COMMUNICATIONS IN COMPUTATIONAL PHYSICS
Vol. **6**, No. 1, pp. 203-230

Commun. Comput. Phys.
July 2009

A Family of Characteristic Discontinuous Galerkin Methods for Transient Advection-Diffusion Equations and Their Optimal-Order L^2 Error Estimates

Kaixin Wang[1], Hong Wang[2,*], Mohamed Al-Lawatia[3] and Hongxing Rui[1]

[1] *School of Mathematics and System Sciences, Shandong University, Jinan, Shandong 250100, China.*
[2] *Department of Mathematics, University of South Carolina, Columbia, South Carolina 29208, USA.*
[3] *Department of Mathematics and Statistics, Sultan Qaboos University, P.O. Box 36, Al-Khod 123, Sultanate of Oman.*

Received 20 November 2007; Accepted (in revised version) 25 October 2008

Communicated by Chi-Wang Shu

Available online 24 November 2008

Abstract. We develop a family of characteristic discontinuous Galerkin methods for transient advection-diffusion equations, including the characteristic NIPG, OBB, IIPG, and SIPG schemes. The derived schemes possess combined advantages of Eulerian-Lagrangian methods and discontinuous Galerkin methods. An optimal-order error estimate in the L^2 norm and a superconvergence estimate in a weighted energy norm are proved for the characteristic NIPG, IIPG, and SIPG scheme. Numerical experiments are presented to confirm the optimal-order spatial and temporal convergence rates of these schemes as proved in the theorems and to show that these schemes compare favorably to the standard NIPG, OBB, IIPG, and SIPG schemes in the context of advection-diffusion equations.

AMS subject classifications: 35R32, 37N30, 65M12, 76N15

Key words: Advection-diffusion equation, characteristic method, discontinuous Galerkin method, numerical analysis, optimal-order L^2 error estimate, superconvergence estimate.

1 Introduction

Transient advection-diffusion equations admit solutions with moving steep fronts and complicated structures. Classic numerical methods tend to generate numerical solutions

*Corresponding author. *Email addresses:* kx.wang@mail.sdu.edu.cn (K. Wang), hwang@math.sc.edu (H. Wang), allawati@squ.edu.om (M. Al-Lawatia), hxrui@sdu.edu.cn (H. Rui)

with spurious oscillations, excessive numerical diffusion, or a combination of both. Since their introduction in 1970s [3, 18, 30], discontinuous Galerkin methods have become a topic of extensive research on numerically solving differential equations. These methods use piecewise polynomial trial and test functions which may be discontinuous across cell boundaries in their weak formulations and are locally mass conservative. They are inherently adaptive to local high order approximations and can capture moving steep fronts and shock discontinuities in the solution to advection-diffusion equations via the use of discontinuous approximating spaces. They are well suited for handling unstructured meshes and nonmatching grids for multidimensional problems.

Optimal-order convergence in the energy norm has been proved for a variety of primal discontinuous Galerkin methods for elliptic and parabolic differential equations [7, 17,20,22,23]. Optimal L^2 convergence has been established [22,23] for the symmetric interior penalty Galerkin (SIPG) method [30]. However, numerical experiments [7,23] reveal that the nonsymmetric discontinuous Galerkin methods, including the Oden-Babuška-Baumann (OBB) formulation [7,17], the nonsymmetric interior penalty Galerkin (NIPG) method [19], and the incomplete interior penalty Galerkin (IIPG) method [12,22,23], are generally not optimal in the L^2 norm. An optimal-order error estimate in the L^2 norm has been proved for the NIPG and IIPG methods with odd degree polynomials for one-dimensional elliptic problems with a uniform partition [16].

In this paper we develop a family of characteristic discontinuous Galerkin methods for one-dimensional transient advection-diffusion equations by using an Eulerian-Lagrangian approach within a primal discontinuous Galerkin framework [4,9] (instead of the dual formulation [10,11]). These include the characteristic SIPG method, the characteristic NIPG method, the characteristic IIPG method, and the characteristic OBB method. The developed methods retain the numerical advantages of the discontinuous Galerkin methods. Further, they stabilize the numerical approximations and generate accurate numerical solutions, even if large time steps and coarse spatial grids are used. In this paper we prove an optimal-order L^2 error estimate and a superconvergence estimate for the characteristic NIPG, SIPG, and IIPG methods. Numerical results are presented to show the convergence behavior of the proposed schemes.

The rest of the paper is organized as follows: In Section 2 we derive a reference equation satisfied by the true solution to problem (2.1). In Section 3 we develop a family of characteristic discontinuous Galerkin methods. In Section 4 we present the preliminaries used in the error analysis. Section 5 contains the main error estimate of this article. In Section 6 we present preliminary example runs to show the performance of the scheme. Section 7 contains summary and conclusions. Finally, we present the proof of auxiliary lemmas in the appendix.

2 A cell-based characteristic reference equation

We consider the initial-boundary value problem

$$\begin{aligned} &u_t+(Vu)_x-(Du_x)_x=f(x,t), \quad x\in(a,b), \quad t\in(0,T],\\ &Vu(a,t)-Du_x(a,t)=g(t), \quad -Du_x(b,t)=h(t), \quad t\in(0,T],\\ &u(x,0)=u_0(x), \quad x\in(a,b). \end{aligned} \tag{2.1}$$

Here $V(x,t)$ is the velocity field; $f(x,t)$, $u_0(x)$, $g(t)$, and $h(t)$ are the prescribed source and sink term, the initial data, and the inflow and outflow boundary data, respectively. The diffusion coefficient D satisfies

$$0<D_{\min}\le D(x,t)\le D_{\max}<+\infty.$$

To show the treatment of general inflow and outflow flux boundary conditions, we assume that $V(a,t)$ and $V(b,t)$ are positive so that $x=a$ and $x=b$ are the inflow and outflow boundaries, respectively.

Define a quasiuniform space-time partition $a=x_0<x_1<x_2<\cdots<x_I=b$ and $0=t_0<t_1<t_2<\cdots<t_N=T$ with $h_i=x_i-x_{i-1}$, $h=\max_{1\le i\le I}h_i$, $\Delta t_n=t_n-t_{n-1}$ and $\Delta t=\max_{1\le n\le N}\Delta t_n$ being the sizes of spatial mesh and time steps. For simplicity we drop the subscripts when it is clear from the context. Note that for a general velocity field, the initial-value problem $dx/dt=V(x,t)$ with $x(\bar{t})=\bar{x}$ cannot be solved analytically. Thus, we use an Euler quadrature to define the approximate characteristic that passes through $\bar{x}$ at $\bar{t}$ by

$$r(s;\bar{x},\bar{t})=\bar{x}-V(\bar{x},\bar{t})(\bar{t}-s). \tag{2.2}$$

From now on, we shall assume that the curve $r(s;\bar{x},\bar{t})$ is defined by Eq. (2.2).

2.1 A reference equation on an interior control volume

In this subsection we derive a reference equation for the true solution $u(x,t)$ to the problem (2.1), on the following space-time control volume

$$\Omega_i^n=\{(x,t):r(t;x_{i-1},t_n)<x<r(t;x_i,t_n),t_{n-1}<t\le t_n\},$$

which is aligned with the characteristic curve defined by (2.2). Note that the lateral boundary of Ω_i^n is determined by the characteristic curves $r(t;x_{i-1},t_n)$ and $r(t;x_i,t_n)$ from $t=t_n$ to $t=t_{n-1}$. We let x^* be the foot of the characteristics at time step t_{n-1} with head x at time t_n and $\tilde{x}$ be the head of the characteristics at time step t_n with the foot x at time t_{n-1}, i.e., $x^*=r(t_{n-1};x,t_n)$ and $x=r(t_{n-1};\tilde{x},t_n)$.

To derive a cell-based reference equation, we choose the space-time test functions $w(x,t)$ to be smooth within Ω_i^n and to vanish outside Ω_i^n in the weak formulation. Due to the use of characteristics, the test functions $w(x,t)$ are not necessarily continuous across the time steps t_n and t_{n-1}. In the formulation, we require the test functions to be left-continuous with respect to time. Namely, the test functions $w(x,t)$ satisfy

$$w(x,t_n)=\lim_{t\to t_n-0}w(x,t) \quad \text{but} \quad w(x,t_{n-1})\neq\lim_{t\to t_{n-1}+0}w(x,t)$$

in general. We use the notation $w(x,t_{n-1}^+)=\lim_{t\to t_{n-1}+0}w(x,t)$ to account for the possible discontinuity of $w(x,t)$ in time at time t_{n-1}. We multiply the governing equation in (2.1) by $w(x,t)$ and integrate the resulting equation by parts over Ω_i^n to obtain a weak formulation

$$\begin{aligned}&\int_{x_{i-1}}^{x_i}u(x,t_n)w(x,t_n)dx+\int_{\Omega_i^n}Du_xw_xdxdt\\&\quad-\int_{t_{n-1}}^{t_n}(Du_xw)(r(t;x_i,t_n),t)dt+\int_{t_{n-1}}^{t_n}(Du_xw)(r(t;x_{i-1},t_n),t)dt\\=&\int_{x_{i-1}^*}^{x_i^*}u(x,t_{n-1})w(x,t_{n-1}^+)dx+\int_{\Omega_i^n}fwdxdt-E_i^{(1)}(u,w).\end{aligned}\tag{2.3}$$

Here $E_i^{(1)}(u,w)$ is defined by

$$\begin{aligned}E_i^{(1)}(u,w)=&-\int_{t_{n-1}}^{t_n}(uw)(r(t;x_{i-1},t_n),t)(V(r(t;x_{i-1},t_n),t)-r_t(t;x_{i-1},t_n))dt\\&+\int_{t_{n-1}}^{t_n}(uw)(r(t;x_i,t_n),t)(V(r(t;x_i,t_n),t)-r_t(t;x_i,t_n))dt-\int_{\Omega_i^n}u(w_t+Vw_x)dxdt.\end{aligned}$$

We define the test functions to be constant along the characteristics [8, 24], i.e.,

$$w_t(r(t;x,t_n),t)+V(x,t_n)w_r(r(t;x,t_n),t)=0,\quad t\in[t_{n-1},t_n],\quad x\in[x_{i-1},x_i].\tag{2.4}$$

In the evaluation of source and diffusion terms we reserve x for points in $[x_{i-1},x_i]$ at time t_n representing the heads of characteristics. We use the variable $y=r(t;x,t_n)$ to represent the spatial coordinate of an arbitrary point at time $t\in(t_{n-1},t_n)$. We apply Euler quadrature to evaluate the source and sink term in Eq. (2.3) to obtain

$$\begin{aligned}\int_{\Omega_i^n}f(y,t)w(y,t)dydt=&\int_{x_{i-1}}^{x_i}\int_{t_{n-1}}^{t_n}f(r(t;x,t_n),t)w(r(t;x,t_n),t)r_x(t;x,t_n)dtdx\\=&\Delta t\int_{x_{i-1}}^{x_i}f(x,t_n)w(x,t_n)dx-E_i^{(f)}(w),\end{aligned}$$

with the local truncation error $E_i^{(f)}(w)$ for the source term being given by

$$E_i^{(f)}(w)=\int_{x_{i-1}}^{x_i}w(x,t_n)\int_{t_{n-1}}^{t_n}\left\{\int_t^{t_n}\frac{df}{d\theta}(r(\theta;x,t_n),\theta)d\theta+f(r(t;x,t_n),t)V_x(x,t_n)(t_n-t)\right\}dtdx.$$

In the derivation we have used the fact $r_x(t;x,t_n)=1-V_x(x,t_n)(t_n-t)$. We evaluate the diffusion term similarly as

$$\begin{aligned}\int_{\Omega_i^n}D(y,t)u_y(y,t)w_y(y,t)dydt=&\int_{x_{i-1}}^{x_i}\int_{t_{n-1}}^{t_n}D(r(t;x,t_n),t)u_y(r(t;x,t_n),t)w_x(x,t_n)dtdx\\=&\Delta t\int_{x_{i-1}}^{x_i}D(x,t_n)u_x(x,t_n)w_x(x,t_n)dx\\&-\int_{x_{i-1}}^{x_i}w_x(x,t_n)\int_{t_{n-1}}^{t_n}\int_t^{t_n}\frac{d}{d\theta}(Du_y)(r(\theta;x,t_n),\theta)d\theta dtdx.\end{aligned}\tag{2.5}$$

