Third order implicit-explicit Runge-Kutta local discontinuous Galerkin methods with suitable boundary treatment for convection-diffusion problems with Dirichlet boundary conditions

Haijin Wang[†]   Qiang Zhang [‡]   Chi-Wang Shu[§]

April 8, 2018

Abstract

To avoid the order reduction when third order implicit-explicit Runge-Kutta time discretization is used together with the local discontinuous Galerkin (LDG) spatial discretization, for solving convection-diffusion problems with time-dependent Dirichlet boundary conditions, we propose a strategy of boundary treatment at each intermediate stage in this paper. The proposed strategy can achieve optimal order of accuracy by numerical verification. Also by suitably setting numerical flux on the boundary in the LDG methods, and by establishing an important relationship between the gradient and interface jump of the numerical solution with the independent numerical solution of the gradient and the given boundary conditions, we build up the unconditional stability of the corresponding scheme, in the sense that the time step is only required to be upper bounded by a suitable positive constant, which is independent of the mesh size.

keywords. local discontinuous Galerkin method, implicit-explicit time discretization, convection-diffusion equation, Dirichlet boundary condition, order reduction.

AMS. 65M12, 65M15, 65M60

1   Introduction

The local discontinuous Galerkin (LDG) method was introduced by Cockburn and Shu [10], motivated by the work of Bassi and Rebay [4] for solving the compressible Navier-Stokes equations. It was designed for solving convection-diffusion problems initially, and later it gained wide applications in many high order partial differential equations, for example, KdV-type equations [27], bi-harmonic equations [28, 11], the fifth order dispersion equation

[28], and so on. More applications of the LDG schemes can be found in the review article [26] and the reference therein.

With respect to the theoretical studies about LDG schemes, most of them pay attention to model problems with periodic boundary conditions (BCs), as far as the authors know. Since many applications in practice are non-periodic, it is important to study LDG schemes in non-periodic situations. For convection-diffusion problems with Dirichlet BCs, optimal error estimate of the semi-discrete LDG method has been studied in [8]. Wang and Zhang [24] presented an optimal error estimate for a fully-discrete LDG scheme, where the third order total variation diminishing (TVD) explicit Runge-Kutta (RK) method [19, 13] was adopted in time-discretization, and suitable boundary treatment at each intermediate stage of explicit third order RK time marching was proposed.

As for convection-diffusion problems, explicit RK methods are stable and efficient for solving convection-dominated problems. However, for problems which are not convection-dominated, explicit time discretization will suffer from a stringent time step restriction for stability [25]. To overcome the small time step restriction, we considered a type of implicit-explicit (IMEX) RK schemes [3] in [21, 22], where the convection and diffusion parts were treated explicitly and implicitly, respectively. The corresponding IMEX-LDG schemes were shown to be unconditionally stable for periodic BCs.

In this paper, we consider the third order IMEX RK time-discretization [5] coupled with LDG spatial discretization for one-dimensional convection-diffusion problems with time-dependent Dirichlet BCs. Stability as well as error estimates will be investigated. The main difficulties come from two aspects, one is about how to set numerical flux on element interfaces, the other is about how to avoid the order reduction due to improper boundary treatment at each intermediate stage of high order (≥ 3) IMEX RK schemes.

The first difficulty has been overcome by Castillo *et al.* in [8], where a suitable numerical flux was defined to ensure the stability and optimal error estimates of the semi-discrete LDG scheme. In this paper we will adopt a similar numerical flux as in [8], based on which we establish an important relationship between the gradient and interface jump of the numerical solution with the independent numerical solution of the gradient and the given boundary conditions, which plays a key role in stability and error analysis.

To put the second difficulty in proper perspective, let us briefly describe the background of order reduction. It occurs when a RK method is used together with the method of lines for the fully discretization of an initial boundary value problem [12, 18, 20, 14]. To recover the full order of accuracy, researchers have proposed several strategies of boundary corrections for explicit and implicit RK methods. Some representative works for explicit RK methods are in [7, 1, 17, 6, 29], and a representative work for implicit RK methods is [2]. The ideas of boundary remedies for explicit and implicit methods are essentially the same, i.e, to examine the truncation errors made by the s-th order RK method when no intermediate-stage boundary conditions are enforced, and then to mimic these errors to at least s-th order when prescribing boundary conditions [17].

However, there has been no such work on boundary corrections for IMEX RK methods, to the best of our knowledge. The main difficulty lies in that, the time discretization for convection and diffusion are different, the conversion between spatial and temporal derivatives are not trivial, compared with the fully explicit or implicit situations, so it is rather difficult to derive consistent intermediate boundary conditions [7] solely from

the physical boundary condition and its time derivatives. Similar difficulty occurs when using inverse Lax-Wendroff procedure for numerical boundary conditions of convection-diffusion equations [15]. Nevertheless, the ideas of boundary remedies for explicit and implicit methods can be adopted here.

In this paper, we will first study boundary remedies $g_{\text{im}}^{n,\ell}$ for the purely third order implicit scheme (which is the implicit part of the third order IMEX RK method [5]) for the diffusion model, where the boundary remedies only contain information of physical boundary condition and its time derivatives. Then based on $g_{\text{im}}^{n,\ell}$, we propose a strategy of boundary corrections for the third order IMEX RK method [5] for the convection-diffusion model, where extra terms involving high order spatial derivatives need to be approximated properly, thus, certain temporal-spatial constraint is required to observe optimal error estimates in time. Even though we are not able to prove the optimal error estimates in time for the proposed scheme, it behaves very well in numerical experiments.

The remaining of this paper is organized as follows. In Section 2 we present the semi-discrete and fully-discrete LDG scheme for the model problem. A strategy of boundary treatment and the corresponding numerical results are given in Section 3. Sections 4 is about the stability and error estimates for the corresponding scheme. Finally, we give concluding remarks in Section 5.

2 The LDG method

2.1 The semi-discrete LDG scheme

In this subsection we present the definition of semi-discrete LDG schemes for the linear convection-diffusion problem with time-dependent Dirichlet boundary condition

$$U_t + cU_x = dU_{xx}, \qquad\qquad (x,t) \in Q_T = (a,b) \times (0,T], \qquad (2.1a)$$

$$U(x,0) = U_0(x), \qquad\qquad x \in \Omega = (a,b), \qquad\qquad (2.1b)$$

$$U(a,t) = U_a(t), \quad U(b,t) = U_b(t), \qquad t \in (0,T]. \qquad\qquad (2.1c)$$

Here the constants c and $d > 0$ are convection and diffusion coefficients, respectively. Without loss of generality, we assume $c > 0$ in this paper. The initial solution $U_0(x)$ is assumed to be in $L^2(\Omega)$.

Let $Q = \sqrt{d}U_x$ and define $(h_U^c, h_U^d, h_Q) := (cU, -\sqrt{d}Q, -\sqrt{d}U)$. The LDG scheme starts from the following equivalent first-order differential system

$$U_t + (h_U^c)_x + (h_U^d)_x = 0, \quad Q + (h_Q)_x = 0, \qquad (x,t) \in Q_T, \qquad (2.2)$$

with the same initial condition (2.1b) and boundary condition (2.1c).

Let $\mathcal{T}_h = \{I_j = (x_{j-1}, x_j)\}_{j=1}^N$ be the partition of Ω, where $x_0 = a$ and $x_N = b$ are the two boundary endpoints. Denote the cell length as $h_j = x_j - x_{j-1}$ for $j = 1, \ldots, N$, and define $h = \max_j h_j$. We assume $\mathcal{T}_h$ is quasi-uniform in this paper, that is, there exists a positive constant ρ such that for all j there holds $h_j/h \geq \rho$, as h goes to zero.

Associated with this mesh, we define the discontinuous finite element space

$$V_h = \left\{ v \in L^2(\Omega) : v|_{I_j} \in \mathcal{P}_k(I_j), \, \forall j = 1, \ldots, N \right\}, \qquad (2.3)$$

where $\mathcal{P}_k(I_j)$ denotes the space of polynomials in I_j of degree at most k. Note that the functions in this space are allowed to have discontinuities across element interfaces. At each element interface point, for any piecewise function p, there are two traces along the right-hand and left-hand, denoted by p^+ and p^-, respectively. The jump is denoted by $[\![p]\!] = p^+ - p^-$.

The semi-discrete LDG scheme is defined as follows: for any $t > 0$, find the numerical solution $\boldsymbol{w}(t) := (u(t), q(t)) \in V_h \times V_h$ (where the argument x is omitted), such that

$$(u_t, v)_j = \mathcal{H}_j(u, v) + \mathcal{L}_j(\boldsymbol{w}, v), \tag{2.4a}$$

$$(q, r)_j = \mathcal{K}_j(u, r), \tag{2.4b}$$

hold in each cell I_j, $j = 1, \ldots, N$, for any test functions $\boldsymbol{z} = (v, r) \in V_h \times V_h$. Here $(\cdot, \cdot)_j$ is the inner product in $L^2(I_j)$ and

$$\mathcal{H}_j(u, v) = (h_u^c, v_x)_j - (\hat{h}_u^c)_j v_j^- + (\hat{h}_u^c)_{j-1} v_{j-1}^+, \tag{2.5a}$$

$$\mathcal{L}_j(\boldsymbol{w}, v) = (h_u^d, v_x)_j - (\hat{h}_u^d)_j v_j^- + (\hat{h}_u^d)_{j-1} v_{j-1}^+, \tag{2.5b}$$

$$\mathcal{K}_j(u, r) = (h_q, r_x)_j - (\hat{h}_q)_j r_j^- + (\hat{h}_q)_{j-1} r_{j-1}^+, \tag{2.5c}$$

where $\hat{h}_u^c, \hat{h}_u^d$ and $\hat{h}_q$ are numerical flux defined at every element boundary point. Let $g_a = U_a$ and $g_b = U_b$ be given Dirichlet boundary conditions, we would like to define the numerical flux in the similar way as that in [8, 24], which is listed in Table 1.

Table 1: The definition of the numerical flux at each element boundary point.

