ in uniform meshes and nonuniform meshes, respectively. We observe an optimal convergence rate of $k + 1$ for the L^2 error $\|e\|_0$, a superconvergence rate of $k + 2$ for the downwind edge average error e_d for $k \ge 1$ in uniform meshes, which confirm our theoretical results in Theorem 2.6. As for the cell average error e_c, Table 4.1 demonstrate a superconvergence rate of $2k + 1$ for $k = 1, 2$ and $2k$ for $k = 3$ in uniform meshes, which is better than the theoretical result $k + 2$ given in (2.27). While in nonuniform meshes, Table 4.2 shows the optimal convergence rate $k + 1$ for all the errors $\|e\|_0, e_c, e_d$, and the superconvergence phenomena for the cell average error and edge average error disappear.

Example 2. We solve the following linear variable equation with periodic boundary condition

$$\begin{cases} u_t + (\alpha(x, y)u)_x + (\beta(x, y)u)_y = f, & (x, y, t) \in [0, 2\pi] \times [0, 2\pi] \times (0, 1), \\ u(x, y, 0) = \sin(x + y), \end{cases}$$

where $\alpha(x, y) = \sin(x + y) + 2, \beta(x, y) = \cos(x + y) - 2$, and f is chosen such that the exact solution is $u(x, y, t) = \sin(x + y - 2t)$.

The computational results in uniform and nonuniform meshes are given in Table 4.3 and Table 4.4, respectively, from which, we can observe similar results as given in Example 1 for the linear constant coefficient problem, which indicates that optimal error estimates in the L^2 norm and superconvergence results for the cell/edge average errors hold true for the linear variable coefficient equation in uniform meshes.

Example 3. We consider the following nonlinear equation with periodic boundary condition

$$\begin{cases} u_t + (u^3)_x - (e^u)_y = f, & (x, y, t) \in [0, 2\pi] \times [0, 2\pi] \times (0, 1), \\ u(x, y, 0) = \sin(x + y). \end{cases}$$

Table 4.1

Errors and convergence rates for the constant coefficient equation in uniform meshes

	$N \times N$	$\|e\|_0$	rate	e_c	rate	e_d	rate
	16×16	1.49e-0	–	2.08e-1	–	3.51e-1	–
	32×32	8.32e-1	0.84	1.20e-1	0.80	1.95e-1	0.84
$k = 0$	64×64	4.42e-1	0.91	6.44e-2	0.90	1.03e-1	0.92
	128×128	2.28e-1	0.96	3.34e-2	0.95	5.32e-2	0.96
	256×256	1.16e-1	0.98	1.70e-2	0.97	2.70e-2	0.98
	16×16	8.42e-2	–	4.19e-3	–	1.01e-2	–
	32×32	2.07e-2	2.03	5.73e-4	2.87	1.38e-3	2.87
$k = 1$	64×64	5.14e-3	2.01	7.34e-5	2.96	1.77e-4	2.96
	128×128	1.28e-3	2.00	9.24e-6	2.99	2.23e-5	2.99
	256×256	3.21e-4	2.00	1.16e-6	3.00	2.80e-6	3.00
	16×16	4.72e-3	–	2.05e-5	–	1.67e-4	–
	32×32	5.85e-4	3.01	6.54e-7	4.97	1.08e-5	3.95
$k = 2$	64×64	7.29e-5	3.00	2.06e-8	4.99	6.79e-7	3.99
	128×128	9.11e-6	3.00	6.44e-10	5.00	4.25e-8	4.00
	256×256	1.14e-6	3.00	2.00e-11	5.01	2.66e-9	4.00
	16×16	2.84e-4	–	1.54e-7	–	7.17e-6	–
	32×32	1.75e-5	4.02	3.53e-9	5.45	1.93e-7	5.22
$k = 3$	64×64	1.09e-6	4.00	5.53e-11	6.00	6.46e-9	4.90
	128×128	6.82e-8	4.00	8.62e-13	6.00	1.99e-10	5.02
	256×256	4.26e-9	4.00	1.64e-14	5.71	6.09e-12	5.03

Table 4.2

Errors and convergence rates for the constant coefficient equation in nonuniform meshes