Here we have used the fact that $w(y,t)=w(x,t_n)$ for $y=r(t;x,t_n)$ and so

$$w_y(y,t)=w_y(x,t_n)=w_x(x,t_n)\frac{dx}{dy}=\frac{w_x(x,t_n)}{r_x(t;x,t_n)}.$$

Next we evaluate the two inter element diffusive flux terms in Eq.(2.3) as

$$\begin{aligned}&\int_{t_{n-1}}^{t_n}(Du_x\,w)(r(t;x_{i-1},t_n),t)dt-\int_{t_{n-1}}^{t_n}(Du_x\,w)(r(t;x_i,t_n),t)dt\\&=\Delta t(Du_x)(x_{i-1},t_n)w(x_{i-1}^+,t_n)-\Delta t(Du_x)(x_i,t_n)w(x_i^-,t_n)+E_i^{(2)}(u,w).\end{aligned}\tag{2.6}$$

Here $E_i^{(2)}(u,w)$ is defined by

$$\begin{aligned}E_i^{(2)}(u,w)=&w(x_i^-,t_n)\int_{t_{n-1}}^{t_n}\int_t^{t_n}\frac{d}{d\theta}(Du_x)(r(\theta;x_i,t_n),\theta)d\theta dt\\&-w(x_{i-1}^+,t_n)\int_{t_{n-1}}^{t_n}\int_t^{t_n}\frac{d}{d\theta}(Du_x)(r(\theta;x_{i-1},t_n),\theta)d\theta dt.\end{aligned}$$

We incorporate the preceding terms into Eq. (2.3) to obtain

$$\begin{aligned}&\int_{x_{i-1}}^{x_i}u(x,t_n)w(x,t_n)dx+\Delta t\int_{x_{i-1}}^{x_i}D(x,t)u_x(x,t)w_x(x,t_n)dx\\&\quad+\Delta tDu_x(x_{i-1},t_n)w(x_{i-1}^+,t_n)-\Delta tDu_x(x_i,t_n)w(x_i^-,t_n)\\&=\int_{x_{i-1}^*}^{x_i^*}u(x,t_{n-1})w(x,t_{n-1}^+)dx+\Delta t\int_{x_{i-1}}^{x_i}f(x,t_n)w(x,t_n)dx+E_i(u,w).\end{aligned}\tag{2.7}$$

Here the local truncation error $E_i(u,w)$ is given by

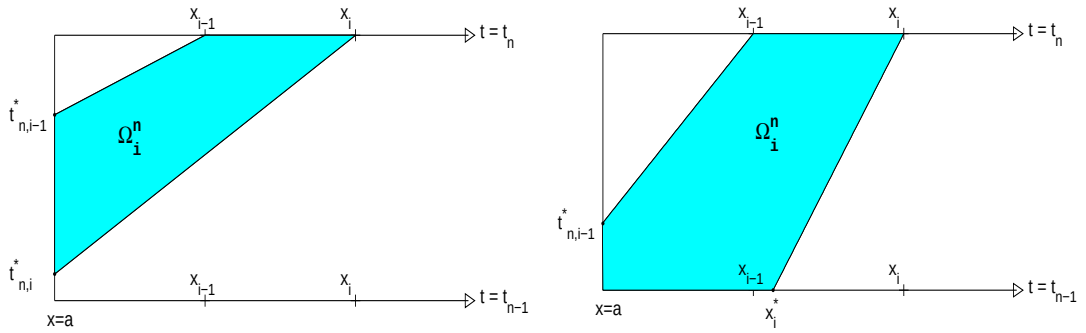
$$\begin{aligned}E_i(u,w)=&\int_{x_{i-1}}^{x_i}w_x(x,t_n)\int_{t_{n-1}}^{t_n}\int_t^{t_n}\frac{d}{d\theta}(Du_y)(r(\theta;x,t_n),\theta)d\theta dtdx\\&-E_i^{(1)}(u,w)-E_i^{(2)}(u,w)-E_i^{(f)}(w).\end{aligned}$$

2.2 Treatment of boundary condition

When the space-time prism Ω_i^n under consideration intersects the inflow boundary $x=a$ during the time period $[t_{n-1},t_n]$, we use $t_n^*(x)$ to denote the time instant if $r(\theta;x,t_n)$ intersects the boundary $x=a$, i.e.,

$$a=r(t_n^*(x);x,t_n)\quad\text{and}\quad t_n^*(x)=t_{n-1}\quad\text{otherwise.}$$

We let $t_{n,i-1}^*=t_n^*(x_{i-1})$. However, the curve $r(\theta;x_i,t_n)$ could either intersect the inflow boundary $x=a$ (see Fig. 1(a)) or falls inside the domain $[a,b]$ at time t_{n-1} (see Fig. 1(b)). Without loss of generality, we assume that the approximate characteristic curve $r(\theta;x_i,t_n)$



(a) $r(t^*(x_i);x_i,t_n)$ meets the boundary (b) x_i^* falls inside $[a,b]$

Figure 1: Illustration on control volumes Ω_i^n intersecting inflow boundary.

backtracks to the inflow boundary $x=a$ during $[t_{n-1},t_n]$. We then evaluate each term in Eq. (2.3) in parallel to what we did earlier to obtain

$$\begin{aligned}&\int_{x_{i-1}}^{x_i} u(x,t_n)w(x,t_n)dx+\int_{x_{i-1}}^{x_i}\Delta t(x)D(x,t_n)u_x(x,t_n)w_x(x,t_n)\,dt\\&\quad+\Delta t(x_{i-1})Du_x(x_{i-1},t_n)w(x_{i-1}^+,t_n)-\Delta t(x_i)Du_x(x_i,t_n)w(x_i^-,t_n)\\&=\int_{x_{i-1}}^{x_i}\Delta t(x)f(x,t_n)w(x,t_n)dx+\int_{t_{n,i}^*}^{t_{n,i-1}^*}g(t)w(a,t)dt+E_i(u,w).\end{aligned}\tag{2.8}$$

The local truncation error $E_i(u,w)$ is defined similarly to that in Eq. (2.7), with the lower limit t_{n-1} in the first integral being replaced by $t_n^*(x)$, the lower limit t_{n-1} in inter-element advective and diffusive fluxes being replaced by $t_{n,i-1}^*$ and $t_{n,i}^*$, and $\Delta t(x)=t_n-t_n^*(x)$.

Next we consider the outflow boundary. At the space-time outflow boundary $\{(b,t);t\in[t_{n-1},t_n]\}$, we define the space-time prism $\Omega^{(o)}$ by the characteristics $r(t;b,t_n)$, $[t_{n-1},t_n]$ and $[b^*,b]$. We now define the test functions w on this control volume as follows:

$$w(r(s;b,t),s)=w(b,t),\quad s\in(t_{n-1},t],\quad t\in[t_{n-1},t_n].\tag{2.9}$$

Thus, once the test functions $w(b,t)$ are specified for $t\in[t_{n-1},t_n]$, they are determined completely within the control volume $\Omega^{(o)}$ by the constraint (2.9).

We integrate (2.1) multiplied by w over $\Omega^{(o)}$ and approximate the relevant temporal integrals by an Euler quadrature to obtain

$$\begin{aligned}&\int_{t_{n-1}}^{t_n} Vu(b,t)w(b,t)dt+\Delta tDu_x(b,t_n)w(b,t_n^-)\\&=\int_{b^*}^{b}u(x,t_{n-1})w(x,t_{n-1}^+)dx+\int_{t_{n-1}}^{t_n}(t-t_{n-1})Vf(b,t)w(b,t)dt\\&\quad-\int_{t_{n-1}}^{t_n}h(t)w(b,t)dt-\int_{t_{n-1}}^{t_n}(t-t_{n-1})h(t)w_t(b,t)dt-E_{(o)}(u,w).\end{aligned}\tag{2.10}$$

Here $E_{(o)}(u,w)=-E^{(1)}_{(o)}(u,w)+E^{(2)}_{(o)}(u,w)+E^{(3)}_{(o)}(u,w)+E^{(f)}_{(o)}(w)$, with

$$E^{(1)}_{(o)}(u,w)=\int_{t_{n-1}}^{t_n}(uw)(r(t;b,t_n),t)(V(r(t;b,t_n),t)-r_t(t;b,t_n))dt+\int_{\Omega^{(o)}}u(w_t+Vw_x)dxdt,$$

$$E^{(2)}_{(o)}(u,w)=-w(b,t_n^-)\int_{t_{n-1}}^{t_n}\int_t^{t_n}\frac{d}{d\tau}(Du_x)(r(\tau;b,t_n),\tau)d\tau dt,$$

$$\begin{aligned}E^{(3)}_{(o)}(u,w)=&\int_{t_{n-1}}^{t_n}w_t(b,t)\int_{t_{n-1}}^{t}\Big[\int_\theta^t\frac{d}{d\tau}(Du_y)(r(\tau;b,t),\tau)d\tau d\theta dt\\&+\mathcal{O}(\Delta t)(Du_y)(r(\theta;b,t),\theta)\Big]d\theta dt,\end{aligned}$$

$$\begin{aligned}E^{(f)}_{(o)}(w)=&\int_{t_{n-1}}^{t_n}w(b,t)V(b,t)\int_{t_{n-1}}^{t}\int_\theta^t\frac{df}{d\tau}(r(\tau;b,t),\tau)d\tau d\theta dt\\&+\int_{t_{n-1}}^{t_n}w(b,t)V_t(b,t)\int_{t_{n-1}}^{t}(t-\theta)f(r(\theta;b,t),\theta)d\theta dt.\end{aligned}$$

In (2.10) we approximate the diffusion term by its value at the outflow boundary. Moreover, w_x is now a normal derivative and needs to be replaced by the time derivative.

$$\begin{aligned}&\int_{\Omega^{(o)}}Du_y(y,t)w_y(y,t)dydt\\=&-\int_{t_{n-1}}^{t_n}\int_{t_{n-1}}^{t}Du_y(r(\theta;b,t),\theta)w_y(r(\theta;b,t),\theta)r_t(\theta;b,t)d\theta dt\\=&-\int_{t_{n-1}}^{t_n}w_t(b,t)\int_{t_{n-1}}^{t}Du_y(r(\theta;b,t),\theta)(1+\mathcal{O}(\Delta t))d\theta dt\\=&\int_{t_{n-1}}^{t_n}(t-t_{n-1})h(t)w_t(b,t)dt+E^{(3)}_{(o)}(u,w).\end{aligned}$$

3 A family of characteristic discontinuous Galerkin methods

3.1 A global reference equation

We begin by formulating a reference equation for the true solution u to problem (2.1). Note that the test functions $w(x,t)$ in Eqs. (2.7), (2.8), and (2.10) are smooth only within each space-time control volume Ω_i^n and could be discontinuous across each node x_i for $i=1,2,\cdots,I-1$. On the other hand, the inter-element diffusive flux should be continuous. Introduce the average $\{z\}(x_i,t_n)=\frac{1}{2}(z(x_i^-,t_n)+z(x_i^+,t_n))$ and the jump $[[z]](x_i,t_n)=z(x_i^+,t_n)-z(x_i^-,t_n)$ at $i=1,2,\cdots,I-1$ for any function $z(x,t_n)$.

We sum Eqs. (2.7), (2.8) (for $i=1,\cdots,I$), and (2.10) and express the diffusive flux term in the form $\{Du_x\}(x_i,t_n)[[w]](x_i,t_n)$ [4, 17]. We then use the idea in [17, 19] to add an asymmetric term $[[u]](x_i,t_n)\{Dw_x\}(x_i,t_n)$ to the inter-element diffusive flux. Finally, we utilize the continuity of $u(x_i,t_n)$ to introduce an inter-element penalty term

$\alpha_i[[u]](x_i,t_n)[[w]](x_i,t_n)$ to derive the following characteristic NIPG formulation

$$\begin{aligned}
&\int_a^b u(x,t_n)w(x,t_n)dx+\sum_{i=1}^{I}\int_{x_{i-1}}^{x_i}\Delta t(x)D(x,t_n)u_x(x,t_n)w_x(x,t_n)dx\\
&\quad+\int_{t_{n-1}}^{t_n}Vu(b,t)w(b,t)dt+\sum_{i=1}^{I}\alpha_i[[u]](x_i,t_n)[[w]](x_i,t_n)\\
&\quad+\sum_{i=1}^{I}\Delta t(x_i)\Big(\{Du_x\}(x_i,t_n)[[w]](x_i,t_n)-[[u]](x_i,t_n)\{Dw_x\}(x_i,t_n)\Big)\\
=&\int_a^b u(x,t_{n-1})w(x,t_{n-1}^+)dx+\int_a^b\Delta t(x)f(x,t_n)w(x,t_n)dx\\
&\quad+\int_{t_{n-1}}^{t_n}(t-t_{n-1})Vf(b,t)w(b,t)dt+\int_{t_{n-1}}^{t_n}g(t)w(a,t)dt\\
&\quad-\int_{t_{n-1}}^{t_n}h(t)w(b,t)dt-\int_{t_{n-1}}^{t_n}(t-t_{n-1})h(t)w_t(b,t)dt+E(u,w).
\end{aligned}\tag{3.1}$$

Here $E(u,w)=\sum_{i=1}^{I}E_i(u,w)+E_{(o)}(u,w)$, and the term $\Delta t(a)(Du_x)(a,t_n)w(a^+,t_n)$ vanishes due to $\Delta t(a)=0$. Because $w(x,t_n)$ may be discontinuous across inter-element boundary, the summation of the element integrals cannot be simply written as an integral on the entire domain.