	$j = 0$	$j = 1, \cdots, N-1$	$j = N$
$\hat{h}_u^c$	cg_a	cu_j^-	cu_N^-
$\hat{h}_u^d$	$-\sqrt{d}q_0^+$	$-\sqrt{d}q_j^+$	$-\sqrt{d}\left[q_N^- - \frac{\sqrt{d}}{h}(u_N^- - g_b)\right]$
$\hat{h}_q$	$-\sqrt{d}g_a$	$-\sqrt{d}u_j^-$	$-\sqrt{d}g_b$

Remark 2.1. In Table 1, the standard "upwinding" numerical flux was taken for the discretization of convection, which was used in the original DG method. On the interior element interfaces, the "alternating " numerical flux was adopted for the discretization of diffusion, this simple setting works well in the LDG method for periodic BCs. For Dirichlet BCs, we only make some modifications on the boundary points. Specifically, with regard to $\hat{h}_q$ at $j = N$, we use $-\sqrt{d}g_b$ instead of $-\sqrt{d}u_N^-$ to ensure the stability of the scheme, since the interior elements should communicate with boundary; with regard to $\hat{h}_u^d$, we define $q_N^+ = q_N^-$, the penalty term $\frac{d}{h}(u_N^- - g_b)$ is used to enhance the stability and guarantee the optimal accuracy of the scheme. For more details of the numerical fluxes, we refer the readers to [8], and the detailed description given in [24].

In (2.4) and below, we drop the argument t if there is no confusion. The initial condition $u(x, 0)$ can be taken as any approximation of the given initial solution $U_0(x)$, for example,

the local Gauss-Radau projection of $U_0(x)$. Please refer to (4.24) for more details. We have now defined the semi-discrete LDG scheme.

To write the above scheme in compact form, we denote by

$$(v, w) = \sum_{j=1}^{N} (v, w)_j$$

the inner product in $L^2(\Omega)$, and by

$$\langle v, w \rangle = \sum_{j=1}^{N-1} v_j w_j$$

the L^2-inner product on all interior mesh grids, and denote $\boldsymbol{g} = (g_a, g_b)$. Summing up the variational formulations (2.4) over $j = 1, 2, \ldots, N$, we can get the semi-discrete LDG scheme in global form: for any $t > 0$, find the numerical solution $\boldsymbol{w} = (u, q) \in V_h \times V_h$, such that

$$(u_t, v) = \mathcal{H}(\boldsymbol{g}; u, v) + \mathcal{L}(\boldsymbol{g}; \boldsymbol{w}, v), \tag{2.6a}$$

$$(q, r) = \mathcal{K}(\boldsymbol{g}; u, r), \tag{2.6b}$$

hold for any $\boldsymbol{z} = (v, r) \in V_h \times V_h$. Here $\Xi = \sum_{j=1}^{N} \Xi_j$ for $\Xi = \mathcal{H}, \mathcal{L}, \mathcal{K}$. For convenience of further analysis, we would like to divide these operators into two parts, i.e, the interior part and the boundary part, to be specific

$$\mathcal{H}(\boldsymbol{g}; u, v) = \mathcal{H}_{\text{int}}(u, v) + \mathcal{H}_{\text{bry}}(\boldsymbol{g}; v),$$

$$\mathcal{L}(\boldsymbol{g}; \boldsymbol{w}, v) = \mathcal{L}_{\text{int}}(\boldsymbol{w}, v) + \mathcal{L}_{\text{bry}}(\boldsymbol{g}; v),$$

$$\mathcal{K}(\boldsymbol{g}; u, r) = \mathcal{K}_{\text{int}}(u, r) + \mathcal{K}_{\text{bry}}(\boldsymbol{g}; r),$$

where

$$\mathcal{H}_{\text{int}}(u, v) = c\left[(u, v_x) + \langle u^-, [\![v]\!]\rangle - u_N^- v_N^-\right] = c\left[-(u_x, v) - \langle [\![u]\!], v^+\rangle - u_0^+ v_0^+\right], \tag{2.7a}$$

$$\mathcal{L}_{\text{int}}(\boldsymbol{w}, v) = -\sqrt{d}\left[(q, v_x) + \langle q^+, [\![v]\!]\rangle - q_N^- v_N^- + q_0^+ v_0^+\right] - \frac{d}{h} u_N^- v_N^-, \tag{2.7b}$$

$$\mathcal{K}_{\text{int}}(u, r) = -\sqrt{d}\left[(u, r_x) + \langle u^-, [\![r]\!]\rangle\right]; \tag{2.7c}$$

and

$$\mathcal{H}_{\text{bry}}(\boldsymbol{g}; v) = c g_a v_0^+, \quad \mathcal{L}_{\text{bry}}(\boldsymbol{g}; v) = \frac{d}{h} g_b v_N^-, \quad \mathcal{K}_{\text{bry}}(\boldsymbol{g}; r) = \sqrt{d}(g_b r_N^- - g_a r_0^+). \tag{2.7d}$$

2.2   Properties of the LDG spatial discretization

In this subsection, we give some properties of the LDG scheme (2.6). Let us first introduce some notations. We define

$$[\![v]\!]_{\text{int}}^2 = \sum_{j=1}^{N-1} [\![v]\!]_j^2, \quad \text{and} \quad [\![v]\!]^2 = [\![v]\!]_{\text{int}}^2 + (v_N^-)^2 + (v_0^+)^2, \tag{2.8}$$

for arbitrary v belonging to the (mesh-dependent) broken Sobolev space

$$H^1(\mathcal{T}_h) = \left\{ \phi \in L^2(\Omega) : \phi|_{I_j} \in H^1(I_j),\ \forall j = 1, \ldots, N \right\}.$$

For any function $v \in V_h$, we have the inverse inequality

$$\|v\|_{\partial I_j} \leq \sqrt{\mu h^{-1}} \|v\|_{I_j}, \qquad \forall j = 1, 2, \cdots, N, \tag{2.9}$$

where $\|v\|_{\partial I_j} = \sqrt{(v_{j-1}^+)^2 + (v_j^-)^2}$ is the L^2-norm on the boundary of I_j, $\|v\|_{I_j}$ is the L^2-norm in I_j, and $\mu > 0$ is the inverse constant which is independent of v, h and j.

Next we present the following properties of LDG spatial discretization.

Lemma 2.1. *For any $w, v_2 \in H^1(\mathcal{T}_h)$ and $\boldsymbol{z}_1 = (v_1, r_1) \in H^1(\mathcal{T}_h) \times H^1(\mathcal{T}_h)$, there hold the following equalities*

$$\mathcal{H}_{\mathrm{int}}(w, w) = -\frac{c}{2} [\![w]\!]^2, \tag{2.10}$$

$$\mathcal{L}_{\mathrm{int}}(\boldsymbol{z}_1, v_2) + \mathcal{K}_{\mathrm{int}}(v_2, r_1) = -\frac{d}{h} v_{1,N}^- v_{2,N}^-. \tag{2.11}$$

Proof. From (2.7a), we get

$$\begin{aligned}
\mathcal{H}_{\mathrm{int}}(w, w) &= c \left[(w, w_x) + \langle w^-, [\![w]\!] \rangle - (w_N^-)^2 \right] \\
&= \frac{c}{2} \left[(w_N^-)^2 - \langle [\![w^2]\!], 1 \rangle - (w_0^+)^2 \right] + c \langle w^-, [\![w]\!] \rangle - c(w_N^-)^2 \\
&= -\frac{c}{2} \left[(w_N^-)^2 + \langle [\![w]\!], [\![w]\!] \rangle + (w_0^+)^2 \right] = -\frac{c}{2} [\![w]\!]^2.
\end{aligned} \tag{2.12}$$

From (2.7b) and (2.7c), we get

$$\begin{aligned}
&\mathcal{L}_{\mathrm{int}}(\boldsymbol{z}_1, v_2) + \mathcal{K}_{\mathrm{int}}(v_2, r_1) \\
&= -\sqrt{d} \left[(r_1, (v_2)_x) + \langle (r_1)^+, [\![v_2]\!] \rangle - r_{1,N}^- v_{2,N}^- + r_{1,0}^+ v_{2,0}^+ \right] \\
&\quad - \sqrt{d} \left[(v_2, (r_1)_x) + \langle (v_2)^-, [\![r_1]\!] \rangle \right] - \frac{d}{h} v_{1,N}^- v_{2,N}^- \\
&= -\frac{d}{h} v_{1,N}^- v_{2,N}^-,
\end{aligned} \tag{2.13}$$

by integrating by parts. $\qquad\square$

Corollary 2.1. *For any $\boldsymbol{z} = (v, r) \in H^1(\mathcal{T}_h) \times H^1(\mathcal{T}_h)$, we have*

$$\mathcal{L}_{\mathrm{int}}(\boldsymbol{z}, v) + \mathcal{K}_{\mathrm{int}}(v, r) = -\frac{d}{h} (v_N^-)^2. \tag{2.14}$$

By simply using the Cauchy-Schwarz inequality and the inverse inequality (2.9), we can directly obtain the following two lemmas whose proofs are trivial, so we omit the details to save space.

Lemma 2.2. *For any $w, v \in V_h$, there hold the following inequalities*

$$|\mathcal{H}_{\mathrm{int}}(w, v)| \leq c \left(\|w_x\| + \sqrt{\mu h^{-1}} ([\![w]\!]_{\mathrm{int}} + |w_0^+|) \right) \|v\|, \tag{2.15a}$$

$$|\mathcal{H}_{\mathrm{int}}(w, v)| \leq c \left(\|v_x\| + \sqrt{\mu h^{-1}} ([\![v]\!]_{\mathrm{int}} + |v_N^-|) \right) \|w\|. \tag{2.15b}$$

Lemma 2.3. *For any $v, r \in V_h$, we have*

$$|\mathcal{H}_{\text{bry}}(\boldsymbol{g}; v)| \leq c\sqrt{\mu h^{-1}}|g_a|\|v\|, \tag{2.16}$$

$$|\mathcal{L}_{\text{bry}}(\boldsymbol{g}; v)| \leq \frac{d}{h}|g_b||v_N^-|, \tag{2.17}$$

$$|\mathcal{K}_{\text{bry}}(\boldsymbol{g}; r)| \leq \sqrt{d}\sqrt{\mu h^{-1}}(|g_a| + |g_b|)\|r\|. \tag{2.18}$$

The next lemma establishes the important relationship between $\|u_x\|$, $[\![u]\!]$ and $\|q\|$, which plays a key role in the stability analysis.