	$N \times N$	$\|e\|_0$	rate	e_c	rate	e_d	rate
	16×16	1.54e-0	–	2.13e-1	–	3.55e-1	–
	32×32	8.72e-1	0.88	1.25e-1	0.83	2.00e-1	0.89
$k = 0$	64×64	4.65e-1	0.96	6.75e-2	0.94	1.07e-1	0.96
	128×128	2.40e-1	0.98	3.50e-2	0.97	5.49e-2	0.98
	256×256	1.22e-1	0.98	1.78e-2	0.97	2.79e-2	0.97
	16×16	1.04e-1	–	6.80e-3	–	1.35e-2	–
	32×32	2.35e-2	2.12	1.25e-3	2.42	2.79e-3	2.25
$k = 1$	64×64	5.83e-3	2.03	2.54e-4	2.32	5.59e-4	2.34
	128×128	1.50e-3	2.15	5.81e-5	2.34	1.21e-4	2.43
	256×256	3.75e-4	2.00	1.40e-5	2.05	2.94e-5	2.04
	16×16	7.53e-3	–	2.99e-4	–	8.27e-4	–
	32×32	7.70e-4	3.39	2.17e-5	3.90	6.28e-5	3.83
$k = 2$	64×64	9.84e-5	3.00	2.47e-6	3.16	8.12e-6	2.98
	128×128	1.25e-5	3.04	3.43e-7	2.91	1.03e-6	3.05
	256×256	1.55e-6	3.01	3.97e-8	3.05	1.26e-7	3.03
	16×16	4.22e-4	–	7.15e-6	–	2.76e-5	–
	32×32	2.76e-5	4.52	7.33e-7	3.77	2.29e-6	4.12
$k = 3$	64×64	1.89e-6	3.91	5.32e-8	3.83	1.68e-7	3.82
	128×128	1.22e-7	4.04	3.74e-9	3.91	1.17e-8	3.92
	256×256	7.72e-9	3.98	2.24e-10	4.06	6.78e-10	4.11

Table 4.3

Errors and convergence rates for the linear variable coefficient equation in uniform meshes

	$N \times N$	$\|e\|_0$	rate	e_c	rate	e_d	rate
$k = 0$	16×16	2.00e-0	–	2.98e-1	–	4.57e-1	–
	32×32	1.23e-0	0.70	1.88e-1	0.66	2.82e-1	0.70
	64×64	7.21e-1	0.78	1.11e-1	0.76	1.64e-1	0.78
	128×128	4.03e-1	0.84	6.26e-2	0.83	9.18e-2	0.84
	256×256	2.17e-1	0.89	3.39e-2	0.89	4.94e-2	0.89
$k = 1$	16×16	1.00e-1	–	9.06e-3	–	1.47e-2	–
	32×32	2.22e-2	2.18	1.42e-3	2.67	2.20e-3	2.74
	64×64	5.24e-3	2.08	1.90e-4	2.90	2.89e-4	2.92
	128×128	1.29e-3	2.02	2.41e-5	2.98	3.65e-5	2.99
	256×256	3.21e-4	2.01	3.01e-6	3.00	4.57e-6	3.00
$k = 2$	16×16	4.64e-3	–	4.32e-5	–	1.83e-4	–
	32×32	5.74e-4	3.02	1.39e-6	4.96	1.12e-5	4.03
	64×64	7.15e-5	3.00	4.32e-8	5.00	6.95e-7	4.01
	128×128	8.92e-6	3.00	1.34e-9	5.01	4.34e-8	4.00
	256×256	1.11e-6	3.00	4.18e-11	5.01	2.71e-9	4.00
$k = 3$	16×16	2.84e-4	–	3.84e-6	–	9.16e-6	–
	32×32	1.73e-5	4.04	1.20e-8	8.32	2.55e-7	5.17
	64×64	1.08e-6	4.01	8.67e-11	7.11	7.32e-9	5.12
	128×128	6.72e-8	4.00	1.26e-12	6.10	2.24e-10	5.03
	256×256	4.20e-9	4.00	3.25e-14	5.28	6.90e-12	5.02

Table 4.4

Errors and convergence rates for the linear variable coefficient equation in nonuniform meshes