3.2 Characteristic NIPG schemes

Let the finite element space S_h consist of piecewise polynomials of degree up to $m(\geq 1)$ with respect to the spatial partition in $[a,b]$. Furthermore, each function in S_h is constant in the time interval $[t_{n-1},t_n]$ at $x=b$. If we replace the true solution u and the test function w in (3.1) by their corresponding approximations in S_h and drop the local truncation error term $E(u,w)$ in (3.1), we obtain a family of characteristic NIPG schemes: find $u_h\in S_h$, such that $\forall w\in S_h$ the following equation holds

$$\begin{aligned}
&\int_a^b u_h(x,t_n)w(x,t_n)dx+\sum_{i=1}^{I}\int_{x_{i-1}}^{x_i}\Delta t(x)D(x,t_n)u_{hx}(x,t_n)w_x(x,t_n)dx\\
&\quad+\int_{t_{n-1}}^{t_n}V(b,t)u_h(b,t)w(b,t)dt+\sum_{i=1}^{I}\alpha_i[[u_h]](x_i,t_n)[[w]](x_i,t_n)\\
&\quad+\sum_{i=1}^{I}\Delta t(x_i)\Big(\{Du_{hx}\}(x_i,t_n)[[w]](x_i,t_n)-[[u_h]](x_i,t_n)\{Dw_x\}(x_i,t_n)\Big)\\
=&\int_a^b u_h(x,t_{n-1})w(x,t_{n-1}^+)dx+\int_a^b\Delta t(x)f(x,t_n)w(x,t_n)dx\\
&\quad+\int_{t_{n-1}}^{t_n}(t-t_{n-1})V(b,t)f(b,t)w(b,t)dt+\int_{t_{n-1}}^{t_n}g(t)w(a,t)dt-\int_{t_{n-1}}^{t_n}h(t)w(b,t)dt.
\end{aligned}\tag{3.2}$$

Here α_i is a parameter for the interior penalty term, which can be prescribed independent of the problem (refer to Theorem 5.1). In (3.2) we have used the fact that all the functions in S_h are constant in time at the outflow boundary $x=b$ so that the sixth term on the right side of (3.1) vanishes. Recall that the test function w is defined on $[a,b]$ at time step t_n and at the outflow boundary $x=b$ for $t\in[t_{n-1},t_n]$ and is extended by constant along the characteristics (2.4). Hence, a characteristic tracking from t_{n-1} (or the inflow boundary $x=a$) to t_n (or the outflow boundary $x=b$) needs to be used in the evaluation of the test function w in the first and fourth terms on the right side of (3.1). To insure that the characteristics do not intersect with each other, we impose the constraint $|V_x(x,t)|\Delta t<1$ on the time step Δt [15, 27]. Finally, since the penalty term and the diffusion term in this scheme are of the same order, the introduction of the penalty term will not hurt the condition number of the characteristic NIPG scheme.

3.3 Characteristic OBB, SIPG, and IIPG schemes

Choosing $\alpha_i\equiv 0$ in the characteristic NIPG scheme (3.2) yields the characteristic OBB scheme. On the other hand, if we do not add the term $[[u]](x_i,t_n)\{Dw_x\}(x_i,t_n)$ to the inter-element diffusive flux in (3.1) and (3.2), we end up with the following characteristic IIPG scheme:

$$\begin{aligned}
&\int_a^b u_h(x,t_n)w(x,t_n)dx+\sum_{i=1}^{I}\int_{x_{i-1}}^{x_i}\Delta t(x)D(x,t_n)u_{hx}(x,t_n)w_x(x,t_n)dx\\
&\quad+\int_{t_{n-1}}^{t_n}V(b,t)u_h(b,t)w(b,t)dt+\sum_{i=1}^{I}\alpha_i[[u_h]](x_i,t_n)[[w]](x_i,t_n)\\
&\quad+\sum_{i=1}^{I}\Delta t(x_i)\{Du_{hx}\}(x_i,t_n)[[w]](x_i,t_n)\\
=&\int_a^b u_h(x,t_{n-1})w(x,t_{n-1}^+)dx+\int_a^b\Delta t(x)f(x,t_n)w(x,t_n)dx\\
&+\int_{t_{n-1}}^{t_n}(t-t_{n-1})V(b,t)f(b,t)w(b,t)dt+\int_{t_{n-1}}^{t_n}g(t)w(a,t)dt-\int_{t_{n-1}}^{t_n}h(t)w(b,t)dt.
\end{aligned}\tag{3.3}$$

Finally, if we add a symmetric term $[[u]](x_i,t_n)\{Dw_x\}(x_i,t_n)$ to the inter-element diffusive flux in (3.1) and (3.2), we obtain the following characteristic SIPG scheme:

$$\begin{aligned}
&\int_a^b u_h(x,t_n)w(x,t_n)dx+\sum_{i=1}^{I}\int_{x_{i-1}}^{x_i}\Delta t(x)D(x,t_n)u_{hx}(x,t_n)w_x(x,t_n)dx\\
&\quad+\int_{t_{n-1}}^{t_n}V(b,t)u_h(b,t)w(b,t)dt+\sum_{i=1}^{I}\alpha_i[[u_h]](x_i,t_n)[[w]](x_i,t_n)\\
&\quad+\sum_{i=1}^{I}\Delta t(x_i)\Big(\{Du_{hx}\}(x_i,t_n)[[w]](x_i,t_n)+[[u_h]](x_i,t_n)\{Dw_x\}(x_i,t_n)\Big)
\end{aligned}$$

$$=\int_a^b u_h(x,t_{n-1})w(x,t_{n-1}^+)dx+\int_a^b \Delta t(x)f(x,t_n)w(x,t_n)dx$$
$$+\int_{t_{n-1}}^{t_n}(t-t_{n-1})V(b,t)f(b,t)w(b,t)dt+\int_{t_{n-1}}^{t_n}g(t)w(a,t)dt-\int_{t_{n-1}}^{t_n}h(t)w(b,t)dt. \quad (3.4)$$

4 Preliminaries

Let $W_p^m(a,b)$ be the Sobolev spaces consisting of functions whose derivatives up to order-m are p-th integrable on (a,b), equipped with their standard norms; and $H^m=W_2^m$. For any Banach space X, we introduce Sobolev spaces involving time

$$W_p^m(t_1,t_2;X):=\Big\{f(x,t)\ :\Big\|\frac{\partial^k f}{\partial t^k}(\cdot,t)\Big\|_X\in L^p(t_1,t_2),\ 0\le k\le m,\ 1\le p\le\infty\Big\},$$

$$\|f\|_{W_p^m(t_1,t_2;X)}:=\begin{cases}\Big(\sum_{k=0}^m\int_{t_1}^{t_2}\Big\|\frac{\partial^k f}{\partial t^k}(\cdot,t)\Big\|_X^p dt\Big)^{1/p}, & 1\le p<\infty,\\ \max\limits_{0\le k\le m}\operatorname{esssup}_{(t_1,t_2)}\Big\|\frac{\partial^k f}{\partial t^k}(\cdot,t)\Big\|_X, & p=\infty.\end{cases}$$

When it is clear from the context, we use $\|\cdot\|$, $\|\cdot\|_{H^m}$, and $\|\cdot\|_{W_\infty^m}$ to denote $\|\cdot\|_{L^2(a,b)}$, $\|\cdot\|_{H^m(a,b)}$, and $\|\cdot\|_{W_\infty^m(a,b)}$, respectively. We also introduce the discrete (semi-) norms

$$\|f\|_h=\Big(\sum_{i=1}^I\int_{x_{i-1}}^{x_i}\Delta t_n(x)D(x,t_n)f_x^2(x,t_n)dx\Big)^{\frac12},$$
$$\|f(b,t)\|_{L_V^2(t_1,t_2)}=\Big(\int_{t_1}^{t_2}Vf^2(b,t)dt\Big)^{1/2},$$
$$\|f\|_{\widehat{L}^2(0,T;H_D^1)}=\Big(\sum_{n=1}^N\sum_{i=1}^I\int_{x_{i-1}}^{x_i}\Big(\Delta t_n(x)D(x,t_n)f_x^2(x,t_n)dx$$
$$+\Delta t_n(x_i)h^{-1}D(x_i,t_n)[[f]]^2(x_i,t_n)\Big)\Big)^{\frac12},$$
$$\|f\|_{\widehat{L}^\infty(0,T;L^2)}=\max_{0\le n\le N}\|f(x,t_n)\|_{L^2(a,b)}.$$

Let $\Pi p\in S_h$ be the interpolation of $p\in H^{m+2}$. The following estimates hold

$$\begin{aligned}&\|\Pi p-p\|_{H^k(x_{i-1},x_i)}\le Ch^{l-k}\|p\|_{H^l(x_{i-1},x_i)}, && 0\le k\le 2,\quad k\le l\le m+1,\\ &\|\phi\|_{W_\infty^k(x_{i-1},x_i)}\le Ch^{-1/2}\|\phi\|_{W_2^k(x_{i-1},x_i)} && \forall\phi\in S_h,\quad k=0,1.\end{aligned} \quad (4.1)$$

Furthermore, we have the following superconvergence [29, 31–33] for all $\phi\in S_h$:

$$\Big|\sum_{i=1}^I\int_{x_{i-1}}^{x_i}D(x,t_n)(\Pi p-p)_x(x)\phi_x(x)dx\Big|$$
$$\le Ch^{m+1}\|p\|_{L^\infty(0,T;H^{m+2})}\Big[\sum_{i=1}^I\Big(\int_{x_{i-1}}^{x_i}\phi_x^2(x)dx\Big)^{\frac12}+\|\phi\|\Big]. \quad (4.2)$$

When the finite element space S_h consists of piecewise polynomials of $m\geq 3$, we can choose Πp to interpolate p and p_x at the nodes x_{i-1} and x_i on each interval $[x_{i-1},x_i]$, e.g., $(\Pi p-p)(x_i)=(\Pi p-p)_x(x_i)=0$. Then for $x\in[x_{i-1},x_i]$, we have

$$(\Pi p-p)(x)=-\int_x^{x_i}(\Pi p-p)_y(y)dy=\int_x^{x_i}\int_y^{x_i}(\Pi p-p)_{zz}(z)dzdy. \tag{4.3}$$

Then we introduce the following notations:

$$\begin{aligned}
&e(x,t_n)=u_h(x,t_n)-u(x,t_n),\ \ \xi(x,t_n)=u_h(x,t_n)-\Pi u(x,t_n),\quad x\in[a,b),\\
&\eta(x,t)=\Pi u(x,t)-u(x,t),\quad (x,t)\in[a,b)\times(t_{n-1},t_n],\\
&e(b,t)=u_h(b,t_n^-)-u(b,t),\ \ \xi(b,t)=u_h(b,t_n^-)-\Pi u(b,t_n),\ \ \eta(b,t)=\Pi u(b,t_n)-u(b,t).
\end{aligned} \tag{4.4}$$

In this paper, we use ε to denote an arbitrary small positive number, and C to denote a general positive constant that could be assumed different values at different occurrences.

5 Optimal-order L^2 estimates and superconvergence estimates

In this section we derive optimal-order error estimates in the L^2 norm and superconvergence estimates in the discrete energy norm for the characteristic NIPG scheme (3.2), the characteristic IIPG scheme (3.3), and the characteristic SIPG scheme (3.4).