Lemma 2.4. *Suppose $\boldsymbol{w} = (u, q) \in V_h \times V_h$ is the solution of the scheme (2.6), then there exists a positive constant C_μ independent of h but maybe depending on the inverse constant μ, such that*

$$\|u_x\| + \sqrt{\mu h^{-1}}([\![u]\!]_{\text{int}} + |u_0^+|) \leq \frac{C_\mu}{\sqrt{d}}\|q\| + \sqrt{\mu h^{-1}}(|g_a| + |g_b| + |u_N^-|). \tag{2.19}$$

Proof. From (2.4b), (2.5c) and the definition of the numerical flux $\hat{h}_q$ defined in Table 1, we have

$$(q, r)_j = \begin{cases} -\sqrt{d}\left[(u, r_x)_1 - u_1^- r_1^- + g_a r_0^+\right], & j = 1, \\ -\sqrt{d}\left[(u, r_x)_j - u_j^- r_j^- + u_{j-1}^- r_{j-1}^+\right], & j = 2, \cdots, N-1, \\ -\sqrt{d}\left[(u, r_x)_N - g_b r_N^- + u_{N-1}^- r_{N-1}^+\right], & j = N. \end{cases}$$

Integrating by parts gives rise to

$$(q, r)_j = \begin{cases} \sqrt{d}\left[(u_x, r)_1 + (u_0^+ - g_a)r_0^+\right], & j = 1, \\ \sqrt{d}\left[(u_x, r)_j + [\![u]\!]_{j-1}r_{j-1}^+\right], & j = 2, \cdots, N-1, \\ \sqrt{d}\left[(u_x, r)_N + [\![u]\!]_{N-1}r_{N-1}^+ + (g_b - u_N^-)r_N^-\right], & j = N. \end{cases} \tag{2.20}$$

From Lemma 2.4 in [21], we can get

$$\|u_x\|_{I_j} + \sqrt{\mu h^{-1}}|[\![u]\!]_{j-1}| \leq \frac{C_\mu}{\sqrt{d}}\|q\|_{I_j}, \tag{2.21}$$

for $j = 2, \cdots, N-1$. Similarly, we can derive

$$\|u_x\|_{I_1} + \sqrt{\mu h^{-1}}|u_0^+| \leq \frac{C_\mu}{\sqrt{d}}\|q\|_{I_1} + \sqrt{\mu h^{-1}}|g_a|. \tag{2.22}$$

To obtain the result for $j = N$, we first take

$$r(x) = u_x(x) - (-1)^k u_x^+(x_{N-1})L_k(\xi),$$

in (2.20) for $j = N$, with $\xi = \frac{2x - (x_{N-1} + x_N)}{h_N}$, where $L_k(\xi)$ is the standard Legendre polynomial of degree k in $[-1, 1]$. It is obvious that $(u_x, r)_N = \|u_x\|_{I_N}^2$ since $(u_x, L_k)_N = 0$, and

$r_{N-1}^+ = 0$, $r_N^- = (u_x)_N^- - (-1)^k (u_x)_{N-1}^+$ because $L_k(-1) = (-1)^k$ and $L_k(1) = 1$. Then we have

$$\|u_x\|_{I_N}^2 = \frac{1}{\sqrt{d}}(q, r)_N - (g_b - u_N^-)r_N^-$$

$$\leq \frac{C_\mu}{\sqrt{d}}\|q\|_{I_N}\|u_x\|_{I_N} + \sqrt{\mu h^{-1}}(|u_N^-| + |g_b|)\|u_x\|_{I_N},$$

due to the inverse inequality (2.9) and the fact that $\|L_k\|_{I_N} \leq Ch^{1/2}$. Hence

$$\|u_x\|_{I_N} \leq \frac{C_\mu}{\sqrt{d}}\|q\|_{I_N} + \sqrt{\mu h^{-1}}(|u_N^-| + |g_b|). \tag{2.23}$$

Then taking $r = 1$ in (2.20) for $j = N$, we obtain

$$[\![u]\!]_{N-1} = \frac{1}{\sqrt{d}}(q, 1)_N - (u_x, 1)_N - (g_b - u_N^-).$$

Thus it follows from Cauchy-Schwarz inequality that

$$|[\![u]\!]_{N-1}| \leq Ch^{1/2}\left(\frac{1}{\sqrt{d}}\|q\|_{I_N} + \|u_x\|_{I_N}\right) + (|u_N^-| + |g_b|).$$

As a result, from (2.23) we have

$$\|u_x\|_{I_N} + \sqrt{\mu h^{-1}}|[\![u]\!]_{N-1}| \leq \frac{C_\mu}{\sqrt{d}}\|q\|_{I_N} + \sqrt{\mu h^{-1}}(|u_N^-| + |g_b|). \tag{2.24}$$

At last, combining the above results and summing up over $j = 1, \cdots, N$, we get the desired result (2.19). □

2.3   Fully discrete IMEX LDG scheme

In this paper we would like to adopt the third order IMEX RK time marching method [5] to update the semi-discrete LDG scheme (2.6). We call the corresponding fully-discrete LDG scheme as the IMEX-RK3-LDG scheme in this paper.

Let $\{t^n = n\tau\}_{n=0}^M$ be an uniform partition of the time interval $[0, T]$, with time step τ. The time step could actually change from step to step, but in this paper we take it as a constant for simplicity. Given (u^n, q^n), we would like to find the numerical solution at the next time level t^{n+1}, through three intermediate solutions $(u^{n,\ell}, q^{n,\ell})$, $\ell = 1, 2, 3$. In detail, for any function $(v, r) \in V_h \times V_h$

$$(u^{n,\ell}, v) = (u^n, v) + \tau \sum_{i=0}^{3} \left[a_{\ell i} \mathcal{H}(\boldsymbol{g}^{n,i}; u^{n,i}, v) + \hat{a}_{\ell i} \mathcal{L}(\boldsymbol{g}^{n,i}; \boldsymbol{w}^{n,i}, v) \right], \tag{2.25a}$$

$$(u^{n+1}, v) = (u^n, v) + \tau \sum_{i=0}^{3} \left[b_i \mathcal{H}(\boldsymbol{g}^{n,i}; u^{n,i}, v) + \hat{b}_i \mathcal{L}(\boldsymbol{g}^{n,i}; \boldsymbol{w}^{n,i}, v) \right], \tag{2.25b}$$

$$(q^{n,\ell}, r) = \mathcal{K}(\boldsymbol{g}^{n,\ell}; u^{n,\ell}, r), \quad \text{for} \quad \ell = 1, 2, 3, \tag{2.25c}$$

where the coefficients are given in the following table

$$
\begin{array}{c|cccc||cccc|c}
a_{\ell i} & \gamma & 0 & 0 & 0 & 0 & \gamma & 0 & 0 & \\
& \frac{1+\gamma}{2} - \alpha_1 & \alpha_1 & 0 & 0 & 0 & \frac{1-\gamma}{2} & \gamma & 0 & \hat{a}_{\ell i} \\
& 0 & 1-\alpha_2 & \alpha_2 & 0 & 0 & \beta_1 & \beta_2 & \gamma & \\
\hline
b_i & 0 & \beta_1 & \beta_2 & \gamma & 0 & \beta_1 & \beta_2 & \gamma & \hat{b}_i
\end{array}
\tag{2.26}
$$

The left half of the table lists $a_{\ell i}$ and b_i, with the columns from left to right corresponding to $i = 0, 1, 2, 3$, and the first three rows from top to bottom corresponding to $\ell = 1, 2, 3$. Similarly, the right half lists $\hat{a}_{\ell i}$ and $\hat{b}_i$. Noting that we use explicit time discretization for the operator $\mathcal{H}$ since $a_{\ell i} = 0$ for all $i \geq \ell$, and we use diagonally implicit time discretization for the operator $\mathcal{L}$, where $\hat{a}_{ii} = \gamma$ and $\hat{a}_{\ell i} = 0$ for all $i > \ell$.

In the above time discretization scheme, $\gamma \approx 0.435866521508459$ is the middle root of $6x^3 - 18x^2 + 9x - 1 = 0$. And $\beta_1 = -\frac{3}{2}\gamma^2 + 4\gamma - \frac{1}{4}$, $\beta_2 = \frac{3}{2}\gamma^2 - 5\gamma + \frac{5}{4}$. The parameter α_1 is chosen as -0.35 and $\alpha_2 = \frac{\frac{1}{3} - 2\gamma^2 - 2\beta_2\alpha_1\gamma}{\gamma(1-\gamma)}$. Compared with other third order IMEX-RK schemes [3, 16], the advantage of this scheme is that there are fewer intermediate stages for the implicit part, and hence it is more efficient.

The notation $\boldsymbol{g}^{n,\ell} = (g_a^{n,\ell}, g_b^{n,\ell})$ is used to represent the numerical boundary conditions at $x = a$ and $x = b$, at each intermediate time level $t^{n,\ell}$. As we have mentioned, improper boundary treatment may affect the accuracy of the scheme. So the setting of $\boldsymbol{g}^{n,\ell}$ is one of the most important aspects of this paper, which will be discussed in next section.

3 Strategy of boundary treatment

In this section, we will propose a boundary treatment strategy for the IMEX-RK3-LDG scheme (2.25), by firstly studying the boundary treatment for purely implicit third order scheme, i.e, the implicit part of (2.25). To simplify notations, we use $u(x,t)$ to denote the exact solution and $g(t)$ to denote the given boundary condition in this section, these notations are only valid in this section.