	$N \times N$	$\|e\|_0$	rate	e_c	rate	e_d	rate
$k = 0$	16×16	2.05e-0	–	3.06e-1	–	4.66e-1	–
	32×32	1.28e-0	0.68	1.95e-1	0.65	2.91e-1	0.69
	64×64	7.56e-1	0.85	1.16e-1	0.84	1.70e-1	0.86
	128×128	4.19e-1	0.88	6.50e-2	0.86	9.50e-2	0.87
	256×256	2.26e-1	0.88	3.51e-2	0.88	5.10e-2	0.88
$k = 1$	16×16	1.16e-1	–	7.06e-3	–	1.69e-2	–
	32×32	2.55e-2	2.22	1.88e-3	1.94	3.62e-3	2.26
	64×64	6.21e-3	2.27	5.17e-4	2.07	8.95e-4	2.24
	128×128	1.48e-3	1.93	1.18e-4	1.99	2.10e-4	1.96
	256×256	3.52e-4	2.07	2.77e-5	2.09	5.02e-5	2.06
$k = 2$	16×16	6.89e-3	–	2.37e-4	–	7.03e-4	–
	32×32	7.44e-4	3.18	1.54e-5	3.91	6.16e-5	3.48
	64×64	9.72e-5	3.15	2.61e-6	2.75	8.60e-6	3.05
	128×128	1.24e-5	2.97	3.47e-7	2.91	1.14e-6	2.92
	256×256	1.66e-6	2.91	4.20e-8	3.05	1.56e-7	2.87
$k = 3$	16×16	4.70e-4	–	1.85e-5	–	4.65e-5	–
	32×32	2.99e-5	4.20	1.31e-6	4.04	2.99e-6	4.19
	64×64	1.94e-6	3.87	8.48e-8	3.87	1.94e-7	3.87
	128×128	1.17e-7	4.09	5.06e-9	4.11	1.17e-8	4.09
	256×256	7.19e-9	4.03	3.01e-10	4.07	6.69e-10	4.13

TABLE 4.5
Errors and convergence rates for the nonlinear equation in uniform meshes

	$N \times N$	$\|e\|_0$	rate	e_c	rate	e_d	rate
	16×16	1.77e-0	–	2.57e-1	–	4.02e-1	–
	32×32	1.09e-0	0.70	1.63e-1	0.66	2.48e-1	0.70
$k = 0$	64×64	6.32e-1	0.78	9.66e-2	0.76	1.45e-1	0.78
	128×128	3.52e-1	0.85	5.42e-2	0.83	8.03e-2	0.85
	256×256	1.90e-1	0.89	2.93e-2	0.89	4.32e-2	0.89
	16×16	1.15e-1	–	1.10e-2	–	2.05e-2	–
	32×32	2.41e-2	2.25	1.84e-3	2.58	3.31e-3	2.63
$k = 1$	64×64	5.42e-3	2.15	2.61e-4	2.81	4.64e-4	2.83
	128×128	1.30e-3	2.06	3.53e-5	2.89	6.17e-5	2.91
	256×256	3.22e-4	2.02	4.72e-6	2.90	8.08e-6	2.93
	16×16	4.88e-3	–	8.43e-5	–	3.15e-4	–
	32×32	5.91e-4	3.05	4.11e-6	4.36	1.80e-5	4.13
$k = 2$	64×64	7.31e-5	3.02	1.81e-7	4.51	9.29e-7	4.27
	128×128	9.11e-6	3.00	7.05e-9	4.68	5.23e-8	4.15
	256×256	1.14e-6	3.00	2.52e-10	4.80	3.13e-9	4.06
	16×16	2.87e-4	–	1.41e-6	–	1.09e-5	–
	32×32	1.77e-5	4.02	2.83e-8	5.64	3.20e-7	5.09
$k = 3$	64×64	1.10e-6	4.01	2.46e-10	6.85	8.77e-9	5.19
	128×128	6.88e-8	4.00	2.70e-12	6.51	2.81e-10	4.96
	256×256	4.30e-9	4.00	4.76e-14	5.82	8.79e-12	5.00

Again we choose a special f such that the exact solution to this equation is $u(x, y, t) = \sin(x + y - 2t)$.

Tables 4.1-4.2 present errors and the corresponding convergence rates for $k = 0, 1, 2, 3$ in uniform meshes and nonuniform meshes, respectively. Similar to the linear equations, the L^2 error is convergent with $(k + 1)$-th order, and the error for the downwind edge average e_d is superconvergent, with an order of $k + 2$ in uniform meshes. All these results are consistent with our theoretical findings given in Theorem 3.1. While the convergence rate for the cell average error e_c is slightly better than our theoretical result in Theorem 3.1, which is $2k + 1$ for $k = 1, 2$ and $2k$ for $k = 3$. As for the nonuniform meshes, the expected $(k+1)$-th order of accuracy are observed and superconvergence results no longer exist.

5. Concluding remarks. We have studied the error estimates and superconvergence behavior of the DG solution on uniform Cartesian meshes for linear and nonlinear 2D hyperbolic equations using upwind fluxes and $\mathbb{P}^k$ elements. Optimal error estimates in the L^2 norm and superconvergence for the cell average and downwind edge average with an order of $\mathcal{O}(h^{k+2})$ are derived, under the condition that the wind direction does not change sign on the whole domain.