Theorem 5.1. *Let u be the solution to problem (2.1) and u_h be the solution of the characteristic NIPG scheme (3.2) with $\alpha_i=\omega D(x_i,t_n)\Delta t(x_i)h^{-1}$. Further, the spatial partition is uniform for $m=1$. The following optimal-order L^2 estimate and superconvergence estimate hold for any $\omega>0$, except for $m=2$ when a suboptimal-order estimate holds with the h^{m+1} in (5.1) replaced by h^2*

$$\begin{aligned}
&\|u_h-u\|_{\widehat{L}^\infty(0,T;L^2)}+\|u_h-u\|_{\widehat{L}^2(0,T;H_D^1)}+\|(u_h-u)(b,t)\|_{L_V^2(0,T)}\\
\leq& C\Delta t\Big(\|u\|_{L^\infty(0,T;H^2)}+\|u_t(b,t)\|_{L^2(0,T)}+\|u\|_{L^2(0,T;W_\infty^1)}+\Big\|\frac{du}{dt}\Big\|_{L^2(0,T;W_\infty^1)}\\
&+\|f\|_{L^2(0,T;L^2)}+\Big\|\frac{df}{dt}\Big\|_{L^2(0,T;L^2)}\Big)+Ch^{m+1}\Big(\|u\|_{L^\infty(0,T;H^{m+2})}+\|u\|_{H^1(0,T;H^{m+1})}\Big).
\end{aligned} \tag{5.1}$$

Proof. We subtract Eq. (3.1) from Eq. (3.2) and use the fact $w_t(b,t)=0$ to obtain an error equation on $e=u_h-u$

$$\begin{aligned}
&\int_a^b e(x,t_n)w(x,t_n)dx+\sum_{i=1}^I\int_{x_{i-1}}^{x_i}\Delta t(x)D(x,t_n)e_x(x,t_n)w_x(x,t_n)dx\\
&\quad+\int_{t_{n-1}}^{t_n}Ve(b,t)w(b,t)dt+\sum_{i=1}^I\Delta t(x_i)h^{-1}D(x_i,t_n)[[e]](x_i,t_n)[[w]](x_i,t_n)\\
&\quad+\sum_{i=1}^I\Delta t(x_i)D(x_i,t_n)\Big(\{e_x\}(x_i,t_n)[[w]](x_i,t_n)-[[e]](x_i,t_n)\{w_x\}(x_i,t_n)\Big)\\
=&\int_a^b e(x,t_{n-1})w(x,t_{n-1}^+)dx-E(u,w).
\end{aligned} \tag{5.2}$$

We decompose the error $e=u_h-u$ as $e=\xi+\eta$. Since the estimate for η is known, we need only to derive an error estimate for ξ. We choose $w=\xi$ and use the fact that $\eta(x_i,t_n)=0$ and $\eta_x(x_i,t_n)=0$ for $i=1,2,\cdots,I$ to rewrite Eq. (5.2) in terms of ξ and η

$$\begin{aligned}
&\|\xi(x,t_n)\|^2+\|\xi(x,t_n)\|_h^2+\|\xi(b,t)\|^2_{L_V^2(t_{n-1},t_n)}+\sum_{i=1}^{I}\Delta t(x_i)h^{-1}D(x_i,t_n)[[\xi]]^2(x_i,t_n)\\
=&\int_a^b\xi(x,t_{n-1})\xi(x,t_{n-1}^+)dx+\int_a^b\eta(x,t_{n-1})\xi(x,t_{n-1}^+)dx-\int_a^b\eta(x,t_n)\xi(x,t_n)dx\\
&-\int_{t_{n-1}}^{t_n}V\eta(b,t)\xi(b,t)dt-\sum_{i=1}^{I}\int_{x_{i-1}}^{x_i}\Delta t(x)D(x,t_n)\eta_x(x,t_n)\xi_x(x,t_n)dx-E(u,\xi). \qquad (5.3)
\end{aligned}$$

Now we estimate the right-hand side of (5.3) term by term. We begin by bounding a shifted L^2-norm of $\xi(x,t_{n-1}^+)=\xi(\tilde{x},t_n)$. We use $y=r(t_{n-1};x,t_n)$ to denote the variable at time step t_{n-1} and x to denote the variable at time step t_n.

$$\begin{aligned}
\int_a^b\xi^2(y,t_{n-1}^+)dy&=\int_{\tilde{a}}^b\xi^2(x,t_n)r_x(t_{n-1};x,t_n)dx+\int_{t_{n-1}}^{t_n}\xi^2(b,t)r_t(t_{n-1};b,t)dt\\
&\le(1+C\Delta t)\big(\|\xi(x,t_n)\|^2_{L^2(\tilde{a},b)}+\|\xi(b,t)\|^2_{L_V^2(t_{n-1},t_n)}\big). \qquad (5.4)
\end{aligned}$$

Then, the first term on the right-hand side of (5.3) can be bounded as follows:

$$\begin{aligned}
&\Big|\int_a^b\xi(x,t_{n-1})\xi(x,t_{n-1}^+)dx\Big|\\
\le&\frac{1}{2}\|\xi(x,t_{n-1})\|^2+\frac{1}{2}(1+C\Delta t)\Big(\|\xi(x,t_n)\|^2_{L^2(\tilde{a},b)}+\|\xi(b,t)\|^2_{L_V^2(t_{n-1},t_n)}\Big). \qquad (5.5)
\end{aligned}$$

We use Cauchy inequality to bound the fourth term on the right side of (5.3) by

$$\begin{aligned}
\int_{t_{n-1}}^{t_n}V\eta(b,t)\xi(b,t)dt&\le\varepsilon\int_{t_{n-1}}^{t_n}V\xi^2(b,t)dt+C\int_{t_{n-1}}^{t_n}V\eta^2(b,t)dt\\
&\le\varepsilon\|\xi(b,t)\|^2_{L_V^2(t_{n-1},t_n)}+C(\Delta t)^2\|u(b,t)\|^2_{H^1(t_{n-1},t_n)}.
\end{aligned}$$

The estimates of the remaining terms on the right-hand side of (5.3) are lengthy and will be presented in Lemmas A.1-A.3 in the Appendix.

$$\begin{aligned}
&\Big|\int_a^b\eta(x,t_{n-1})\xi(x,t_{n-1}^+)dx-\int_a^b\eta(x,t_n)\xi(x,t_n)dx\Big|\\
\le&\varepsilon\|\xi(x,t_n)\|_h^2+C\Delta t\|\xi(x,t_n)\|^2+C\Delta t\|\xi(b,t)\|^2_{L_V^2(t_{n-1},t_n)}\\
&+Ch^{2m+2}(\|u\|^2_{H^1(t_{n-1},t_n;H^{m+1})}+\Delta t\|u\|^2_{L^\infty(0,T;H^{m+1})})+C(\Delta t)^3\|u\|^2_{L^\infty(0,T;H^2)}, \qquad (5.6)\\
&\Big|\sum_{i=1}^{I}\int_{x_{i-1}}^{x_i}\Delta t(x)D(x,t_n)\eta_x(x,t_n)\xi_x(x,t_n)dx\Big|\\
\le&\varepsilon\|\xi(x,t_n)\|_h^2+C\Delta t\|\xi(x,t_n)\|^2+C(\Delta t)^3\|u\|^2_{L^\infty(0,T;H^2)}+C\Delta th^{2m+2}\|u\|^2_{L^\infty(0,T;H^{m+2})}, \qquad (5.7)
\end{aligned}$$

and

$$\begin{aligned}\big|E(u,\xi)\big| \leq & C\Delta t\|\xi(x,t_n)\|^2+\varepsilon\|\xi(x,t_n)\|_h^2+C\Delta t\|\xi(b,t)\|^2_{L^2_V(t_{n-1},t_n)}\\ &+\varepsilon\sum_{i=1}^{I}\Delta t(x_i)h^{-1}D(x_i,t_n)[[\xi]]^2(x_i,t_n)+C(\Delta t)^2\Big(\Big\|\frac{du}{dt}\Big\|^2_{L^2(t_{n-1},t_n;W^1_\infty)}\\ &+\|u\|^2_{L^2(t_{n-1},t_n;W^1_\infty)}+\Big\|\frac{df}{dt}\Big\|^2_{L^2(t_{n-1},t_n;L^2)}+\|f\|^2_{L^2(t_{n-1},t_n;L^2)}\Big).\end{aligned}\tag{5.8}$$

We incorporate Eqs. (5.5)-(5.8) into Eq. (5.3) to obtain

$$\begin{aligned}&\|\xi(x,t_n)\|^2+\|\xi(x,t_n)\|_h^2+\|\xi(b,t)\|^2_{L^2_V(t_{n-1},t_n)}+\sum_{i=1}^{I}\Delta t_n(x_i)h^{-1}D(x_i,t_n)[[\xi]]^2(x_i,t_n)\\ \leq&\Big(\frac{1}{2}+C\Delta t\Big)(\|\xi(x,t_{n-1})\|^2+\|\xi(x,t_n)\|^2)+\Big(\frac{1}{2}+\varepsilon+C\Delta t\Big)\|\xi(b,t)\|^2_{L^2_V(t_{n-1},t_n)}\\ &+\varepsilon\|\xi(x,t_n)\|_h^2+\varepsilon\sum_{i=1}^{I}\Delta t_n(x_i)h^{-1}D(x_i,t_n)[[\xi]]^2(x_i,t_n)\\ &+Ch^{2m+2}\Big(\Delta t\|u\|^2_{L^\infty(0,T;H^{m+2})}+\|u\|^2_{H^1(t_{n-1},t_n;H^{m+1})}\Big)\\ &+C(\Delta t)^2\Big(\Delta t\|u\|^2_{L^\infty(0,T;H^2)}+\|u(b,t)\|^2_{H^1(t_{n-1},t_n)}+\|u\|^2_{L^2(t_{n-1},t_n;W^1_\infty)}\\ &+\Big\|\frac{du}{dt}\Big\|^2_{L^2(t_{n-1},t_n;W^1_\infty)}+\|f\|^2_{L^2(t_{n-1},t_n;L^2)}+\Big\|\frac{df}{dt}\Big\|^2_{L^2(t_{n-1},t_n;L^2)}\Big).\end{aligned}$$

We choose $\varepsilon=\frac{1}{4}$ and sum the resulting inequality on n to rewrite the proceeding inequality as ($N_1\leq N$)

$$\begin{aligned}&\|\xi(x,t_{N_1})\|^2+\|\xi\|^2_{\widehat{L}^2(0,t_{N_1};H^1_D)}+\|\xi(b,t)\|^2_{L^2_V(0,t_{N_1})}\\ \leq& C\sum_{n=0}^{N_1-1}\Delta t_n\|\xi(x,t_n)\|^2+Ch^{2m+2}\Big(\|u\|^2_{L^\infty(0,T;H^{m+2})}+\|u\|^2_{H^1(0,T;H^{m+1})}\Big)\\ &+C(\Delta t)^2\Big(\|u\|^2_{L^\infty(0,T;H^2)}+\|u(b,t)\|^2_{H^1(0,T)}+\|u\|^2_{L^2(0,T;W^1_\infty)}\\ &+\Big\|\frac{du}{dt}\Big\|^2_{L^2(0,T;W^1_\infty)}+\|f\|^2_{L^2(0,T;L^2)}+\Big\|\frac{df}{dt}\Big\|^2_{L^2(0,T;L^2)}\Big).\end{aligned}\tag{5.9}$$

We apply Gronwall's inequality to obtain

$$\begin{aligned}&\|\xi\|_{\widehat{L}^\infty(0,T;L^2)}+\|\xi\|_{\widehat{L}^2(0,T;H^1_D)}+\|\xi(b,t)\|_{L^2_V(0,T)}\\ \leq& C\Delta t\Big(\|u\|_{L^\infty(0,T;H^2)}+\|u(b,t)\|_{H^1(0,T)}+\|u\|_{L^2(0,T;W^1_\infty)}+\Big\|\frac{du}{dt}\Big\|_{L^2(0,T;W^1_\infty)}\\ &+\|f\|_{L^2(0,T;L^2)}+\Big\|\frac{df}{dt}\Big\|_{L^2(0,T;L^2)}\Big)+Ch^{m+1}\Big(\|u\|_{L^\infty(0,T;H^{m+2})}+\|u\|_{H^1(0,T;H^{m+1})}\Big).\end{aligned}\tag{5.10}$$

We combine (5.10) with the estimate for η to finish the proof for (5.1). Note that we used the condition $\{\eta_x(x_i,t_n)\}=0$ before (5.3), which requires the auxiliary function Πu to interpolate u and u_x at the nodes x_{i-1} and x_i on each interval $[x_{i-1},x_i]$, i.e., $m\geq 3$.