3.1 The purely implicit case

To make the idea clear enough, we consider the linear diffusion problem

$$u_t = u_{xx}$$

with boundary conditions

$$u|_{\partial\Omega} = g(t).$$

The third order implicit scheme implemented in interior reads as follows

$$u^{n,1} = u^n + \gamma\tau u_{xx}^{n,1}, \tag{3.1a}$$

$$u^{n,2} = u^n + \frac{1-\gamma}{2}\tau u_{xx}^{n,1} + \gamma\tau u_{xx}^{n,2}, \tag{3.1b}$$

$$u^{n,3} = u^n + \beta_1\tau u_{xx}^{n,1} + \beta_2\tau u_{xx}^{n,2} + \gamma\tau u_{xx}^{n,3}, \tag{3.1c}$$

$$u^{n+1} = u^{n,3}, \tag{3.1d}$$

where $\beta_1 + \beta_2 + \gamma = 1$. Replacing $u_{xx}^{n,\ell}$ with $u_t^{n,\ell}$ (only consider intermediate stages), we get

$$u^{n,1} = u^n + \gamma\tau u_t^{n,1}, \tag{3.2a}$$

$$u^{n,2} = u^n + \frac{1-\gamma}{2}\tau u_t^{n,1} + \gamma\tau u_t^{n,2}, \tag{3.2b}$$

$$u^{n,3} = u^n + \beta_1\tau u_t^{n,1} + \beta_2\tau u_t^{n,2} + \gamma\tau u_t^{n,3}. \tag{3.2c}$$

Letting $u_t^{n,\ell} \approx u_t(t^{n,\ell})$ and extending the above scheme up to boundary, we get the strategy of boundary treatment at intermediate stages

$$g_{\text{im}}^{n,1} = g(t^n) + \gamma\tau g_t(t^{n,1}), \tag{3.3a}$$

$$g_{\text{im}}^{n,2} = g(t^n) + \frac{1-\gamma}{2}\tau g_t(t^{n,1}) + \gamma\tau g_t(t^{n,2}), \tag{3.3b}$$

$$g_{\text{im}}^{n,3} = g(t^n) + \beta_1\tau g_t(t^{n,1}) + \beta_2\tau g_t(t^{n,2}) + \gamma\tau g_t(t^{n,3}), \tag{3.3c}$$

where $t^{n,1} = t^n + \gamma\tau, t^{n,2} = t^n + \frac{1+\gamma}{2}\tau, t^{n,3} = t^n + \tau$.

Suppose u is smooth enough so that Taylor expansion can be carried out. From (3.2), we get $u_t^{n,\ell} = u_t(t^{n,\ell}) + \mathcal{O}(\tau^2)$ by Taylor expansion, so we have

$$g_{\text{im}}^{n,\ell} = g^{n,\ell} + \mathcal{O}(\tau^3), \tag{3.4}$$

where $g^{n,\ell}$ is the corresponding reference boundary condition.

In Table 2 we list errors and orders of accuracy for the IMEX-RK3-LDG scheme (2.25) with the boundary treatment (3.3), for solving $u_t = u_{xx}$ on $[-1, 1]$ with Dirichlet boundary conditions which are given by the exact solution $u(x,t) = e^{-t}\sin(x)$. The computing time is $T = 5$, time step is $\tau = 0.5h$ and piecewise quadratic polynomials are used in the spatial discretization. We also test the conventional treatment

$$g_{\text{ex}}^{n,\ell} = g(t^{n,\ell}), \qquad \text{for} \quad \ell = 1, 2, 3 \tag{3.5}$$

as a comparison. From the table we can observe that the conventional treatment will lose accuracy while the proposed strategy (3.3) can recover the third order accuracy.

Table 2: Errors and orders of accuracy of strategies (3.3) and (3.5) for $u_t = u_{xx}$.

N	strategy (3.3)				strategy (3.5)			
	L^∞ error	L^∞ order	L^2 error	L^2 order	L^∞ error	L^∞ order	L^2 error	L^2 order
20	1.16E-07	-	4.15E-08	-	2.14E-07	-	1.18E-07	-
40	1.42E-08	3.02	5.19E-09	3.00	4.37E-08	2.29	2.20E-08	2.42
80	1.78E-09	3.00	6.51E-10	2.99	9.54E-09	2.20	4.32E-09	2.35
160	2.22E-10	3.00	1.02E-11	2.99	2.24E-09	2.09	8.76E-10	2.30
320	2.78E-11	3.00	1.02E-11	3.00	5.43E-10	2.05	1.81E-10	2.28
640	3.48E-12	3.00	1.28E-12	3.00	1.33E-10	2.03	3.76E-11	2.26

3.2　The implicit-explicit case

In this subsection, we consider the simple linear convection-diffusion model problem

$$u_t = cu_x + du_{xx}$$

with boundary condition

$$u|_{\partial\Omega} = g(t),$$

where c and $d > 0$ are coefficients of convection and diffusion, respectively. The third order IMEX RK scheme implemented in interior reads

$$u^{n,1} = u^n + \gamma\tau cu_x^n + \gamma\tau du_{xx}^{n,1}, \tag{3.6a}$$

$$u^{n,2} = u^n + (\frac{1+\gamma}{2} - \alpha_1)\tau cu_x^n + \alpha_1\tau cu_x^{n,1} + \frac{1-\gamma}{2}\tau du_{xx}^{n,1} + \gamma\tau du_{xx}^{n,2}, \tag{3.6b}$$

$$u^{n,3} = u^n + (1 - \alpha_2)\tau cu_x^{n,1} + \alpha_2\tau cu_x^{n,2} + \beta_1\tau du_{xx}^{n,1} + \beta_2\tau du_{xx}^{n,2} + \gamma\tau du_{xx}^{n,3}, \tag{3.6c}$$

$$u^{n+1} = u^n + \beta_1\tau(cu_x^{n,1} + du_{xx}^{n,1}) + \beta_2\tau(cu_x^{n,2} + du_{xx}^{n,2}) + \gamma\tau(cu_x^{n,3} + du_{xx}^{n,3}). \tag{3.6d}$$

We only need to consider the intermediate stages, we would like to take the first stage as an example to show the idea. Note that $u_t^{n,1} = cu_x^{n,1} + du_{xx}^{n,1}$, so from (3.6a) we have

$$u^{n,1} = u^n + \gamma\tau cu_x^{n,1} + \gamma\tau du_{xx}^{n,1} - \gamma\tau c(u_x^{n,1} - u_x^n)$$
$$= u^n + \gamma\tau u_t^{n,1} - \gamma\tau c(u_x^{n,1} - u_x^n).$$

Taking derivative with respect to x on both sides of (3.6a) we get

$$u_x^{n,1} = u_x^n + \gamma\tau cu_{xx}^n + \gamma\tau du_{xxx}^{n,1} = u_x^n + \gamma\tau(cu_{xx}^n + du_{xxx}^n) + \mathcal{O}(\tau^2).$$

Hence

$$u^{n,1} = u^n + \gamma\tau u_t^{n,1} - \gamma^2\tau^2 c(cu_{xx}^n + du_{xxx}^n) + \mathcal{O}(\tau^3). \tag{3.7a}$$

Similarly,

$$u^{n,2} = u^n + \frac{1-\gamma}{2}\tau u_t^{n,1} + \gamma\tau u_t^{n,2} + (\alpha_1 - 1)\gamma\tau^2 c(cu_{xx}^n + du_{xxx}^n) + \mathcal{O}(\tau^3), \tag{3.7b}$$

$$u^{n,3} = u^n + \beta_1\tau u_t^{n,1} + \beta_2\tau u_t^{n,2} + \gamma\tau u_t^{n,3} + \chi\tau^2 c(cu_{xx}^n + du_{xxx}^n) + \mathcal{O}(\tau^3), \tag{3.7c}$$

where $\chi = (\alpha_2 - \beta_2)\frac{1+\gamma}{2} - (\alpha_2 + \beta_1)\gamma$.

Remark 3.1. Notice that the expressions in (3.7) contain derivatives with respect to spatial variable, so we have to approximate these derivatives $u_{xx}^n = u_{xx}(t^n), u_{xxx}^n = u_{xxx}(t^n)$ by their inner approximations when extending the above scheme to boundary. These extra terms prevent us from obtaining boundary corrections which solely contain the physical boundary condition and its time derivatives, compared with fully explicit or implicit cases. Moreover, the approximations must be at least first order (in time) to recover third order accuracy of the scheme. However, approximations of these derivatives are done in space, hence, certain restrictions on the time step with respect to the mesh size may be required.

Letting $u_t^{n,\ell} \approx u_t(t^{n,\ell})$ and extending the inner scheme (3.7) up to the boundary, by taking suitable approximations for u_{xx}^n and u_{xxx}^n, we get the strategy of boundary treatment for the third order IMEX RK scheme (4.19), which is a modification of (3.3) and is given as

$$g_{\text{imex}}^{n,\ell} = g_{\text{im}}^{n,\ell} + A_\ell \tau^2 \mathrm{R}^n, \quad \text{for} \quad \ell = 1, 2, 3, \tag{3.8}$$

where

$$A_1 = -\gamma^2, \quad A_2 = (\alpha_1 - 1)\gamma, \quad A_3 = \chi = (\alpha_2 - \beta_2)\frac{1+\gamma}{2} - (\alpha_2 + \beta_1)\gamma$$

and R^n is any first order approximation of $c(cu_{xx}^n + du_{xxx}^n)$. Next we will give a method to approximate u_{xx}^n and u_{xxx}^n at the boundaries.

Approximation for u_{xx}^n. Since piecewise quadratic polynomials are always adopted in the spatial discretization to match the third order accuracy in time, we can approximate u_{xx} by

$$u_{xx}(x_a) \approx P_1''(x_a), \quad u_{xx}(x_b) \approx P_N''(x_b), \tag{3.9}$$

where x_a and x_b are the boundary points, and P_1 and P_N are the quadratic polynomials in the first and the last cells, respectively. This approximation is of order $\mathcal{O}(h)$, so, to ensure the first order approximation in time, we require the time step $\tau = \mathcal{O}(h)$ in the computation.

Approximation for u_{xxx}^n. Since $k = 2$, the second derivative of the numerical solution is piecewise constant on each cell, we can use the difference quotient of the second order derivatives in neighboring cells to approximate u_{xxx} on the boundaries. For example, we can simply adopt

$$u_{xxx}(x_a) \approx (P_2'' - P_1'')/h, \quad u_{xxx}(x_b) \approx (P_N'' - P_{N-1}'')/h. \tag{3.10}$$

In Tables 3 and 4 we list the errors and orders of accuracy for the IMEX-RK3-LDG scheme (2.25) with the boundary treatment (3.8), for solving $u_t = cu_x + u_{xx}$ on $[-1, 1]$ with Dirichlet boundary conditions given by the exact solution $u(x,t) = e^{-t}\sin(x + ct)$. The computing time is $T = 5$, time step is $\tau = 0.1h$ for $c = 1$ and $\tau = 0.5h$ for $c = 0.1$, piecewise quadratic polynomials is used in the spatial discretization. We also test the conventional treatment (3.5) as a comparison. From these tables we can observe that the conventional treatment will lose accuracy while the proposed strategy (3.8) can recover the third order accuracy.