Comparing with the counterpart $\mathbb{Q}^k$ element, the degree of freedom for the $\mathbb{P}^k$ element is almost cut in half. However, a interesting and surprising result is that the superconvergence property for $\mathbb{P}^k$ element still remains on the uniform meshes. Especially, evidences from our numerical experiments show that the highest superconvergence rate for the cell average error can reach as high as the counterpart $\mathbb{Q}^k$ element, which turns out to be $\mathcal{O}(h^{2k+1})$ in some special cases (e.g., $k = 1, 2$). It is surprising that the loss of degree of freedom is not at the expense of the accuracy for

Table 4.6
Errors and convergence rates for the nonlinear equation in uniform meshes

	$N\times N$	$\|e\|_0$	rate	e_c	rate	e_d	rate
$k=0$	16×16	1.82e-0	–	2.64e-1	–	4.10e-1	–
	32×32	1.13e-0	0.78	1.69e-1	0.73	2.55e-1	0.78
	64×64	6.53e-1	0.76	9.95e-2	0.74	1.48e-1	0.76
	128×128	3.65e-1	0.86	5.62e-2	0.84	8.29e-2	0.86
	256×256	1.98e-1	0.91	3.06e-2	0.90	4.48e-2	0.91
$k=1$	16×16	1.41e-1	–	1.38e-2	–	2.48e-2	–
	32×32	2.80e-2	2.43	2.86e-3	2.36	4.94e-3	2.43
	64×64	6.52e-3	2.09	5.97e-4	2.25	9.37e-4	2.39
	128×128	1.55e-3	2.15	1.39e-4	2.18	2.27e-4	2.11
	256×256	3.98e-4	1.96	3.94e-5	1.82	5.93e-5	1.94
$k=2$	16×16	6.75e-3	–	2.30e-4	–	7.01e-4	–
	32×32	7.83e-4	3.23	1.87e-5	3.76	6.87e-5	3.48
	64×64	9.76e-5	2.86	2.18e-6	2.95	7.51e-6	3.04
	128×128	1.26e-5	3.16	3.33e-7	2.90	1.08e-6	2.99
	256×256	1.62e-6	2.96	3.83e-8	3.12	1.43e-7	2.92
$k=3$	16×16	4.52e-4	–	1.23e-5	–	3.44e-5	–
	32×32	2.70e-5	3.92	7.72e-7	3.86	1.91e-6	4.02
	64×64	1.88e-6	4.30	7.27e-8	3.81	1.50e-7	4.11
	128×128	1.20e-7	4.16	4.17e-9	4.33	1.02e-8	4.08
	256×256	7.74e-9	3.96	2.62e-10	3.99	6.72e-10	3.92

superconvergence on the uniform meshes. However, different to the $\mathbb{Q}^k$ element, where superconvergence result are the same in both the uniform and nonuniform meshes, it seems that the superconvergence for $\mathbb{P}^k$ element is dependent upon the mesh and the superconvergence phenomenon disappears when the mesh is nonuniform.

Extension of this work to nonuniform meshes and to triangulations is interesting and challenging, and constitutes our future work.

6. Appendix.

This section is dedicated to the proof of Lemma 2.4.

Proof. We first prove (2.21)-(2.22) for $l=1$. In each $\tau_{i,j}$, we suppose

$$\zeta_1|_{\tau_{i,j}} = \sum_{1\le p+q\le k} c_{p,q}^{i,j} L_{i,p}(x)L_{j,q}(y). \tag{6.1}$$

By choosing $v = L_{i,p}(x)L_{j,q(}y), 1\le p+q\le k$ in (2.18), we easily obtain a linear system for the coefficients $c_{p,q}^{i,j}$ and thus,

$$h|c_{p,q}^{i,j}| \lesssim \int_{\tau_j^y} |\zeta_0^-|_{i+\frac12,y}| + |\zeta_0^-|_{i-\frac12,y}|dy + \int_{\tau_i^x} |\zeta_0^-|_{x,j+\frac12}| + |\zeta_0^-|_{x,j-\frac12}|dx,$$

which yields, together with the trace inequality,

$$\|\zeta_1\|_{0,\tau}^2 \le |\tau| \sum_{1\le p+q\le k} |c_{p,q}^{i,j}|^2 \le \|\zeta_0\|_{0,\tau}^2 + h^2|\zeta_0|_{1,\tau}^2.$$

Recalling the definition of ζ_0 and using the estimates for ζ_0, we have

$$\|\zeta_1\|_0 \lesssim h^{k+1}\|u\|_{k+1}.$$

The (2.21) holds ture for $l = 1$.