For m = 1 or 2, we have to choose Πu to be the Lagrange interpolation of the true solution u and so have the following extra term to bound since $\{\eta_x\}(x_i,t_n)\}\neq 0$ in general

$$-\sum_{i=1}^{I-1}\Delta t(x_i)D(x_i,t_n)\{\eta_x\}(x_i,t_n)[[\xi]](x_i,t_n). \tag{5.11}$$

We use the Peano kernel theorem to get

$$\eta(x,t_n)\Big|_{[x_{i-1},x_i]}=\int_{x_{i-1}}^{x_i}G_i^{(m)}(y;x)\frac{\partial^{m+1}u}{\partial y^{m+1}}(y,t_n)dy,\quad 1\leq i\leq I,\quad m=1,2. \tag{5.12}$$

Here $G_i^{(1)}(y;x)=(x_i-x)(y-x_{i-1})/h_i$ for $y\in[x_{i-1},x]$ or $(x-x_{i-1})(x_i-y)/h_i$ otherwise. We differentiate this expression with respect to x to obtain the following expansion for a uniform spatial partition

$$\{\eta_x\}(x_i,t_n)=\frac{1}{2}\int_{x_i}^{x_{i+1}}\frac{x_{i+1}-y}{h}\int_{x_i}^{y}u_{zzz}(z)dzdy-\frac{1}{2}\int_{x_{i-1}}^{x_i}\frac{y-x_{i-1}}{h}\int_{x_i}^{y}u_{zzz}(z)dzdy. \tag{5.13}$$

We use this estimate to bound (5.11) by

$$\varepsilon\sum_{i=1}^{I-1}\Delta t(x_i)h^{-1}D(x_i,t_n)[[\xi]]^2(x_i,t_n)+C\sum_{i=1}^{I-1}\Delta t(x_i)h^4D(x_i,t_n)\int_{x_{i-1}}^{x_i}u_{xxx}^2(x,t_n)dx. \tag{5.14}$$

The rest can be estimated as before and we finish the proof for the case $m=1$. For the case of $m=2$, $G_i^{(2)}(y;x)$ is of the form

$$G_i^{(2)}(y;x):=\begin{cases}\dfrac{2}{h^2}(x_i-y)^2(x-x_{i-1})(x-x_{i-1/2})-(x-y)^2\\ \quad-\dfrac{4}{h^2}(x_{i-1/2}-y)^2(x-x_{i-1})(x-x_i), & y\in(x_{i-1},x),\ x<x_{i-1/2},\\ \dfrac{2}{h^2}(x_i-y)^2(x-x_{i-1})(x-x_{i-1/2})\\ \quad-\dfrac{4}{h^2}(x_{i-1/2}-y)^2(x-x_{i-1})(x-x_i), & y\in(x,x_{i-1/2}),\\ \dfrac{2}{h^2}(x_i-y)^2(x-x_{i-1})(x-x_{i-1/2})-(x-y)^2, & y\in(x_{i-1/2},x),\\ \dfrac{2}{h^2}(x_i-y)^2(x-x_{i-1})(x-x_{i-1/2}), & y\in(x,x_i),\ x>x_{i-1/2}.\end{cases} \tag{5.15}$$

We differentiate $\eta(x,t_n)$ with respect to x in (5.12) incorporated with (5.15) on intervals $[x_{i-1},x_i]$ and $[x_i,x_{i+1}]$ and use the decomposition $u_{yyy}(y,t_n)=u_{yyy}(x_i,t_n)+(u_{yyy}(y,t_n)-$

$u_{yyy}(x_i,t_n)$) to compute the average $\{\eta_x\}(x_i,t_n)$ to obtain

$$\begin{aligned}
&\{\eta_x\}(x_i,t_n)\\
&=-\frac{1}{4h}\Big(\int_{x_{i-1}}^{x_{i-1/2}}(x_{i-1}-y)^2u_{yyy}(y,t_n)dy+\int_{x_{i-1/2}}^{x_i}(x_i-y)(3y-2x_{i-1}-x_i)u_{yyy}(y,t_n)dy\\
&\quad+\int_{x_i}^{x_{i+1/2}}(y-x_i)(2x_{i+1}+x_i-3y)u_{yyy}(y,t_n)dy+\int_{x_{i-1/2}}^{x_{i+1}}(x_{i+1}-y)^2u_{yyy}(y,t_n)dy\Big)\\
&=-\frac{h^2}{48}u_{xxx}(x_i,t_n)+\mathcal{O}(h^3).
\end{aligned}$$

This expansion is of the same order as (5.13), so (5.14) and hence a suboptimal-order estimate holds for $m=2$. □

Remark 5.1. The suboptimal-order L^2 estimate proved for $m=2$ coincides with the numerical observations [7,23]. The Peano kernel approach (5.12) would prove an optimal-order L^2 estimate only for the odd-order NIPG schemes under the assumption of a uniform spatial partition. We alleviate these restrictions via the use of the Hermite interpolation in the proof.

Remark 5.2. Theorem 5.1 holds for any fixed $\omega>0$ with the generic constant C in the theorem depending on ω. However, the theorem does not cover the characteristic OBB scheme, even though numerical results show that this is the case. The current proof will lead to a suboptimal-order error estimate for the characteristic OBB schemes.

Theorem 5.2. *Assume that the condition of Theorem* 5.1 *holds. Let u_h be the solution of the characteristic IIPG scheme* (3.3) *or the characteristic SIPG scheme* (3.4) *with $\alpha_i=\omega D(x_i,t_n)\Delta t(x_i)h^{-1}$. Further, the spatial partition is uniform for $m=1$. Then there exists a sufficiently large constant $\omega_0>0$ such that for any $\omega\geq\omega_0$, the optimal-order L^2 estimate and the superconvergence estimate* (5.1) *hold, except for $m=2$ when a suboptimal-order estimate holds with the h^{m+1} in* (5.1) *replaced by h^2.*

Proof. If we compare the characteristic NIPG scheme (3.2) with the characteristic IIPG scheme (3.3) and the characteristic SIPG scheme (3.4), we see that we need only to bound one extra term as follows in addition to those in (5.3)

$$\begin{aligned}
&\Big|\sum_{i=1}^{I}\Delta t(x_i)D(x_i,t_n)\{\xi_x\}(x_i,t_n)[[\xi]](x_i,t_n)\Big|\\
&\leq\varepsilon\sum_{i=1}^{I}\Delta t(x_i)h^{-1}D(x_i,t_n)\big(\xi_x^2(x_i^-,t_n)+\xi_x^2(x_i^+,t_n)\big)+\frac{C}{4\varepsilon}\sum_{i=1}^{I}\Delta t(x_i)hD(x_i,t_n)[[\xi]]^2(x_i,t_n)\\
&\leq\varepsilon\sum_{i=1}^{I}h\big(\Delta t(x_i)D(x_i,t_n)\xi_x^2(x_i^-,t_n)+\Delta t(x_{i-1})D(x_{i-1},t_n)\xi_x^2(x_{i-1}^+,t_n)\big)\\
&\quad+\frac{C}{4\varepsilon}\sum_{i=1}^{I}\Delta t(x_i)h^{-1}D(x_i,t_n)[[\xi]]^2(x_i,t_n)
\end{aligned}$$

$$\leq \varepsilon \sum_{i=1}^{I} \int_{x_{i-1}}^{x_i} \Delta t(x) D(x,t_n) \xi_x^2(x,t_n) dx + \frac{C}{4\varepsilon} \sum_{i=1}^{I} \Delta t(x_i) h^{-1} D(x_i,t_n) [[\xi]]^2(x_i,t_n).$$

If we choose $\varepsilon = 1/4$ and $\omega_0 = C/\varepsilon$, we manage to hide these two terms in (5.9). The rest of proof in Theorem 5.1 goes through. □

Remark 5.3. Theorems 5.1 and 5.2 give the optimal-order L^2 estimates for the standard NIPG, IIPG, and SIPG schemes for parabolic equations, since (2.1) reduces to a parabolic problem and the characteristic DG schemes reduces to the standard DG schemes in the case of $V \equiv 0$ in (2.1). Furthermore, a similar argument gives an optimal-order estimate and superconvergence estimate for the DG schemes for elliptic problems.

6 Numerical experiments

We carry out numerical experiments to illustrate the performance of the characteristic NIPG scheme, the characteristic OBB scheme, the characteristic IIPG scheme, and the characteristic SIPG scheme developed in this paper. In Section 6.1 we conduct numerical experiments to investigate the spatial and temporal convergence rates of the cubic characteristic NIPG scheme as a representative of high-order characteristic DG schemes, as well as the linear characteristic NIPG, OBB, IIPG, and SIPG schemes. In Section 6.2 we compare the performance of the characteristic DG scheme with some well-received standard DG schemes.

6.1 Spatial and temporal convergence rates of the characteristic DG schemes

We carry out numerical experiments to observe the spatial and temporal convergence rates of the characteristic discontinuous Galerkin schemes. We present the numerical results only for the cubic characteristic NIPG scheme (3.2) with $\omega = 1$ in the Theorem 5.1, a representative of the characteristic DG schemes developed in this paper.

We use a standard test problem of the transport of a Gaussian pulse subject to Eq. (2.1) with the initial configuration being given by

$$u_o(x) = \exp\left(-\frac{(x-x_c)^2}{2\sigma^2}\right), \tag{6.1}$$

where x_c and σ are the centered and standard deviations of the Gaussian pulse. In the numerical experiments the data are chosen as follows: $(a,b) = (0,2)$, $(0,T) = (0,1)$, $D = 0.0001$, $x_c = 0.5$, and $\sigma = 0.05$. In addition,

$$V(x,t) = 1 + 0.01x$$

is used. We use a linear regression to fit the convergence rates and the associated constants in the error estimates

$$\|u_h(x,T) - u(x,T)\|_{L^p(a,b)} \leq C_\alpha h^\alpha + C_\beta (\Delta t)^\beta, \quad \text{with } p = 1, 2, \infty. \tag{6.2}$$

Table 1: Spatial convergence rate of the cubic characteristic NIPG scheme.

h	Δt	L_2 Error	L_1 Error	L_∞ Error
1/16	1/300	8.9974e-004	4.3598e-004	6.7663e-003
1/18	1/300	5.0648e-004	2.2985e-004	4.3117e-003
1/20	1/300	3.1483e-004	1.3775e-004	2.8747e-003
1/22	1/300	2.1907e-004	9.5743e-005	1.9954e-003
1/24	1/300	1.7160e-004	7.6409e-005	1.4363e-003
1/26	1/300	1.4861e-004	6.8708e-005	1.0688e-003
		$\alpha=4.10$	$\alpha=4.30$	$\alpha=3.90$
		$c_\alpha=79$	$c_\alpha=68$	$c_\alpha=274$

Table 2: Temporal convergence rate of the cubic characteristic NIPG scheme.

h	Δt	L_2 Error	L_1 Error	L_∞ Error
1/500	1/10	3.9124e-003	1.9239e-003	1.0957e-002
1/500	1/14	2.7960e-003	1.3746e-003	7.8319e-003
1/500	1/18	2.1753e-003	1.0693e-003	6.0939e-003
1/500	1/22	1.7801e-003	8.7496e-004	4.9872e-003
1/500	1/26	1.5064e-003	7.4041e-004	4.2207e-003
1/500	1/30	1.3057e-003	6.4172e-004	3.6584e-003
		$\beta=1.00$	$\beta=1.00$	$\beta=1.00$
		$c_\beta=0.04$	$c_\beta=0.02$	$c_\beta=0.11$

Table 3: Spatial convergence rates for the linear characteristic NIPG scheme.

h	Δt	L_2 Error	L_1 Error	L_∞ Error
1/50	1/500	7.9383e-003	3.7963e-003	3.2820e-002
1/60	1/500	4.7482e-003	2.2328e-003	2.0341e-002
1/70	1/500	3.0434e-003	1.4169e-003	1.3335e-002
1/80	1/500	2.0700e-003	9.4668e-004	9.1807e-003
1/90	1/500	1.4808e-003	6.7164e-004	6.5899e-003
1/100	1/500	1.1051e-003	4.9861e-004	4.8990e-003
		$\alpha=2.85$	$\alpha=2.94$	$\alpha=2.75$
		$c_\alpha=560$	$c_\alpha=376$	$c_\alpha=1568$

We perform two kinds of computations. The first tests the spatial convergence rate of the characteristic NIPG scheme, where we fix a small time step Δt and compute the constant C_α and the rate α with respect to h; the other tests the temporal convergence rate, where we choose a small grid size h and calculate the constant C_β and the rate β with respect to Δt. The results are presented in Tables 1 and 2, fitting the pairs (C_α, α) and (C_β, β). These results show that the characteristic NIPG scheme possesses the optimal-order spatial and temporal convergence rates as predicted by Theorem 5.1. Moreover we notice that the generic constant C_β of the temporal error is much smaller than the generic constant C_α of the spatial error. This reflects the strength of the characteristic discontinuous Galerkin schemes.