4   Stability and error estimates

In this section, we would like to present the stability and error estimates of the IMEX-RK3-LDG scheme proposed in the previous sections. To this end, we first give some notations which will be used throughout the remaining of this paper.

For vector-valued function $\mathbf{v} = (v_1, v_2, \cdots, v_\ell)^\top$, we denote

$$|\mathbf{v}|^2 = \sum_{i=1}^{\ell} v_i^2 \quad \text{and} \quad \|\mathbf{v}\|^2 = \sum_{i=1}^{\ell} \|v_i\|^2,$$

Table 3: Errors and orders of accuracy of strategies (3.8) and (3.5) for $u_t = u_x + u_{xx}$.

N	strategy (3.8)				strategy (3.5)			
	L^∞ error	L^∞ order	L^2 error	L^2 order	L^∞ error	L^∞ order	L^2 error	L^2 order
20	1.00E-07	-	2.66E-08	-	5.65E-06	-	1.72E-06	-
40	1.28E-08	2.97	3.32E-09	3.00	1.57E-06	1.85	3.70E-07	2.22
80	1.62E-09	2.98	4.14E-10	3.00	4.22E-07	1.89	7.94E-08	2.22
160	2.05E-10	2.99	5.18E-11	3.00	1.11E-07	1.92	1.70E-08	2.22
320	2.58E-11	2.99	6.47E-12	3.00	2.90E-08	1.94	3.63E-09	2.23
640	3.24E-12	2.99	8.10E-13	3.00	7.45E-09	1.96	7.71E-10	2.23

Table 4: Errors and orders of accuracy of strategies (3.8) and (3.5) for $u_t = 0.1u_x + u_{xx}$.

N	strategy (3.8)				strategy (3.5)			
	L^∞ error	L^∞ order	L^2 error	L^2 order	L^∞ error	L^∞ order	L^2 error	L^2 order
20	1.18E-07	-	4.37E-08	-	4.55E-07	-	1.42E-07	-
40	1.51E-08	2.97	5.51E-09	2.99	1.15E-07	1.98	2.76E-08	2.37
80	1.95E-09	2.95	6.90E-10	3.00	2.92E-08	1.99	5.56E-09	2.31
160	2.51E-10	2.96	8.63E-11	3.00	7.34E-09	1.99	1.15E-09	2.28
320	3.11E-11	3.01	1.08E-11	3.00	1.85E-09	1.99	2.39E-10	2.26
640	3.26E-12	3.26	1.34E-12	3.01	4.63E-10	1.99	5.02E-11	2.25

where $\| \cdot \|$ is the L^2-norm in Ω. We use the standard norms and semi-norms in Sobolev spaces, and also use the notation $L^\infty(H^s)$ to represent the set of functions v such that $\max_{0 \le t \le T} \|v(\cdot, t)\|_{H^s(\Omega)} < \infty$.

Throughout this paper, we denote C as a generic positive constant which is independent of h and τ, also we use $\varepsilon > 0$ to represent an arbitrary small constant, they may have different values in different occurrence.

4.1 Stability analysis

For the convenience of analysis, we would like to introduce a series of notations following [21, 22]

$$\mathbb{E}_1 w^n = w^{n,1} - w^n; \qquad \mathbb{E}_2 w^n = w^{n,2} - 2w^{n,1} + w^n;$$
$$\mathbb{E}_3 w^n = 2w^{n,3} + w^{n,2} - 3w^{n,1}; \qquad \mathbb{E}_4 w^n = w^{n+1} - w^{n,3}, \tag{4.1}$$

for arbitrary w, and after some algebraic manipulations we can rewrite scheme (2.25) into the following compact form

$$(\mathbb{E}_\ell u^n, v) = \Phi_\ell(\boldsymbol{g}_n; \boldsymbol{u}_n, v) + \Psi_\ell(\boldsymbol{g}_n; \boldsymbol{w}_n, v), \quad \text{for} \quad \ell = 1, 2, 3, 4, \tag{4.2a}$$
$$(q^{n,\ell}, r) = \mathcal{K}(\boldsymbol{g}^{n,\ell}; u^{n,\ell}, r), \quad \text{for} \quad \ell = 1, 2, 3 \tag{4.2b}$$

where $\boldsymbol{u}_n = (u^n, u^{n,1}, u^{n,2}, u^{n,3})$, $\boldsymbol{g}_n = (\boldsymbol{g}^n, \boldsymbol{g}^{n,1}, \boldsymbol{g}^{n,2}, \boldsymbol{g}^{n,3})$ and $\boldsymbol{w}_n = (\boldsymbol{w}^{n,1}, \boldsymbol{w}^{n,2}, \boldsymbol{w}^{n,3})$.

$$\Phi_\ell(\boldsymbol{g}_n; \boldsymbol{u}_n, v) = \sum_{i=0}^{3} \delta_{\ell i} \tau \mathcal{H}(\boldsymbol{g}^{n,i}; u^{n,i}, v), \tag{4.3}$$

$$\Psi_\ell(\boldsymbol{g}_n; \boldsymbol{w}_n, v) = \theta_{\ell 1} \tau \mathcal{L}(\boldsymbol{g}^{n,1}; \boldsymbol{w}^{n,1}, v) + \theta_{\ell 2} \tau [\mathcal{L}(\boldsymbol{g}^{n,2}; \boldsymbol{w}^{n,2}, v) - 2\mathcal{L}(\boldsymbol{g}^{n,1}; \boldsymbol{w}^{n,1}, v)]$$
$$+ \theta_{\ell 3} \tau \mathcal{L}(\boldsymbol{g}^{n,3}; \boldsymbol{w}^{n,3}, v), \tag{4.4}$$

for $\ell = 1, 2, 3, 4$. The coefficients $\delta_{\ell i}$ and $\theta_{\ell i}$ are listed in Table 5. See [21, 22] for more details.

Table 5: The coefficients $\delta_{\ell i}$ and $\theta_{\ell i}$ in (4.3) and (4.4).

ℓ \ i	$\delta_{\ell i}$				$\theta_{\ell i}$		
	0	1	2	3	1	2	3
1	γ	0	0	0	γ	0	0
2	$\frac{1-3\gamma}{2} - \alpha_1$	α_1	0	0	$\frac{1-\gamma}{2}$	γ	0
3	$\frac{1-5\gamma}{2} - \alpha_1$	$2(1-\alpha_2) + \alpha_1$	$2\alpha_2$	0	$2(\frac{9}{4} - \frac{11}{4}\gamma - \beta_1)$	$2(1 - \beta_1 - \frac{\gamma}{2})$	2γ
4	0	$\alpha_2 - \beta_2 - \gamma$	$\beta_2 - \alpha_2$	γ	0	0	0

Theorem 4.1. *There exists a positive constant τ_0 independent of h, such that the solution of (2.25) satisfies*

$$\|u^n\|^2 \leq e^{n\tau} \left(\|u^0\|^2 + C\tau \sum_{m=0}^{n-1} h^{-1} |\mathbf{g}^m|^2 \right), \tag{4.5}$$

where C is independent of h and τ, $|\mathbf{g}^m|^2 = \sum_{\ell=1}^{3} |g_a^{m,\ell}|^2 + |g_b^{m,\ell}|^2$.

Proof. Along the same line as the proof in [21], we take the test functions $v_\ell = u^{n,1}, u^{n,2} - 2u^{n,1}, u^{n,3}$ and $2u^{n+1}$ in (4.2), for $\ell = 1, 2, 3, 4$, respectively. Adding them together, we get the energy equation

$$\|u^{n+1}\|^2 - \|u^n\|^2 + \mathcal{S} = \mathcal{T}_c + \mathcal{T}_d, \tag{4.6}$$

where

$$\mathcal{S} = \frac{1}{2} \left(\|\mathbb{E}_1 u^n\|^2 + \|\mathbb{E}_2 u^n\| + \|\mathbb{E}_{31} u^n\|^2 + \|\mathbb{E}_{32} u^n\|^2 + 2\|\mathbb{E}_4 u^n\|^2 \right) \tag{4.7}$$

is the stability term coming from the time-marching, with $\mathbb{E}_{31} u^n = u^{n,3} + u^{n,2} - 2u^{n,1}$, $\mathbb{E}_{32} u^n = u^{n,3} - u^{n,1}$, and

$$\mathcal{T}_c = \sum_{\ell=1}^{4} \Phi_\ell(\boldsymbol{g}_n; \boldsymbol{u}_n, v_\ell), \qquad \mathcal{T}_d = \sum_{\ell=1}^{4} \Psi_\ell(\boldsymbol{g}_n; \boldsymbol{w}_n, v_\ell). \tag{4.8}$$

We will first estimate the term $\mathcal{T}_d$. By the definition (4.4), we get

$$\mathcal{T}_d = \theta_{11} \tau \mathcal{L}(\boldsymbol{g}^{n,1}; \boldsymbol{w}^{n,1}, u^{n,1}) + \theta_{21} \tau \mathcal{L}(\boldsymbol{g}^{n,1}; \boldsymbol{w}^{n,1}, u^{n,2} - 2u^{n,1}) + \theta_{31} \tau \mathcal{L}(\boldsymbol{g}^{n,1}; \boldsymbol{w}^{n,1}, u^{n,3})$$
$$+ \theta_{22} \tau [\mathcal{L}(\boldsymbol{g}^{n,2}; \boldsymbol{w}^{n,2}, u^{n,2} - 2u^{n,1}) - 2\mathcal{L}(\boldsymbol{g}^{n,1}; \boldsymbol{w}^{n,1}, u^{n,2} - 2u^{n,1})]$$
$$+ \theta_{32} \tau [\mathcal{L}(\boldsymbol{g}^{n,2}; \boldsymbol{w}^{n,2}, u^{n,3}) - 2\mathcal{L}(\boldsymbol{g}^{n,1}; \boldsymbol{w}^{n,1}, u^{n,3})] + \theta_{33} \tau \mathcal{L}(\boldsymbol{g}^{n,3}; \boldsymbol{w}^{n,3}, u^{n,3}).$$