To prove (2.22), we denoting

$$\tau_1 = \tau_{i,j},\ \tau_2 = \tau_{i-1,j}, \quad w_1 = \zeta_1|_{\tau_1}, \quad w_2 = \zeta_1|_{\tau_2}.$$

Then

$$w_1 = \sum_{1 \le p+q \le k} c_{p,q}^{i,j} L_{i,p}(x) L_{j,q}(y), \quad w_2 = \sum_{1 \le p+q \le k} c_{p,q}^{i-1,j} L_{i-1,p}(x) L_{j,q}(y).$$

Recalling the definition of ζ_1 in (2.18), we have

$$b_{\tau_1}(w_1, v_1) - b_{\tau_2}(w_2, v_2) = H_{\tau_2}(\zeta_0, v_1) - H_{\tau_1}(\zeta_0, v_2), \quad \forall v_i \in \mathbb{P}_k(\tau_i).$$

By choosing $v_1 = L_{i,r}(x)L_{j,r'}(y), v_2 = L_{i-1,r}(x)L_{j,r'}(y)$ and using the fact that the mesh is uniform, we obtain a linear system for the coefficients $c_{p,q}^{i,j} - c_{p,q}^{i-1,j}$ and thus

$$\begin{aligned} h|c_{p,q}^{i,j} - c_{p,q}^{i-1,j}| \lesssim & \max_{0 \le r' \le k} \sum_{i'=i-1,i} |\int_{\tau_j^y} D_x(\alpha\zeta_0)_{i'}^- L_{j,r'} dy| \\ & + \max_{0 \le r \le k} \sum_{j'=j,j-1} |\int_{\tau_i^x} \beta\zeta_0^-|_{x,j'+\frac{1}{2}} L_{i,r}(x) dx - \int_{\tau_{i-1}^x} \beta\zeta_0^-|_{x,j'+\frac{1}{2}} L_{i-1,r}(x) dx|. \end{aligned}$$

In light of (2.12)-(2.15), we have

$$\sum_{\tau_{i,j}} |D_1 c_{p,q}^{i,j}|^2 \lesssim h^{2k+2} \|u\|_{k+2}^2. \tag{6.2}$$

Following the same argument, we can prove that

$$\sum_{\tau_{i,j}} |D_2 c_{p,q}^{i,j}|^2 \lesssim h^{2k+2} \|u\|_{k+2}^2, \ \sum_{\tau_{i,j}} |D_1 D_2 c_{p,q}^{i,j}|^2 + |D_1^2 c_{p,q}^{i,j}|^2 \lesssim h^{2k+4} \|u\|_{k+3}^2. \tag{6.3}$$

Noticing that

$$\begin{aligned} \sum_{\tau_{i,j}} \int_{\tau_i^x} |D_y(\zeta_1)_j^-|^2 dx + \int_{\tau_j^y} |D_x(\zeta_1)_i^-|^2 dy \lesssim & \sum_{\tau_{i,j}} \sum_{1 \le p+q \le k} h|D_1 c_{p,q}^{i,j}|^2 + h|D_2 c_{p,q}^{i,j}|^2 \\ \lesssim & \ h^{2k+3} \|u\|_{k+2}^2. \end{aligned}$$

Noticing that

$$D_x(\alpha\zeta_1)_i^- = D_x\alpha_i^- \zeta_1^-|_{i+\frac{1}{2},y} + \alpha^-|_{i-\frac{1}{2},y} D_x(\zeta_1)_i^-, \ |D_x\alpha_i^-| \lesssim h,$$

we have

$$\sum_{\tau_{i,j}} \int_{\tau_j^y} |D_x(\alpha\zeta_1)_i^-|^2 dy \lesssim \sum_{\tau_{i,j}} \int_{\tau_j^y} |D_x(\zeta_1)_i^-|^2 + h^2|\zeta_1^-|_{i+\frac{1}{2},y}|^2 dy \lesssim h^{2k+3} \|u\|_{k+2}^2.$$

Similarly, we can prove

$$\sum_{\tau_{i,j}} \int_{\tau_i^x} |D_x(\beta\zeta_1)_j^-|^2 dx \lesssim h^{2k+3} \|u\|_{k+2}^2.$$

Then (2.22) holds true for $l=1$.