Table 4: Temporal convergence rates for the linear characteristic NIPG scheme.

h	Δt	L_2 Error	L_1 Error	L_∞ Error
1/500	1/10	2.0074e-003	9.8546e-004	6.9558e-003
1/500	1/20	1.0200e-003	5.0028e-004	3.5064e-003
1/500	1/30	6.7832e-004	3.3321e-004	2.3336e-003
1/500	1/40	5.2975e-004	2.6052e-004	1.7940e-003
1/500	1/50	4.1509e-004	2.0295e-004	1.4284e-003
1/500	1/60	3.7114e-004	1.8186e-004	1.2480e-003
		$\beta = 1.00$	$\beta = 1.00$	$\beta = 1.00$
		$c_\beta = 0.02$	$c_\beta = 0.01$	$c_\beta = 0.06$

Table 5: Spatial convergence rates for the linear characteristic OBB scheme.

h	Δt	L_2 Error	L_1 Error	L_∞ Error
1/50	1/500	7.9758e-003	3.8167e-003	3.2908e-002
1/60	1/500	4.7769e-003	2.2469e-003	2.0414e-002
1/70	1/500	3.0656e-003	1.4274e-003	1.3390e-002
1/80	1/500	2.0875e-003	9.5484e-004	9.2220e-003
1/90	1/500	1.4950e-003	6.7856e-004	6.6201e-003
1/100	1/500	1.1171e-003	5.0464e-004	4.9210e-003
		$\alpha = 2.84$	$\alpha = 2.93$	$\alpha = 2.75$
		$c_\alpha = 544$	$c_\alpha = 365$	$c_\alpha = 1557$

Table 6: Temporal convergence rates for the linear characteristic OBB scheme.

h	Δt	L_2 Error	L_1 Error	L_∞ Error
1/500	1/10	8.5102e-004	3.3275e-004	5.0638e-003
1/500	1/20	4.1470e-004	1.6647e-004	2.3398e-003
1/500	1/30	2.8711e-004	1.1569e-004	1.6001e-003
1/500	1/40	2.2931e-004	1.0072e-004	1.2024e-003
1/500	1/50	1.6627e-004	6.7155e-005	9.3844e-004
1/500	1/60	1.8144e-004	8.2453e-005	8.8697e-004
		$\beta = 0.95$	$\beta = 0.90$	$\beta = 1.00$
		$c_\beta = 0.007$	$c_\beta = 0.002$	$c_\beta = 0.048$

We present the spatial and temporal convergence rates of the linear characteristic NIPG, OBB, IIPG, and SIPG schemes with $\gamma = 1$ in Tables 3-10. The numerical results with other values of γ are similar and are skipped. We observe that the linear characteristic NIPG, OBB, IIPG, and SIPG schemes possess second-order convergence rate in space and first-order convergence rate in time in the norms of L^1, L^2, and L^∞. We notice again that the generic constant C_β of the temporal error is much smaller than the generic constant C_α of the spatial error. Even though these schemes are second order in space and first order in time, the absolute errors with $\Delta t = 1/10$ are much smaller than those with $h = 1/50$. This reflects the strength of the characteristic discontinuous Galerkin schemes developed in this paper.

Table 7: Spatial convergence rates for the linear characteristic SIPG scheme.

h	Δt	L_2 Error	L_1 Error	L_∞ Error
1/50	1/500	8.2680e-003	3.9794e-003	3.4553e-002
1/60	1/500	4.9673e-003	2.3583e-003	2.1548e-002
1/70	1/500	3.1745e-003	1.4813e-003	1.4149e-002
1/80	1/500	2.1350e-003	9.9393e-004	9.7213e-003
1/90	1/500	1.4980e-003	6.8884e-004	6.9443e-003
1/100	1/500	1.0890e-003	4.9792e-004	5.1291e-003
		$\alpha=2.93$	$\alpha=3.00$	$\alpha=2.76$
		$c_\alpha=796$	$c_\alpha=511$	$c_\alpha=1711$

Table 8: Temporal convergence rates for the linear characteristic SIPG scheme.

h	Δt	L_2 Error	L_1 Error	L_∞ Error
1/500	1/10	2.2614e-002	8.4221e-003	1.2922e-001
1/500	1/20	4.9750e-003	1.9266e-003	2.7681e-002
1/500	1/30	6.0566e-003	2.2704e-003	3.4917e-002
1/500	1/40	1.5927e-003	6.5884e-004	8.5701e-003
1/500	1/50	1.5539e-003	6.1899e-004	8.4374e-003
1/500	1/60	8.1026e-004	3.6274e-004	3.9055e-003
		$\beta=1.75$	$\beta=1.67$	$\beta=1.82$
		$c_\beta=1.29$	$c_\beta=0.39$	$c_\beta=8.86$

Table 9: Spatial convergence rates for the linear characteristic IIPG scheme.

h	Δt	L_2 Error	L_1 Error	L_∞ Error
1/50	1/500	8.0790e-003	3.8667e-003	3.3670e-002
1/60	1/500	4.8288e-003	2.2862e-003	2.0928e-002
1/70	1/500	3.0769e-003	1.4341e-003	1.3727e-002
1/80	1/500	2.0690e-003	9.5176e-004	9.4387e-003
1/90	1/500	1.4558e-003	6.6580e-004	6.7571e-003
1/100	1/500	1.0647e-003	4.8202e-004	5.0062e-003
		$\alpha=2.93$	$\alpha=3.01$	$\alpha=2.76$
		$c_\alpha=778$	$c_\alpha=514$	$c_\alpha=1649$

6.2 Comparison of the characteristic DG schemes with standard DG schemes

In this subsection we compare the performance of the characteristic DG scheme with some well-received standard DG scheme to gain better understanding of the proposed characteristic DG scheme. We carry out the comparison of the characteristic DG scheme with the third-order Runge-Kutta DG scheme with piecewise linear polynomials in space for the case of $D=0$. This is what the original Runge-Kutta DG schemes were developed for. Moreover, in this case all the four characteristic DG schemes reduce to one scheme. We present the results in Tables 11-13. Table 11 contains the numerical results of the popular third-order Runge-Kutta DG scheme with piecewise linear finite elements in space,

Table 10: Temporal convergence rates for the linear characteristic IIPG scheme.

h	Δt	L_2 Error	L_1 Error	L_∞ Error
1/500	1/10	2.1693e-003	1.0380e-003	8.6592e-003
1/500	1/20	1.1047e-003	5.3002e-004	4.3450e-003
1/500	1/30	7.4032e-004	3.5499e-004	2.9679e-003
1/500	1/40	5.8580e-004	2.8325e-004	2.2498e-003
1/500	1/50	4.4944e-004	2.1543e-004	1.7729e-003
1/500	1/60	4.1314e-004	2.0011e-004	1.5400e-003
		$\beta=0.95$	$\beta=1.0$	$\beta=1.0$
		$c_\beta=0.02$	$c_\beta=0.01$	$c_\beta=0.08$

Table 11: Spatial convergence rate of the 3rd-order Runge-Kutta DG scheme with linear elements in space for the pure advection problem.

h	Δt	L_2 Error	L_1 Error	L_∞ Error
1/50	1/500	1.0167e-002	4.8793e-003	4.0606e-002
1/60	1/500	6.2777e-003	3.0331e-003	2.5697e-002
1/70	1/500	4.1156e-003	1.9574e-003	1.7040e-002
1/80	1/500	2.8377e-003	1.3477e-003	1.1779e-002
1/90	1/500	2.0414e-003	9.7044e-004	8.4417e-003
1/100	1/500	1.5223e-003	7.3032e-004	6.2422e-003
		$\alpha=2.75$	$\alpha=2.76$	$\alpha=2.71$
		$c_\alpha=475$	$c_\alpha=239$	$c_\alpha=1655$

Table 12: Spatial convergence rate of the linear characteristic NIPG scheme for the pure advection problem.

h	Δt	L_2 Error	L_1 Error	L_∞ Error
1/50	1/500	5.6355e-003	2.5936e-003	2.4644e-002
1/60	1/500	3.1874e-003	1.4355e-003	1.5356e-002
1/70	1/500	2.0352e-003	9.0762e-004	1.0256e-002
1/80	1/500	1.4110e-003	6.2098e-004	7.2776e-003
1/90	1/500	1.0349e-003	4.5520e-004	5.4093e-003
1/100	1/500	7.8998e-004	3.4526e-004	4.1625e-003
		$\alpha=2.83$	$\alpha=2.90$	$\alpha=2.55$
		$c_\alpha=348$	$c_\alpha=213$	$c_\alpha=533$

which demonstrates the optimal spatial convergence rates of the scheme. We present the numerical results generated by the linear characteristic NIPG scheme in Table 12 to test the spatial convergence rate of the linear characteristic DG scheme. We observe that the linear characteristic DG scheme generates comparable results with the third-order Runge-Kutta DG scheme when very fine time steps are used.

Because the Runge-Kutta DG scheme is explicit and requires the Courant number to be less than one [11], we cannot test the temporal convergence rates unless we use a much higher-order finite elements in space. Hence we skip the test for temporal convergence rate of the Runge-Kutta DG scheme. Table 13 contains the numerical results, which test the temporal convergence rates of the linear characteristic NIPG scheme. These results

Table 13: Temporal convergence rate of the linear characteristic NIPG scheme for the pure advection problem.

h	Δt	L_2 Error	L_1 Error	L_∞ Error
1/500	1/10	4.2599e-003	2.0346e-003	1.2220e-002
1/500	1/15	2.8396e-003	1.3563e-003	8.1472e-003
1/500	1/20	2.1296e-003	1.0172e-003	6.1114e-003
1/500	1/25	1.7036e-003	8.1370e-004	4.8900e-003
1/500	1/30	1.4197e-003	6.7808e-004	4.0758e-003
1/500	1/35	1.2169e-003	5.8121e-004	3.4949e-003
1/500	1/40	1.0648e-003	5.0856e-004	3.0596e-003
1/500	1/45	9.4647e-004	4.5205e-004	2.7209e-003
1/500	1/50	8.5184e-004	4.0685e-004	2.4499e-003
1/500	1/55	7.7442e-004	3.6987e-004	2.2283e-003
1/500	1/60	7.0992e-004	3.3905e-004	2.0439e-003
		$\beta=1.00$	$\beta=1.00$	$\beta=1.00$
		$c_\beta=0.04$	$c_\beta=0.02$	$c_\beta=0.12$

show that the characteristic DG scheme generates accurate numerical solutions even if very large time steps are used (the Courant number is greater than 50 in this example). This justifies the development of the proposed characteristic DG schemes, i.e., which allow the use of large time steps in generating stable and accurate numerical solutions.

7 Concluding remarks

In this article we develop a family of characteristic discontinuous Galerkin methods for one-dimensional transient advection-diffusion equations by using an Eulerian-Lagrangian approach within a primal discontinuous Galerkin framework. These include the characteristic SIPG method, the characteristic NIPG method, the characteristic IIPG method, and the characteristic OBB method. The developed methods retain the numerical advantages of the discontinuous Galerkin methods as well as characteristic methods. Further, we prove an optimal-order error estimate in the L^2 norm and a superconvergence estimate in a weighted energy norm for the characteristic NIPG, SIPG, and IIPG methods. Numerical experiments confirm the optimal-order convergence rates in the L^2 norm as proved in the main theorems.

The preliminary comparison of these characteristic DG schemes with the well received Runge-Kutta DG schemes lead to the following observations as anticipated: (i) These methods generate comparable results when very fine time steps are used. (ii) The characteristic DG schemes generate stable and accurate solutions even if a large time step (e.g., a Courant number ≥ 50) is used. In addition, we emphasize that the Runge-Kutta DG schemes are very flexible and can apply to virtually any applications. The characteristic DG schemes require a well defined velocity field to carry out a characteristic tracking and generate accurate solutions without requiring high-order regularity of the exact solutions. These methods are particularly suited for such applications as subsurface porous

medium flow and transport, in which the exact solutions do not typically have high-order regularity due to the heterogeneity of the porous media. On the other hand, the transport equation in porous medium flow and transport never exhibit shock discontinuities because the governing equations are upscaled on the macroscopic representative elementary volume scale that is larger than the microscopic continuum scale.