Furthermore, by the division of the operator $\mathcal{L}$, we get

$$\mathcal{T}_{\mathrm{d}} = \mathcal{T}_{\mathrm{d}}^{\mathrm{int}} + \mathcal{T}_{\mathrm{d}}^{\mathrm{bry}},$$

where

$$
\begin{aligned}
\mathcal{T}_{\mathrm{d}}^{\mathrm{int}} ={}& \theta_{11}\tau\mathcal{L}_{\mathrm{int}}(\boldsymbol{w}^{n,1}, u^{n,1}) + \theta_{21}\tau\mathcal{L}_{\mathrm{int}}(\boldsymbol{w}^{n,1}, u^{n,2} - 2u^{n,1}) + \theta_{31}\tau\mathcal{L}_{\mathrm{int}}(\boldsymbol{w}^{n,1}, u^{n,3}) \\
&+ \theta_{22}\tau\mathcal{L}_{\mathrm{int}}(\boldsymbol{w}^{n,2} - 2\boldsymbol{w}^{n,1}, u^{n,2} - 2u^{n,1}) + \theta_{32}\tau\mathcal{L}_{\mathrm{int}}(\boldsymbol{w}^{n,2} - 2\boldsymbol{w}^{n,1}, u^{n,3}) \\
&+ \theta_{33}\tau\mathcal{L}_{\mathrm{int}}(\boldsymbol{w}^{n,3}, u^{n,3}), \\
\mathcal{T}_{\mathrm{d}}^{\mathrm{bry}} ={}& \theta_{11}\tau\mathcal{L}_{\mathrm{bry}}(\boldsymbol{g}^{n,1}; u^{n,1}) + \theta_{21}\tau\mathcal{L}_{\mathrm{bry}}(\boldsymbol{g}^{n,1}; u^{n,2} - 2u^{n,1}) + \theta_{31}\tau\mathcal{L}_{\mathrm{bry}}(\boldsymbol{g}^{n,1}; u^{n,3}) \\
&+ \theta_{22}\tau\mathcal{L}_{\mathrm{bry}}(\boldsymbol{g}^{n,2} - 2\boldsymbol{g}^{n,1}; u^{n,2} - 2u^{n,1}) + \theta_{32}\tau\mathcal{L}_{\mathrm{bry}}(\boldsymbol{g}^{n,2} - 2\boldsymbol{g}^{n,1}; u^{n,3}) \\
&+ \theta_{33}\tau\mathcal{L}_{\mathrm{bry}}(\boldsymbol{g}^{n,3}; u^{n,3}).
\end{aligned}
$$

Denote by

$$
\begin{aligned}
\mathcal{T}_{\mathrm{d}}^{\mathrm{bry}'} ={}& \theta_{11}\tau\mathcal{K}_{\mathrm{bry}}(\boldsymbol{g}^{n,1}; q^{n,1}) + \theta_{21}\tau\mathcal{K}_{\mathrm{bry}}(\boldsymbol{g}^{n,1}; q^{n,2} - 2q^{n,1}) + \theta_{31}\tau\mathcal{K}_{\mathrm{bry}}(\boldsymbol{g}^{n,1}; q^{n,3}) \\
&+ \theta_{22}\tau\mathcal{K}_{\mathrm{bry}}(\boldsymbol{g}^{n,2} - 2\boldsymbol{g}^{n,1}; q^{n,2} - 2q^{n,1}) + \theta_{32}\tau\mathcal{K}_{\mathrm{bry}}(\boldsymbol{g}^{n,2} - 2\boldsymbol{g}^{n,1}; q^{n,3}) \\
&+ \theta_{33}\tau\mathcal{K}_{\mathrm{bry}}(\boldsymbol{g}^{n,3}; q^{n,3}).
\end{aligned}
$$

Owing to (2.11) and (4.2b), we can derive

$$\mathcal{T}_{\mathrm{d}}^{\mathrm{int}} = -\frac{d}{h}\tau \mathbf{u}_N^{n\top}\mathbb{A}\mathbf{u}_N^n - \tau\int_{\Omega} \mathbf{q}^{n\top}\mathbb{A}\mathbf{q}^n \mathrm{d}x + \mathcal{T}_{\mathrm{d}}^{\mathrm{bry}'}, \tag{4.9}$$

where $\mathbf{u}_N^n = ((u_N^{n,1})^-, (u_N^{n,2})^- - 2(u_N^{n,1})^-, (u_N^{n,3})^-)^\top$, $\mathbf{q}^n = (q^{n,1}, q^{n,2} - 2q^{n,1}, q^{n,3})^\top$, and

$$\mathbb{A} = \frac{1}{2}\begin{pmatrix} 2\theta_{11} & \theta_{21} & \theta_{31} \\ \theta_{21} & 2\theta_{22} & \theta_{32} \\ \theta_{31} & \theta_{32} & 2\theta_{33} \end{pmatrix}. \tag{4.10}$$

Since all the leading principal minors of $\mathbb{A}$ are positive, we can conclude that $\mathbb{A}$ is positive definite. In fact we can show in the same way that $\mathbb{A} - \frac{\gamma}{4}\mathbb{I}$ is positive definite, where $\mathbb{I}$ is the identity matrix. So

$$\mathcal{T}_{\mathrm{d}}^{\mathrm{int}} \le -\frac{\gamma}{4}\tau\left[\frac{d}{h}|\mathbf{u}_N^n|^2 + \|\mathbf{q}^n\|^2\right] + \mathcal{T}_{\mathrm{d}}^{\mathrm{bry}'}. \tag{4.11}$$

From Lemma 2.3 and the Young's inequality we obtain

$$\left|\mathcal{T}_{\mathrm{d}}^{\mathrm{bry}} + \mathcal{T}_{\mathrm{d}}^{\mathrm{bry}'}\right| \le \varepsilon\tau\left[\frac{d}{h}|\mathbf{u}_N^n|^2 + \|\mathbf{q}^n\|^2\right] + Cdh^{-1}\tau|\mathbf{g}^n|^2. \tag{4.12}$$

Hence

$$\mathcal{T}_{\mathrm{d}} \le \left(\varepsilon - \frac{\gamma}{4}\right)\tau\left[\frac{d}{h}|\mathbf{u}_N^n|^2 + \|\mathbf{q}^n\|^2\right] + Cdh^{-1}\tau|\mathbf{g}^n|^2. \tag{4.13}$$

Next, we are going to estimate $\mathcal{T}_{\mathrm{c}}$. Notice that

$$\mathcal{T}_{\mathrm{c}} = \sum_{\ell=1}^{4}\sum_{i=0}^{3}\delta_{\ell i}\tau\mathcal{H}_{\mathrm{int}}(u^{n,i}, v_\ell) + \sum_{\ell=1}^{4}\sum_{i=0}^{3}\delta_{\ell i}\tau\mathcal{H}_{\mathrm{bry}}(\boldsymbol{g}^{n,i}; v_\ell) = \mathcal{T}_{\mathrm{c}}^{\mathrm{int}} + \mathcal{T}_{\mathrm{c}}^{\mathrm{bry}}.$$

Along the similar line as [21], we get

$$\mathcal{T}_c^{\mathrm{int}} = -\frac{c}{2}\tau\left([\![u^{n,1}]\!]^2 + \frac{3\gamma-1}{2}[\![u^{n,2}-2u^{n,1}]\!]^2 + \frac{5(1-\gamma)}{2}[\![u^{n,3}]\!]^2\right) + \sum_{i=1}^{3} T_i, \qquad (4.14)$$

where we have used the property (2.10), and T_i are given as

$$
\begin{aligned}
T_1 &= 2(\beta_2 - \alpha_2 - \gamma)\tau\mathcal{H}_{\mathrm{int}}(u^{n,1}, \mathbb{E}_4 u^n) - \gamma\tau\mathcal{H}_{\mathrm{int}}(\mathbb{E}_1 u^n, u^{n,1}), \\
T_2 &= 2(\beta_2 - \alpha_2)\tau\mathcal{H}_{\mathrm{int}}(u^{n,2}-2u^{n,1}, \mathbb{E}_4 u^n) + \alpha_1\tau\mathcal{H}_{\mathrm{int}}(\mathbb{E}_1 u^n, u^{n,2}-2u^{n,1}), \\
&\quad + \frac{1-3\gamma}{2}\tau\mathcal{H}_{\mathrm{int}}(\mathbb{E}_2 u^n, u^{n,2}-2u^{n,1}) \\
T_3 &= 2\gamma\tau\mathcal{H}_{\mathrm{int}}(u^{n,3}, \mathbb{E}_4 u^n) + 2\beta_2\mathcal{H}_{\mathrm{int}}(\mathbb{E}_2 u^n, u^{n,3}) \\
&\quad + \left(\alpha_1 + 2\beta_2 - \frac{1-5\gamma}{2}\right)\tau\mathcal{H}_{\mathrm{int}}(\mathbb{E}_1 u^n, u^{n,3}) - \frac{5-9\gamma}{2}\tau\mathcal{H}_{\mathrm{int}}(\mathbb{E}_{32} u^n, u^{n,3}).
\end{aligned}
$$

Denote $C_\star$ as the maximum of the absolute value of all the coefficients in the expression of T_i for $i = 1, 2, 3$, and denote

$$T_0 = \|\mathbb{E}_1 u^n\| + \|\mathbb{E}_2 u^n\| + \|\mathbb{E}_4 u^n\| + \|\mathbb{E}_{32} u^n\|,$$

then by the aid of Lemmas 2.2 and 2.4, we can derive

$$
\begin{aligned}
|T_1| &\le C_\star c\tau\left(\|(u^{n,1})_x\| + \sqrt{\mu h^{-1}}([\![u^{n,1}]\!]_{\mathrm{int}} + |(u_0^{n,1})^+| + |(u_N^{n,1})^-|)\right)T_0 \\
&\le C_\star c\tau\left(\frac{C_\mu}{\sqrt{d}}\|q^{n,1}\| + \sqrt{\mu h^{-1}}(|g_a^{n,1}| + |g_b^{n,1}| + |(u_N^{n,1})^-|)\right)T_0.
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
|T_2| &\le C_\star c\tau\left(\frac{C_\mu}{\sqrt{d}}\|q^{n,2}-2q^{n,1}\| + \sqrt{\mu h^{-1}}(|g_a^{n,2}-2g_a^{n,1}| + |g_b^{n,2}-2g_b^{n,1}|\right. \\
&\quad \left. + |(u_N^{n,2})^- - 2(u_N^{n,1})^-|)\right)T_0, \\
|T_3| &\le C_\star c\tau\left(\frac{C_\mu}{\sqrt{d}}\|q^{n,3}\| + \sqrt{\mu h^{-1}}(|g_a^{n,3}| + |g_b^{n,3}| + |(u_N^{n,3})^-|)\right)T_0.
\end{aligned}
$$