By taking $v = L_{i,1}(x)$ in (2.18), we get

$$2\bar{\alpha}\int_{\tau_j^y}\zeta_1^-|_{i+\frac{1}{2},y}dy = \int_{\tau_j^y}(\alpha\zeta_0^-|_{i+\frac{1}{2},y} + \alpha\zeta_0^-|_{i-\frac{1}{2},y})dy + \int_{\tau_i^x}(\beta\zeta_0^-|_{x,j+\frac{1}{2}} - \beta\zeta_0^-|_{x,j-\frac{1}{2}})L_{i,1}(x)dx,$$

and thus

$$\begin{aligned}2\int_{\tau_j^y}D_x(\bar{\alpha}\zeta_1)_i^-\,dy &= \int_{\tau_j^y}(\alpha\zeta_0^-|_{i+\frac{1}{2},y} - \alpha\zeta_0^-|_{i-\frac{3}{2},y})dy\\ &\quad+\int_{\tau_i^x}D_y(\beta\zeta_0)_j^-\,L_{i,1}dx - \int_{\tau_{i-1}^x}D_y(\beta\zeta_0)_j^-\,L_{i-1,1}dx,\end{aligned}$$

which yields, together with (2.12)-(2.15)

$$\sum_{\tau_{i,j}}|\int_{\tau_j^y}D_x(\bar{\alpha}\zeta_1)_i^-\,dy|^2 \lesssim h^{2(k+3)}\|u\|_{k+3}^2, \sum_{\tau_{i,j}}(\int_{\tau_j^y}\zeta_1^-|_{i+\frac{1}{2},y}dy)^2 \lesssim h^{2(k+2)}\|u\|_{k+2}^2. \tag{6.4}$$

Noticing that

$$\begin{aligned}&\int_{\tau_j^y}D_x((\alpha-\bar{\alpha})\zeta_1)_i^-\,dy = \int_{\tau_j^y}D_x(\alpha-\bar{\alpha})_i^-\,\zeta_1^-|_{i+\frac{1}{2},y} + (\alpha-\bar{\alpha})|_{i-\frac{1}{2},y}^-D_x(\zeta_1)_i^-\,dy\\ &|(\alpha-\bar{\alpha})(x_{i-\frac{1}{2}}^-,y)| \lesssim h, \quad |D_x(\alpha-\bar{\alpha})_i^-| \lesssim h^2, \quad \forall y\in\tau_j^y,\end{aligned}$$

we have

$$\begin{aligned}|\int_{\tau_j^y}D_x(\alpha\zeta_1)_i^-\,dy|^2 &\lesssim |\int_{\tau_j^y}D_x((\alpha-\bar{\alpha})\zeta_1)_i^-\,dy|^2 + |\int_{\tau_j^y}D_x(\bar{\alpha}\zeta_1)_i^-\,dy|^2\\ &\lesssim h^5\int_{\tau_j^y}|\zeta_1^-|_{i+\frac{1}{2},y}|^2dy + h^3\int_{\tau_j^y}|D_x(\zeta_1)_i^-|^2dy + |\int_{\tau_j^y}D_x(\bar{\alpha}\zeta_1)_i^-\,dy|^2\end{aligned}$$

Summing up all $\tau_{i,j}$ and using (2.22) for $l=1$ and (6.4), we have

$$\begin{aligned}\sum_{\tau_{i,j}}|\int_{\tau_j^y}D_x(\alpha\zeta_1)_i^-\,dy|^2 &\lesssim h^5\sum_{\tau_{i,j}}\int_{\tau_j^y}|\zeta_1^-|_{i+\frac{1}{2},y}|^2dy + h^{2k+6}\|u\|_{k+3}^2\\ &\lesssim h^4\|\zeta_1\|_0^2 + h^{2k+6}\|u\|_{k+3}^2 \lesssim h^{2k+6}\|u\|_{k+3}^2.\end{aligned}$$

Here in the second step, we have used the inverse inequality. Following the same argument, there hold

$$\sum_{\tau_{i,j}}|\int_{\tau_i^x}D_y(\beta\zeta_1)_j^-\,dx|^2 \lesssim h^{2(k+3)}\|u\|_{k+3}^2, \ \sum_{\tau_{i,j}}(\int_{\tau_i^x}\zeta_1^-|_{x,j+\frac{1}{2}}dx)^2 \lesssim h^{2(k+2)}\|u\|_{k+2}^2.$$

Combining the last two inequalities and the second inequality of (6.4) together yields the desired result (2.23) for $l=1$.