The developed characteristic discontinuous Galerkin methods can achieve high-order spatial accuracy (with high-order elements) but has only first-order temporal accuracy along the characteristics. A high-order temporal accuracy can be handled in a straightforward manner in the context of Eulerian methods, but it is delicate in the context of characteristic methods due to the application of a characteristic tracking. A second-order-in-time characteristic finite element method was developed in [2], a predict-correct approach can also be employed to develop a high-order-in-time schemes which will be investigated in the near future.

Even though the development and the analysis in this paper are for a simple transient linear advection-diffusion equation in one space dimension, our goal is to develop and analyze these methods for more realistic problems in applications. One important area of applications is the coupled system of partial differential equations arising in porous medium flow such as petroleum reservoir simulation and environmental modeling. In this context, the velocity in the transient advection-diffusion equation is determined from an associated pressure equation that is derived from the continuity equation for the fluid mixture and Darcy's law. On the other hand, the viscosity in the pressure equation depends on the concentration that is the solution to the transient advection-diffusion equation for the concentration [5]. Different characteristic finite difference or finite element methods were developed for the simulation of these systems [1, 13] and were analyzed [14]. A characteristic discontinuous Galerkin method is desired partly due to the following reasons: (i) It could resolve the moving steep fronts between different fluids more effectively; (ii) it is locally mass conservative; and (iii) it can handle physical interfaces between different subdomains with different permeabilities more effectively. This will be investigated in the near future.

Acknowledgments

The authors express their sincere thanks to the referees for their very helpful comments and suggestions, which greatly improved the quality of this paper. This work is supported by National Natural Science Foundation of China No. 10771124 and the research Fund for Doctoral Program of High Education by Station Education Ministry of China No. 20060422006.

Appendix: Auxiliary estimates

In this appendix we prove auxiliary estimates on η that were used in the main error analysis in Section 4.

Lemma A.1. *Under the assumption of Theorem* 5.1*, the estimate* (5.6) *holds.*

Proof. Let $\mathrm{Cr}=\max_{(x,t)\in[a,b]\times[t_{n-1},t_n]}|V(x,t)|\Delta t/h$ be the Courant number. We first consider the case $\mathrm{Cr}\geq 1/2$, which implies $h\leq C\Delta t$. We use (4.1) to bound the left side of (5.6) by

$$\begin{aligned}&\Big|\int_a^b \eta(x,t_{n-1})\xi(x,t_{n-1}^+)\,dx-\int_a^b \eta(x,t_n)\xi(x,t_n)\,dx\Big|\\ &\leq \|\eta(x,t_{n-1})\|\,\|\xi(x,t_{n-1}^+)\|_{L^2(a,b)}+\|\eta(x,t_n)\|\,\|\xi(x,t_n)\|\\ &\leq C\Delta t\|\xi(x,t_n)\|^2+C\Delta t\|\xi(b,t)\|^2_{L^2_V(t_{n-1},t_n)}+C(\Delta t)^3\|u\|^2_{L^\infty(0,T;H^2)}.\end{aligned}\tag{A.1}$$

In the case $\mathrm{Cr}<1/2$, we decompose the left-hand side as follows

$$\begin{aligned}&\int_a^b \eta(x,t_{n-1})\xi(x,t_{n-1}^+)\,dx-\int_a^b \eta(x,t_n)\xi(x,t_n)dx\\ =&\int_a^{b^*} \eta(x,t_{n-1})(\xi(\widetilde{x},t_n)-\xi(x,t_n))dx+\int_{b^*}^b \eta(x,t_{n-1})\xi(x,t_{n-1}^+)dx\\ &-\int_{b^*}^b \eta(x,t_{n-1})\xi(x,t_n)dx-\int_a^b\int_{t_{n-1}}^{t_n}\eta_t(x,t)dt\xi(x,t_n)dx.\end{aligned}\tag{A.2}$$

The last term on the right-hand side is bounded in a standard manner

$$\begin{aligned}&\Big|\int_a^b\int_{t_{n-1}}^{t_n}\eta_t(x,t)dt\xi(x,t_n)dx\Big|\\ &\leq(\Delta t)^{1/2}\|\eta_t\|_{L^2(t_{n-1},t_n;L^2)}\|\xi(x,t_n)\|\\ &\leq C\Delta t\|\xi(x,t_n)\|^2+Ch^{2m+2}\|u\|^2_{H^1(t_{n-1},t_n;H^{m+1})}.\end{aligned}\tag{A.3}$$

The first term on the right-hand side of (A.2) requires careful analysis. As the function ξ exhibits discontinuity in space, one cannot simply integrate this term by parts as in the analysis for the ELLAM type of schemes [25,27,28]. Instead, we decompose this term as two sums of terms such that $\xi(\widetilde{x},t_n)$ and $\xi(x,t_n)$ are continuous within each integral

$$\begin{aligned}&\int_a^{b^*}\eta(x,t_{n-1})(\xi(\widetilde{x},t_n)-\xi(x,t_n))dx\\ =&\sum_{i=1}^I\int_{x_{i-1}}^{x_i^*}\eta(x,t_{n-1})(\xi(\widetilde{x},t_n)-\xi(x,t_n))dx+\sum_{i=1}^{I-1}\int_{x_i^*}^{x_i}\eta(x,t_{n-1})(\xi(\widetilde{x},t_n)-\xi(x,t_n))dx.\end{aligned}\tag{A.4}$$

As $x\in[x_{i-1},x_i^*]$, $\widetilde{x}\in[x_{i-1},x_i]$ is in the same interval. The first term is bounded by

$$\begin{aligned}&\Big|\sum_{i=1}^I\int_{x_{i-1}}^{x_i^*}\eta(x,t_{n-1})(\xi(\widetilde{x},t_n)-\xi(x,t_n))dx\Big|\\ =&\Big|\sum_{i=1}^I\int_{x_{i-1}}^{x_i^*}\eta(x,t_{n-1})\int_x^{\widetilde{x}}\xi_y(y,t_n)dydx\Big|\\ \leq&\varepsilon\|\xi(x,t_n)\|_h^2+C\Delta th^{2m+2}\|u\|^2_{L^\infty(0,T;H^{m+1})}.\end{aligned}\tag{A.5}$$

For $x\in[x_i^*,x_i]$, $\widetilde{x}\in[x_i,x_{i+1}]$. Namely, there is one jump discontinuity between $\xi(x,t_n)$ and $\xi(\widetilde{x},t_n)$ in the second term on the right side of (A.4). Therefore, we rewrite this sum as

$$\begin{aligned}&\sum_{i=1}^{I-1}\int_{x_i^*}^{x_i}\eta(x,t_{n-1})(\xi(\widetilde{x},t_n)-\xi(x,t_n))dx\\&=\sum_{i=1}^{I-1}\int_{x_i^*}^{x_i}\eta(x,t_{n-1})\Big((\xi(\widetilde{x},t_n)-\xi(x_i^+,t_n))+[[\xi]](x_i,t_n)+(\xi(x_i^-,t_n)-\xi(x,t_n))\Big)dx.\end{aligned}\tag{A.6}$$

We use the fact $|x_i-x|=\mathcal{O}(\Delta t)$ for $x\in[x_i^*,x_i]$ and (4.3) to bound the jump term by

$$\begin{aligned}&\sum_{i=1}^{I-1}\int_{x_i^*}^{x_i}|\eta(x,t_{n-1})|dx\ \big|[[\xi]](x_i,t_n)\big|\\&\le C(\Delta t)^2\sum_{i=1}^{I-1}\int_{x_{i-1}}^{x_i}|\eta_{xx}(x,t_{n-1})|dx\ \big|[[\xi]](x_i,t_n)\big|\\&\le C\Delta t\|\xi(x,t_n)\|^2+C(\Delta t)^3\|u\|^2_{L^\infty(0,T;H^2)}.\end{aligned}\tag{A.7}$$

Here we have used the inverse inequality. On the other hand, we know

$$\begin{aligned}&\Big|\sum_{i=1}^{I-1}\int_{x_i^*}^{x_i}\eta(x,t_{n-1})((\xi(\widetilde{x},t_n)-\xi(x_i^+,t_n))+(\xi(x_i^-,t_n)-\xi(x,t_n)))dx\Big|\\&=\Big|\sum_{i=1}^{I-1}\int_{x_i^*}^{x_i}\eta(x,t_{n-1})\Big(\int_{x_i}^{\widetilde{x}}\xi_y(y,t_n)dy+\int_x^{x_i}\xi_y(y,t_n)dy\Big)dx\Big|\\&\le\varepsilon\|\xi(x,t_n)\|_h^2+C\Delta t h^{2m+2}\|u\|^2_{L^\infty(0,T;H^{m+1})}.\end{aligned}\tag{A.8}$$

We use (4.3) to bound the second and third terms on the right-hand side of (A.2) by

$$\begin{aligned}&\Big|\int_{b^*}^b\eta(x,t_{n-1})\xi(x,t_n)\ dx-\int_{b^*}^b\eta(x,t_{n-1})\xi(x,t_{n-1}^+)\ dx\Big|\\&=\Big|\int_{b^*}^b\int_x^b\int_y^b\eta_{zz}(z,t_{n-1})dzdy\big(\xi(x,t_n)-\xi(x,t_{n-1}^+)\big)dx\Big|\\&\le C\Delta t\|\xi(x,t_n)\|^2+C\Delta t\|\xi(b,t)\|^2_{L_V^2(t_{n-1},t_n)}+C(\Delta t)^3\|u\|^2_{L^\infty(0,T;H^2)}.\end{aligned}\tag{A.9}$$

We combine these estimates to conclude the proof. □

Lemma A.2. *Under the assumption of Theorem* 5.1*, the estimate* (5.7) *is valid.*

Proof. In the case $\mathrm{Cr}\ge 1/2$ we have

$$\begin{aligned}&\Big|\sum_{i=1}^{I}\int_{x_{i-1}}^{x_i}\Delta t(x)D(x,t_n)\eta_x(x,t_n)\xi_x(x,t_n)dx\Big|\\&\le\varepsilon\|\xi(x,t_n)\|_h^2+C\Delta t h^2\|u\|^2_{L^\infty(0,T;H^2)}\le\varepsilon\|\xi(x,t_n)\|_h^2+C(\Delta t)^3\|u\|^2_{L^\infty(0,T;H^2)}.\end{aligned}\tag{A.10}$$

When $\mathrm{Cr}<1/2$, $\widetilde{a}-a<h/2$. Thus, $\Delta t(x)=\Delta t$ on $[\tilde{a},x_1]$ and $[x_{i-1},x_i]$ for $i=2,3,\cdots,I$. Consequently,

$$\begin{aligned}&\sum_{i=1}^{I}\int_{x_{i-1}}^{x_i}\Delta t(x)D(x,t_n)\eta_x(x,t_n)\xi_x(x,t_n)dx\\=&\sum_{i=1}^{I}\int_{x_{i-1}}^{x_i}\Delta tD(x,t_n)\eta_x(x,t_n)\xi_x(x,t_n)dx\\&-\int_a^{\tilde{a}}(t_n^*(x)-t_{n-1})D(x,t_n)\eta_x(x,t_n)\xi_x(x,t_n)dx.\end{aligned}\tag{A.11}$$

We utilize the superconvergence (4.2) to bound the first term on the right-hand side by

$$\begin{aligned}&\left|\sum_{i=1}^{I}\int_{x_{i-1}}^{x_i}\Delta tD(x,t_n)\eta_x(x,t_n)\xi_x(x,t_n)dx\right|\\\le&C\Delta th^{m+1}\|u\|^2_{L^\infty(0,T;H^{m+2})}\Big[\sum_{i=1}^{I}\Big(\int_{x_{i-1}}^{x_i}\xi_x^2(x,t_n)dx\Big)^{\frac{1}{2}}+\|\xi(x,t_n)\|\Big]\\\le&\varepsilon\|\xi(x,t_n)\|_h^2+C\Delta t\|\xi(x,t_n)\|^2+C\Delta th^{2m+2}\|u\|^2_{L^\infty(0,T;H^{m+2})}.\end{aligned}\tag{A.12}$$

We bound the second term on the right-hand side of (A.11) by

$$\begin{aligned}&\Big|\int_a^{\widetilde{a}}(t_n^*(x)-t_{n-1})D(x,t_n)\eta_x(x,t_n)\xi_x(x,t_n)dx\Big|\\=&\Big|\int_a^{\widetilde{a}}(t_n^*(x)-t_{n-1})D(x,t_n)\int_{\hat{x}_{1/2}}^{x}\eta_{yy}(y,t_n)dy\xi_x(x,t_n)dx\Big|\\\le&\varepsilon\|\xi(x,t_n)\|_h^2+C(\Delta t)^3\|u\|^2_{L^\infty(0,T;H^2)}.\end{aligned}\tag{A.13}$$