Then using the Young's inequality, we obtain

$$\left|\sum_{i=1}^{3} T_i\right| \le \varepsilon\tau\left[\|\mathbf{q}^n\|^2 + \frac{d}{h}|\mathbf{u}_N^n|^2 + \frac{d}{h}|\mathbf{g}^n|^2\right] + C\frac{c^2}{d}\tau\mathcal{S}, \qquad (4.15)$$

where $\mathcal{S}$ is defined in (4.7). Furthermore, from Lemma 2.3 and the Young's inequality we obtain

$$|\mathcal{T}_c^{\mathrm{bry}}| \le \varepsilon\tau\sum_{\ell=1}^{4}\|v_\ell\|^2 + Cc^2 h^{-1}\tau|\mathbf{g}^n|^2 \le \varepsilon\tau\left[\|u^n\|^2 + \mathcal{S}\right] + Cc^2 h^{-1}\tau|\mathbf{g}^n|^2. \qquad (4.16)$$

Thus

$$\mathcal{T}_c \le \varepsilon\tau\left[\|u^n\|^2 + \|\mathbf{q}^n\|^2 + \frac{d}{h}|\mathbf{u}_N^n|^2\right] + C\frac{c^2}{d}\tau\mathcal{S} + C(c^2+d)h^{-1}\tau|\mathbf{g}^n|^2. \tag{4.17}$$

As a result, owing to (4.6), (4.13) and (4.17), and by choosing $\varepsilon \le \frac{\gamma}{8}$ and letting $\tau \le \tau_0$ for some $\tau_0 \propto \frac{d}{c^2}$, we can get

$$\|u^{n+1}\|^2 - \|u^n\|^2 \le \varepsilon\tau\|u^n\|^2 + C\tau h^{-1}|\mathbf{g}^n|^2. \tag{4.18}$$

Thus the use of discrete Gronwall's inequality yields (4.5). $\qquad\square$

4.2 Error estimates

4.2.1 Reference function and the main result

To proceed with the error analysis, we follow [21] to introduce four reference functions, denoted by $\boldsymbol{W}^{(\ell)} = (U^{(\ell)}, Q^{(\ell)})$, $\ell = 0, 1, 2, 3$, associated with the third order IMEX RK time discretization. In detail, $U^{(0)} = U$ is the exact solution of the problem (2.1) and then we define

$$U^{(\ell)} = U^{(0)} + \tau\sum_{i=0}^{3}(a_{\ell i}cU_x^{(i)} + \hat{a}_{\ell i}\sqrt{d}Q_x^{(i)}), \quad \text{for} \quad \ell = 1, 2, 3 \tag{4.19a}$$

where

$$Q^{(\ell)} = \sqrt{d}U_x^{(\ell)}, \quad \text{for} \quad \ell = 1, 2, 3. \tag{4.19b}$$

For any indices n and ℓ under consideration, the reference function at each stage time level is defined as $\boldsymbol{W}^{n,\ell} = (U^{n,\ell}, Q^{n,\ell}) = \boldsymbol{W}^{(\ell)}(x, t^n)$.

Assume the exact solution U satisfies the following smoothness:

$$D_t^\ell U \in L^\infty(H^{k+2}), (\ell = 0, 1), \quad D_t^2 U \in L^\infty(H^{k+1}) \quad \text{and} \quad D_t^4 U \in L^\infty(L^2), \tag{4.20}$$

where $D_t^\ell U$ is the ℓ-th order time derivative of U. We have the following lemma.

Lemma 4.1. *Let $\boldsymbol{W} = (U, Q)$ be the solution of problem (2.1), assume U satisfies the smoothness assumption (4.20). Denote $\boldsymbol{W}_n = (\boldsymbol{W}^{n,1}, \boldsymbol{W}^{n,2}, \boldsymbol{W}^{n,3})$, $\boldsymbol{U}_n = (U^n, U^{n,1}, U^{n,2}, U^{n,3})$ and $\boldsymbol{G}_n = (\boldsymbol{G}^n, \boldsymbol{G}^{n,1}, \boldsymbol{G}^{n,2}, \boldsymbol{G}^{n,3})$. Then for any function $(v, r) \in V_h \times V_h$, there hold*

$$(\mathbb{E}_\ell U^n, v) = \Phi_\ell(\boldsymbol{G}_n; \boldsymbol{U}_n, v) + \Psi_\ell(\boldsymbol{G}_n; \boldsymbol{W}_n, v) + \delta_{4\ell}(\zeta^n, v), \quad \text{for} \quad \ell = 1, 2, 3, 4, \tag{4.21a}$$

$$(Q^{n,\ell}, r) = \mathcal{K}(\boldsymbol{G}^{n,\ell}; U^{n,\ell}, r) \quad \text{for} \quad \ell = 1, 2, 3, \tag{4.21b}$$

where $\boldsymbol{G}^{n,\ell} = (U_a^\ell(t^n), U_b^\ell(t^n))$ is the reference boundary condition. Here $\delta_{4\ell}$ is the Kronecker symbol and ζ^n is the local truncation error in each step of the third order IMEX RK time-marching (4.19). Besides, there exists a bounding constant $C > 0$ depending on the regularity of U, independent of n, h and τ, such that

$$\|\zeta^n\| \le C\tau^4. \tag{4.22}$$

Proof. The proof is straightforward by the considered PDE and the definitions of the reference functions (4.19), so we omit it. Similar analysis can be found in [30, 31, 24]. $\qquad\square$

Theorem 4.2. *Let U be the exact solution of problem (2.1) which satisfies the smoothness assumption (4.20), and u^n is the numerical solution of scheme (2.25). There exists a positive constant τ_0 depending only on the convection and diffusion coefficients but not on h, such that if $\tau \leq \tau_0$, then there holds the following error estimate*

$$\max_{n\tau \leq T} \|U(t^n) - u^n\|^2 \leq e^{n\tau}\left[C(c^2 + d)\tau \sum_{m=0}^{n-1} h^{-1}|\boldsymbol{\theta}^m|^2 + C(h^{2k+2} + \tau^6)\right], \qquad (4.23)$$

where T is the final computing time and the bounding constant $C > 0$ is independent of h and τ, $|\boldsymbol{\theta}^m|^2 = \sum_{\ell=1}^{3}\left(|\theta_a^{m,\ell}|^2 + |\theta_b^{m,\ell}|^2\right)$ with $(\theta_a^{m,\ell}, \theta_b^{m,\ell}) = \boldsymbol{G}^{m,\ell} - \boldsymbol{g}^{m,\ell}$ the error due to the boundary treatment.

Remark 4.1. From this theorem we see that, optimal error estimates in both space and time can be obtained if the boundary errors $\theta_a^{n,\ell}, \theta_b^{n,\ell}$ are of order $\mathcal{O}(h^{1/2}\tau^3)$. It is not easy to satisfy this condition in practice, actually the order of boundary errors can be relaxed to $\mathcal{O}(\tau^3)$ in computation, see the numerical results in Section 3. For the boundary treatment strategy proposed in Section 3, we could not show the optimal error estimate in time, the main reason is that the inverse inequality (2.9) is used to deal with boundary terms.

4.2.2   Proof of Theorem 4.2

We will use two Gauss-Radau projections, from $H^1(\mathcal{T}_h)$ to V_h, denoted by π_h^- and π_h^+ respectively. For any function $p \in H^1(\mathcal{T}_h)$, the projection $\pi_h^{\pm}p$ is defined as the unique element in V_h such that, for any $j = 1, 2, \cdots, N$

$$(\pi_h^- p - p, v)_{I_j} = 0, \quad \forall v \in \mathcal{P}_{k-1}(I_j), \quad (\pi_h^- p)_j^- = p_j^-; \qquad (4.24a)$$

$$(\pi_h^+ p - p, v)_{I_j} = 0, \quad \forall v \in \mathcal{P}_{k-1}(I_j), \quad (\pi_h^+ p)_{j-1}^+ = p_{j-1}^+. \qquad (4.24b)$$

Denote by $\eta = p - \pi_h^{\pm}p$ the projection error. By a standard scaling argument [9], it is easy to obtain the following approximation property

$$\|\eta\| + h^{1/2}\|\eta\|_{\partial\mathcal{T}_h} \leq Ch^{\min(k+1,s)}\|p\|_{H^s(\Omega)}, \qquad (4.25)$$

where the bounding constant $C > 0$ is independent of h and p, $\|\cdot\|_{\partial\mathcal{T}_h} = \sqrt{\sum_j \|\cdot\|_{\partial I_j}^2}$.

At each stage time, we denote the error between the exact (reference) solution and the numerical solution by $\boldsymbol{e}^{n,\ell} = (e_u^{n,\ell}, e_q^{n,\ell}) = (U^{n,\ell} - u^{n,\ell}, Q^{n,\ell} - q^{n,\ell})$. As the standard treatment in finite element analysis, we would like to divide the error in the form $\boldsymbol{e} = \boldsymbol{\xi} - \boldsymbol{\eta}$, where

$$\boldsymbol{\eta} = (\eta_u, \eta_q) = (\pi_h^- U - U, \pi_h^+ Q - Q), \quad \boldsymbol{\xi} = (\xi_u, \xi_q) = (\pi_h^- U - u, \pi_h^+ Q - q), \qquad (4.26)$$

here we have dropped the superscripts n and ℓ for simplicity, and $\pi_h^{\pm}$ are Gauss-Radau projections defined in (4.24). Thus, owing to (2.7) and the definition of Gauss-Radau projections, we have

$$\mathcal{H}_{\text{int}}(\eta_u, v) = 0, \qquad \mathcal{K}_{\text{int}}(\eta_u, v) = 0, \qquad \mathcal{L}_{\text{int}}(\boldsymbol{\eta}, v) = \sqrt{d}\eta_{q,N}^- v_N^-, \qquad (4.27)$$

for any $v \in V_h$. And by the smoothness assumption (4.20), it follows from (4.25) and the linearity of the projections $\pi_h^{\pm}$ that the stage projection errors and their evolutions satisfy

$$\|\eta_u^{n,\ell}\| + \|\eta_q^{n,\ell}\| + h^{1/2}(|\eta_{u,N}^{n,\ell}| + |\eta_{q,N}^{n,\ell}|) \leq Ch^{k+1}, \tag{4.28a}$$

$$\|\mathbb{E}_{\ell+1}\eta_u^n\| \leq Ch^{k+1}\tau, \tag{4.28b}$$

for any n and $\ell = 0, 1, 2, 3$ under consideration.