Following the same argument as what we did for $l=1$, we obtain

$$\|\zeta_2\|_0 \lesssim h(\|\partial_t\zeta_1\|_0 + \|\zeta_1\|_0) + \big(\sum_{\tau_{i,j}}h\int_{\tau_j^y}|D_x(\alpha\zeta_1)_i^-|^2dy + h\int_{\tau_i^x}|D_y(\beta\zeta_1)_j^-|^2dx\big)^{\frac{1}{2}}.$$

Using the estimates (2.21) and (2.22) for $l = 1$, we get

$$\|\zeta_2\|_0 \lesssim h^{k+2}(\|u\|_{k+2} + \|u_t\|_{k+1}) \lesssim h^{k+2}\|u\|_{k+2}.$$

Here in the last step, we have used the fact that $u_t = -(\alpha u)_x - (\beta u)_y$. Consequently, (2.21) holds true for $l = 2$. Similarly, we take $v = L_{i,1}$ in (2.19) and then obtain

$$\begin{aligned}2\int_{\tau_j^y} D_x(\bar{\alpha}\zeta_2)_i^- dy &= (\partial_t\zeta_1, L_{i,1})_{\tau_1} - (\partial_t\zeta_1, L_{i-1,1})_{\tau_2} - \frac{2}{h}(\alpha - \bar{\alpha}, \zeta_1)_{\tau_1} + \frac{2}{h}(\alpha - \bar{\alpha}, \zeta_1)_{\tau_2} \\ &+ \int_{\tau_j^y} D_x^2(\alpha\zeta_1)_i^- dy + \int_{\tau_i^x} D_y(\beta\zeta_1)_j^- L_{i,1}(x)dx - \int_{\tau_{i-1}^x} D_y(\beta\zeta_1)_j^- L_{i-1,1}(x)dx \\ &+ 2\int_{\tau_j^y} (\alpha - \bar{\alpha})\zeta_1^-|_{i+\frac{1}{2},y} - (\alpha - \bar{\alpha})\zeta_1^-|_{i-\frac{1}{2},y})dy := rhs.\end{aligned}$$

In light of (6.1), we have, from a direct calculation,

$$|rhs| \lesssim \sum_{1\le p+q\le k} h^2(|\partial_t D_1 c_{p,q}^{i,j}| + |D_1 c_{p,q}^{i,j}|) + h(|D_2 D_1 c_{p,q}^{i,j}| + |D_1^2 c_{p,q}^{i,j}|).$$

Then we conclude from (6.2)-(6.3) that

$$\sum_{\tau_{i,j}} |\int_{\tau_j^y} D_x(\bar{\alpha}\zeta_2)_i^- dy|^2 \lesssim \sum_{\tau_{i,j}} |rhs|^2 \lesssim h^{2k+6}(\|u\|_{k+3}^2 + \|u_t\|_{k+2}^2).$$

The rest proof of (2.23) and (2.22) for $l = 2$ is similar to that of $l = 1$ and thus we omit it here. □

REFERENCES

[1] S. Adjerid and M. Baccouch, The discontinuous Galerkin method for two-dimensional hyperbolic problems Part I: Superconvergence error analysis, *J. Sci. Comput.*, 33 (2007), pp. 75-113.

[2] S. Adjerid and T. C. Massey, Superconvergence of discontinuous Galerkin solutions for a nonlinear scalar hyperbolic problem, *Comput. Methods Appl. Mech. Engrg.*, 195 (2006), pp. 3331-3346.

[3] S. Adjerid and T. Weinhart, Discontinuous Galerkin error estimation for linear symmetrizable hyperbolic systems, *Math. Comp.*, 80 (2011), pp. 1335-1367.

[4] W. Cao, C.-W. Shu, Yang Yang and Z. Zhang, Superconvergence of discontinuous Galerkin method for 2-D hyperbolic equations, *SIAM J. Numer. Anal.,* 53 (2015), pp. 1651-1671.

[5] W. Cao, C.-W. Shu, Y. Yang and Z. Zhang, Superconvergence of discontinuous Galerkin method for scalar nonlinear hyperbolic equations, *SIAM J. Numer. Anal.*, 56 (2018), pp. 732-765.

[6] W. Cao, Z. Zhang and Q. Zou, Superconvergence of Discontinuous Galerkin method for linear hyperbolic equations, *SIAM. J. Numer. Anal.*, 52 (2014), pp. 2555-2573.

[7] Y. Cheng and C.-W. Shu, Superconvergence and time evolution of discontinuous Galerkin finite element solutions, *J. Comput. Phys.*, 227 (2008), pp. 9612–9627.

[8] B. Cockburn, B. Dong and J. Guzmán, Optimal convergence of the original DG method for the transport-reaction equation on special meshes, *SIAM J. Numer. Anal.*, 46 (2008), pp. 1250-1265.

[9] B. Cockburn, S. Hou and C. Shu, The Runge-Kutta local projection discontinuous Galerkin finite element method for coservation laws, IV : The Multidimensional case, *Math. Comp.*, 54 (1990), pp. 545-581.