Combining these results concludes the proof. □

Lemma A.3. *Under the assumptions of Theorem* 5.1*, the estimate* (5.8) *is valid.*

Proof. Eqs. (2.7), (2.8), (2.10), and the fact that ξ is constant in $\Omega^{(o)}$ concludes that $E(u,\xi)$ has the form

$$\begin{aligned}E(u,\xi)=&-\sum_{i=1}^{I}[[\xi]](x_i,t_n)\int_{t_{n,i}^*}^{t_n}u(r(t;x_i,t_n),t)(V(x_i,t_n)-V(r(t;x_i,t_n),t))dt\\&+\sum_{i=1}^{I}[[\xi]](x_i,t_n)\int_{t_{n,i}^*}^{t_n}\int_t^{t_n}\frac{d}{d\theta}(Du_r)(r(\theta;x_i,t_n),\theta)d\theta dt\\&-\sum_{i=1}^{I}\int_{x_{i-1}}^{x_i}\xi_x(x,t_n)\int_{t_n^*(x)}^{t_n}\int_t^{t_n}\frac{d}{d\theta}(Du_r)(r(\theta;x,t_n),\theta)d\theta dtdx\\&+\sum_{i=1}^{I}\int_{x_{i-1}}^{x_i}\int_{t_n^*(x)}^{t_n}u(\xi_t+V\xi_r)(r(t;x,t_n),t)r_x(t;x,t_n)dtdx-\sum_{i=1}^{I}E_i^{(f)}(\xi)-E_{(o)}^{(f)}(\xi).\end{aligned}\tag{A.14}$$

The first term on the right-hand side of (A.14) is bounded by

$$\begin{aligned}&\Big|\sum_{i=1}^{I}[[\xi]](x_i,t_n)\int_{t_{n,i}^*}^{t_n}u(r(t;x_i,t_n),t)\Big(V(r(t;x_i,t_n),t)-V(x_i,t_n)\Big)dt\Big|\\&\leq C\sum_{i=1}^{I}\Big|[[\xi]](x_i,t_n)\Big|(\Delta t(x_i))^{3/2}\Big(\int_{t_{n,i}^*}^{t_n}u^2(r(t;x_i,t_n),t)dt\Big)^{1/2}\\&\leq\varepsilon\sum_{i=1}^{I}\alpha_i[[\xi]]^2(x_i,t_n)+C(\Delta t)^2\|u\|^2_{L^2(t_{n-1},t_n;L^\infty)}.\end{aligned}\tag{A.15}$$

The second term on the right-hand side of (A.14) is bounded by

$$\begin{aligned}&\Big|\sum_{i=1}^{I}[[\xi]](x_i,t_n)\int_{t_{n,i}^*}^{t_n}\int_{t}^{t_n}\frac{d}{d\theta}(Du_r)(r(\theta;x_i,t_n),\theta)d\theta dt\Big|\\&\leq\varepsilon\sum_{i=1}^{I}\alpha_i[[\xi]]^2(x_i,t_n)+C(\Delta t)^2\Big(\Big\|\frac{du}{dt}\Big\|^2_{L^2(t_{n-1},t_n;W^1_\infty)}+\|u\|^2_{L^2(t_{n-1},t_n;W^1_\infty)}\Big).\end{aligned}\tag{A.16}$$

The third term on the right-hand side of (A.14) is bounded by

$$\begin{aligned}&\Big|\sum_{i=1}^{I}\int_{x_{i-1}}^{x_i}\xi_x(x,t_n)\int_{t_n^*(x)}^{t_n}\int_{t}^{t_n}\frac{d}{d\theta}(Du_r)(r(\theta;x,t_n),\theta)d\theta dtdx\Big|\\&\leq\varepsilon\|\xi(x,t_n)\|_h^2+C(\Delta t)^2\Big(\Big\|\frac{du}{dt}\Big\|^2_{L^2(t_{n-1},t_n;H^1)}+\|u\|^2_{L^2(t_{n-1},t_n;H^1)}\Big).\end{aligned}\tag{A.17}$$

The fifth and sixth terms on the right side of (A.14) are bounded by

$$\begin{aligned}\Big|\sum_{i=1}^{I}E_i^{(f)}(w)+E_{(o)}^{(f)}(w)\Big|&\leq C\Delta t\|\xi(x,t_n)\|^2+C\Delta t\|\xi(b,t)\|^2_{L_V^2(t_{n-1},t_n)}\\&\quad+C(\Delta t)^2\Big(\|f\|^2_{L^2(t_{n-1},t_n;L^2)}+\Big\|\frac{df}{dt}\Big\|^2_{L^2(t_{n-1},t_n;L^2)}\Big).\end{aligned}\tag{A.18}$$

As for the fourth term on the right-hand side of Eq. (A.14), we note that the test function satisfies the approximate adjoint equation

$$\xi_t(r(t;x,t_n),t)+V(x,t_n)\xi_r(r(t;x,t_n),t)=0.$$

We subtract this equation from the adjoint term to bound this term as follows:

$$\begin{aligned}&\Big|\sum_{i=1}^{I}\int_{x_{i-1}}^{x_i}\int_{t_n^*(x)}^{t_n}u\Big(\xi_t+V\xi_r)(r(t;x,t_n),t)r_x(t;x,t_n)dtdx\Big|\\&=\Big|\sum_{i=1}^{I}\int_{x_{i-1}}^{x_i}\int_{t_n^*(x)}^{t_n}u(V(r(t;x,t_n),t)-V(x,t_n))\xi_x(x,t_n)r_x(t;x,t_n)dtdx\Big|\\&\leq\varepsilon\|\xi(x,t_n)\|_h^2+C(\Delta t)^2\|u\|^2_{L^2(t_{n-1},t_n;L^2)}.\end{aligned}\tag{A.19}$$

This completes the proof. □

References

[1] M. Al-Lawatia, K. Wang, A. S. Telyakovskiy and H. Wang, An Eulerian-Lagrangian method for porous medium flow, Int. J. Comput. Sci. Math., 1 (2007), 467-479.

[2] M. Al-Lawatia, R. C. Sharpley and H. Wang, Second-order characteristics methods for advection-diffusion equations and comparison to other schemes, Adv. Water Resources, 22 (1999), 741-768.

[3] D. N. Arnold, An Interior Penalty Finite Element Method with Discontinuous Elements, Ph.D. Thesis, The University of Chicage, Chicago, IL, 1979.

[4] D. N. Arnold, F. Brezzi, B. Cockburn and L. D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, SIAM J. Numer. Anal., 39(5) (2002), 1749-1779.

[5] J. Bear, Dynamics of Fluids in Porous Materials, Elsevier, New York, 1972.

[6] J. P. Boris and D. L. Book, Flux-corrected transport. I. SHASTA, A fluid transport algorithm that works, J. Comput. Phy., 135 (1997), 172-186.

[7] C. E. Baumann and J. T. Oden, Discontinuous hp finite element method for convection-diffusion problems, Comput. Meth. Appl. Mech. Engrg., 175 (1999), 311-341.

[8] M. A. Celia, T. F. Russell, I. Herrera, and R. E. Ewing, An Eulerian-Lagrangian localized adjoint method for the advection-diffusion equation, Adv. Water Resources, 13 (1990), 187-206.

[9] B. Cockburn, G. Karniadakis and C.-W. Shu (Eds.), Discontinuous Galerkin Methods: Theory, Computation, and Applications, Lecture Notes in Computational Science and Engineering, 11, Springer Verlag, Berlin, 2000.

[10] B. Cockburn and C.-W. Shu, The local discontinuous Galerkin method for time-dependent convection-diffusion systems, SIAM J. Numer. Anal., 35(6) (1998), 2440-2463.

[11] B. Cockburn and C.-W. Shu, Runge-Kutta discontinuous Galerkin methods for convection-dominated problems, J. Sci. Comput., 16 (2001), 173-261.

[12] C. Dawson, S. Sun and M. F. Wheeler, Compatible algorithms for coupled flow and transport, Comput. Meth. Appl. Mech. Engrg., 193 (2004), 2565-2580.

[13] R. E. Ewing, T. F. Russell and M. F. Wheeler, Simulation of miscible displacement using mixed methods and a modified method of characteristics, SPE 12241, 7th SPE Symposium on Reservoir Simulation, (1983), 71-81.

[14] R. E. Ewing, T. F. Russell and M. F. Wheeler, Convergence analysis of an approximation of miscible displacement in porous media by mixed finite elements and a modified method of characteristics, Comput. Meth. Appl. Mech. Engrg., 47 (1984), 73-92.

[15] R. E. Ewing and H. Wang, An optimal-order error estimate to Eulerian-Lagrangian localized adjoint method for variable-coefficient advection-reaction problems, SIAM J. Numer. Anal., 33 (1996), 318-348.

[16] M.-G. Larson and A. J. Niklasson, Analysis of a family of discontinuous galerkin methods for elliptic problems: the one dimensional case, Numer. Math., 99(1) (2004), 113-130.

[17] J. T. Oden, I. Babuška and C. E. Baumann, A discontinuous hp finite element method for diffusion problems, J. Comput. Phys., 146 (1998), 491-519.

[18] W. H. Reed and T. R. Hill, Triangular mesh methods for the neutron transport equation, Technical report, Los Alamos Scientific Laboratory, 1973.

[19] B. Rivière, M. F. Wheeler and V. Girault, A priori error estimates for finite element methods based on discontinuous approximation spaces for elliptic problems, SIAM J. Numer. Anal., 39(3) (2001), 902-931.

[20] B. Rivière and M. F. Wheeler, Non conforming methods for transport with nonlinear reac-

tion, Contemporary Mathematics, 295 (2002), 421-432.
[21] C.-W. Shu, Essentially Non-oscillatory and Weighted Essentially Non-oscillatory Schemes for Hyperbolic Conservation Laws, in: B. Cockburn, C. Johnson, C.-W. Shu and E. Tadmor (Eds.), Advanced Numerical Approximation of Nonlinear Hyperbolic Equations, volume 1687 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, Heidelberg, New York, 1998, pp. 325-432.
[22] S. Sun, Discontinuous Galerkin Methods for Reactive Transport in Porous Media, Ph.D. Thesis, The University of Texas at Austin, 2003.
[23] S. Sun and M. F. Wheeler, Symmetric and nonsymmetric discontinuous Galerkin methods for reactive transport in porous media, SIAM J. Numer. Anal., 43(1) (2005), 195-219.
[24] H. Wang, H. K. Dahle, R. E. Ewing, M. S. Espedal, R. C. Sharpley and S. Man, An ELLAM scheme for advection-diffusion equations in two dimensions, SIAM J. Sci. Comput., 20 (1999), 2160-2194.
[25] H. Wang, R. E. Ewing and T. F. Russell, Eulerian-Lagrangian localized methods for convection-diffusion equations and their convergence analysis, IMA J. Numer. Anal., 15 (1995), 405-459.
[26] H. Wang, D. Liang, R. E. Ewing, S. L. Lyons and G. Gin, An approximation to miscible fluid flows in porous media with point sources and sinks by an Eulerian-Lagrangian localized adjoint method and mixed finite element methods, SIAM J. Sci. Comput., 22 (2000), 561-581.
[27] H. Wang and K. Wang, Uniform estimates for Eulerian-Lagrangian methods for singularly perturbed time-dependent problems, SIAM J. Numer. Anal., 45 (2007), 1305-1329.
[28] K. Wang, A uniformly optimal-order error estimate of an ELLAM scheme for unsteady-state advection-diffusion equations, Int. J. Numer. Anal. Model., 5 (2008), 286-302.
[29] L. B. Wahlbin, Superconvergence in Galerkin Finite Element Methods, Springer Lecture Notes in Mathematics 1605, Springer-Verlag, New York, 1995.
[30] M. F. Wheeler, An elliptic collocation-finite element method with interior penalties, SIAM J. Numer. Anal., 15 (1978), 152-161.
[31] N. Yan, Superconvergence Analysis and a Posteriori Error Estimation in Finite Element Methods, Science Press, Beijing, China, 2008.
[32] Z. Zhang, Finite element superconvergent approximation for one-dimensional singularly perturbed problems with graph, Numer. Methods Part. Diff. Eq., 18 (2002), 374-395.
[33] Q. Zhu and Q. Lin, Superconvergence Theory of Finite Element Methods, Hunan Science Press, China, 1989.