[10] B. Cockburn, G. Karniadakis and C.-W. Shu, *The Development of Discontinuous Galerkin Methods*, Springer Berlin Heidelberg, 2000.

[11] B. Cockburn, S. Lin and C. Shu, TVB Runge-Kutta local projection discontinuous Galerkin finite element method for coservation laws, III : One dimensioal systems, *J. Comput. Phys.*, 84 (1989), pp. 90-113.

[12] B. Cockburn and C. Shu, TVB Runge-Kutta local projection discontinuous Galerkin finite element method for coservation laws, II : Genaral framework, *Math. Comp.*, 52 (1989), pp. 411-435.

[13] B. Cockburn and C. Shu, The Runge-Kutta local projection P1-discontinuous Galerkin finite element method for scalar coservation laws, RAIRO Model, *Math. Anal. Numer.*, 25 (1991), pp. 337-361.

[14] B. Cockburn and C. Shu, TVB Runge-Kutta local projection discontinuous Galerkin finite element method for coservation laws, V : Multidimensional systems, *J. Comput. Phys.*, 141 (1998), pp. 199-224.

[15] R.S. Falk and G.R. Richter, Explicit finite element methods for symmetric hyperbolic equations, *SIAM J. Numer. Anal.*, 36 (1999), pp. 935-952.

[16] W. Guo, X. Zhong and J. Qiu, Superconvergence of discontinuous Galerkin and local discontinuous Galerkin methods: eigen-structure analysis based on Fourier approach, *J. Comput. Phys.*, 235 (2013), pp. 458-485.

[17] P. Houston, C. Schwab, and E. Süli, stabilized *hp* finite element methods for first-order hyperbolic problems, *SIAM J. Numer. Anal.*, 37 (2000), pp. 1618-1643.

[18] C. Johnson and J. Pitkaranta, An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation, *Math. Comp.*, 46 (1986), pp. 1-26.

[19] P. LeSaint and P.A. Raviart, *On a finite element method for solving the neutron transport equation,* In C de Boor, editor, Mathematical aspects of finite elements in partial differential equations, pages 89-145, Academic Press, 1974.

[20] Q. Lin, N. Yan, and A. Zhou, An optimal error estimate of the discontinuous Galerkin method, *Journal of Engineering Mathematics*, 13 (1996), pp. 101-105.

[21] Y. Liu, C.-W. Shu, and M. Zhang, Optimal error estimates of the semidiscrete discontinuous Galerkin methods for two dimensional hyperbolic equations on Cartesian meshes using P^k elements, *ESAIM: M2AN*, 54 (2020), pp. 705–726.

[22] X. Meng, C.-W. Shu and B. Wu, Optimal error estimates for discontinuous Galerkin methods based on upwind-biased fluxes for linear hyperbolic equations. Math. Comput. 85 (2016),pp. 1225-1261.

[23] T. Peterson, A note on the convergence of the discontinuous Galerkin method for a scalar hyperbolic equation, *SIAM J. Numer. Anal.*, 28 (1991), pp. 133-140.

[24] W. H. Reed and T. R. Hill, *Triangular Mesh for Neutron Transport Equation,* Los Alamos Scientific Laboratory Report LA-UR-73-479, Los Alamos, NM, 1973.

[25] G.R. Richter, An optimal-order error estimate for the discontinuous Galerkin method, *Math. Comp.*, 50 (1988), pp. 75-88.

[26] C.-W. Shu, *Discontinuous Galerkin method for time-dependent problems: Survey and recent developments*, In: X. Feng O. Karakashian, Y. Xing, Recent Developments in Discontinuous Galerkin Finite Element Methods for Partial Differential Equations. Lecture Notes in The IMA Volumes in Mathematics and its Applications, vol. 157. Cham: Springer, 2014.

[27] Z. Xie and Z. Zhang, Uniform superconvergence analysis of the discontinuous Galerkin method for a singularly perturbed problem in 1-D, *Math. Comp.*, 79 (2010), pp. 35-45.

[28] Y. Yang and C.-W. Shu, Analysis of optimal superconvergence of discontinuous Galerkin method for linear hyperbolic equations, *SIAM J. Numer. Anal.*, 50 (2012), pp. 3110-3133.

[29] Q. Zhang and C.-W. Shu, Error estimates to smooth solutions of Runge-Kutta discontinuous Galerkin methods for scalar conservation laws, *SIAM J. Numer. Anal.*, 42 (2004), pp. 641-666.

[30] A. Zhou and Q. Lin, Optimal and sup erconvergence estimates of the finite element method for a scalar hyperbolic equation, *Acta. Math. Sci.*, 14 (1994), pp. 90-94.