In what follows we will focus our attention on the estimate of the error in the finite element space, say, $\boldsymbol{\xi} \in V_h \times V_h$. To this end, we need to set up the error equations about $\boldsymbol{\xi}^{n,\ell}$. Subtracting those variational forms in Lemma 4.1 from those in the scheme (4.2), in the same order, we obtain the following error equations

$$(\mathbb{E}_\ell \xi_u^n, v) = \Phi_\ell(\boldsymbol{G}_n; \boldsymbol{U}_n, v) - \Phi_\ell(\boldsymbol{g}_n; \boldsymbol{u}_n, v) + \Psi_\ell(\boldsymbol{G}_n; \boldsymbol{W}_n, v) - \Psi_\ell(\boldsymbol{g}_n; \boldsymbol{w}_n, v)$$
$$+ (\mathbb{E}_\ell \eta_u^n + \delta_{4\ell}\zeta^n, v), \quad \text{for} \quad \ell = 1, 2, 3, 4, \tag{4.29a}$$

$$(\xi_q^{n,\ell}, r) = \mathcal{K}(\boldsymbol{G}^{n,\ell}; U^{n,\ell}, r) - \mathcal{K}(\boldsymbol{g}^{n,\ell}; u^{n,\ell}, r) + (\eta_q^{n,\ell}, r), \quad \text{for} \quad \ell = 1, 2, 3. \tag{4.29b}$$

For the convenience of analysis, we denote $\boldsymbol{\theta}^{n,\ell} = \boldsymbol{G}^{n,\ell} - \boldsymbol{g}^{n,\ell} = (\theta_a^{n,\ell}, \theta_b^{n,\ell})$ as the error due to the boundary treatment. Then (4.29b) becomes

$$(\xi_q^{n,\ell}, r) = \mathcal{K}_{\text{bry}}(\boldsymbol{\theta}^{n,\ell}; r) + \mathcal{K}_{\text{int}}(\xi_u^{n,\ell}, r) + (\eta_q^{n,\ell}, r), \quad \text{for} \quad \ell = 1, 2, 3, \tag{4.30}$$

since $\mathcal{K}_{\text{int}}(\eta_u, r) = 0$ by (4.27).

Based on (4.30) and along the similar argument as Lemma 2.4 we get the following important relationship:

$$\|(\xi_u)_x\| + \sqrt{\mu h^{-1}}([\![\xi_u]\!]_{\text{int}} + |\xi_{u,0}^+|) \leq \frac{C_\mu}{\sqrt{d}}(\|\xi_q\| + \|\eta_q\|) + \sqrt{\mu h^{-1}}(|\xi_{u,N}^-| + |\theta_a| + |\theta_b|), \tag{4.31}$$

where we omit the superscript n, ℓ for notational simplicity.

Next we are going to estimate $\boldsymbol{\xi}$ by the energy method, whose procedure is similar as in the stability analysis. Taking $\tilde{v}_\ell = \xi_u^{n,1}, \xi_u^{n,2} - 2\xi_u^{n,1}, \xi_u^{n,3}$ and $2\xi_u^{n+1}$ in (4.29a), for $\ell = 1, 2, 3, 4$ respectively, and adding them together, we obtain the energy equation

$$\|\xi_u^{n+1}\|^2 - \|\xi_u^n\|^2 + \widetilde{\mathcal{S}} = \widetilde{\mathcal{T}_c} + \widetilde{\mathcal{T}_d} + \mathcal{T}_p, \tag{4.32}$$

where

$$\widetilde{\mathcal{S}} = \frac{1}{2}\left(\|\mathbb{E}_1\xi_u^n\|^2 + \|\mathbb{E}_2\xi_u^n\| + \|\mathbb{E}_{31}\xi_u^n\|^2 + \|\mathbb{E}_{32}\xi_u^n\|^2 + 2\|\mathbb{E}_4\xi_u^n\|^2\right). \tag{4.33}$$

$\widetilde{\mathcal{T}_c}$ and $\widetilde{\mathcal{T}_d}$ represent the error of the convection and diffusion parts, respectively, which are given as

$$\widetilde{\mathcal{T}_c} = \sum_{\ell=1}^{4} \Phi_\ell(\boldsymbol{G}_n; \boldsymbol{U}_n, \tilde{v}_\ell) - \Phi_\ell(\boldsymbol{g}_n; \boldsymbol{u}_n, \tilde{v}_\ell), \tag{4.34}$$

$$\widetilde{\mathcal{T}_d} = \sum_{\ell=1}^{4} \Psi_\ell(\boldsymbol{G}_n; \boldsymbol{W}_n, \tilde{v}_\ell) - \Psi_\ell(\boldsymbol{g}_n; \boldsymbol{w}_n, \tilde{v}_\ell). \tag{4.35}$$

$\mathcal{T}_{\mathrm{p}}$ is related to the projection errors which is given as

$$\mathcal{T}_{\mathrm{p}} = \sum_{\ell=1}^{4} (\mathbb{E}_\ell \eta_u^n + \delta_{4\ell} \zeta^n, \tilde{v}_\ell). \tag{4.36}$$

A simple use of the Cauchy-Schwarz inequality, the Young's inequality and (4.28) leads to

$$\mathcal{T}_{\mathrm{p}} \leq \varepsilon\tau \sum_{\ell=1}^{4} \|\tilde{v}_\ell\|^2 + \frac{C}{\tau} \sum_{\ell=1}^{4} \|\mathbb{E}_\ell \eta_u^n + \delta_{4\ell}\zeta^n\|^2 \leq \varepsilon\tau(\|\xi_u^n\|^2 + \widetilde{\mathcal{S}}) + C(h^{2k+2}\tau + \tau^7). \tag{4.37}$$

Next we take the term $\widetilde{\mathcal{T}_{\mathrm{d}}}^1 = \Psi_1(\boldsymbol{G}_n; \boldsymbol{W}_n, \xi_u^{n,1}) - \Psi_1(\boldsymbol{g}_n; \boldsymbol{w}_n, \xi_u^{n,1})$ as an example to illustrate the estimate for $\widetilde{\mathcal{T}_{\mathrm{d}}}$. Noticing that

$$\widetilde{\mathcal{T}_{\mathrm{d}}}^1 = \gamma\tau \left[\mathcal{L}_{\mathrm{int}}(\boldsymbol{\xi}^{n,1}, \xi_u^{n,1}) - \mathcal{L}_{\mathrm{int}}(\boldsymbol{\eta}^{n,1}, \xi_u^{n,1}) + \mathcal{L}_{\mathrm{bry}}(\boldsymbol{\theta}^{n,1}; \xi_u^{n,1}) \right]. \tag{4.38}$$

Owing to Corollary 2.1 and (4.30) we get

$$\begin{aligned}
\mathcal{L}_{\mathrm{int}}(\boldsymbol{\xi}, \xi_u) &= -\frac{d}{h}(\xi_{u,N}^-)^2 - \mathcal{K}_{\mathrm{int}}(\xi_u, \xi_q) \\
&= -\frac{d}{h}(\xi_{u,N}^-)^2 - \|\xi_q\|^2 + (\eta_q, \xi_q) + \mathcal{K}_{\mathrm{bry}}(\boldsymbol{\theta}; \xi_q),
\end{aligned} \tag{4.39}$$

where we have omitted the superscript n, ℓ for the simplicity of notations. Thus owing to (4.27) we have

$$\widetilde{\mathcal{T}_{\mathrm{d}}}^1 = -\frac{d}{h}\gamma\tau((\xi_{u,N}^{n,1})^-)^2 - \gamma\tau\|\xi_q^{n,1}\|^2 + V_1 + V_2, \tag{4.40}$$

where

$$\begin{aligned}
V_1 &= \gamma\tau(\eta_q^{n,1}, \xi_q^{n,1}) - \sqrt{d}\gamma\tau(\eta_{q,N}^{n,1})^-(\xi_{u,N}^{n,1})^-, \\
V_2 &= \gamma\tau\mathcal{K}_{\mathrm{bry}}(\boldsymbol{\theta}^{n,1}; \xi_q^{n,1}) + \gamma\tau\mathcal{L}_{\mathrm{bry}}(\boldsymbol{\theta}^{n,1}; \xi_u^{n,1}).
\end{aligned}$$

Proceeding with similar arguments for the remaining terms and similarly as the estimate for $\mathcal{T}_{\mathrm{d}}$ in the previous section, we can obtain

$$\begin{aligned}
\widetilde{\mathcal{T}_{\mathrm{d}}} &= -\frac{d}{h}\tau\boldsymbol{\xi}_{u,N}^{n\top}\mathbb{A}\boldsymbol{\xi}_{u,N}^n - \tau\int_\Omega \boldsymbol{\xi}_q^{n\top}\mathbb{A}\boldsymbol{\xi}_q^n\,\mathrm{d}x + V_p + V_b \\
&\leq -\frac{\gamma}{4}\tau\left(\frac{d}{h}|\boldsymbol{\xi}_{u,N}^n|^2 + \|\boldsymbol{\xi}_q^n\|^2\right) + V_p + V_b,
\end{aligned} \tag{4.41}$$

where $\boldsymbol{\xi}_{u,N}^n = ((\xi_{u,N}^{n,1})^-, (\xi_{u,N}^{n,2} - 2\xi_{u,N}^{n,1})^-, (\xi_{u,N}^{n,3})^-)^\top$ and $\boldsymbol{\xi}_q^n = (\xi_q^{n,1}, \xi_q^{n,2} - 2\xi_q^{n,1}, \xi_q^{n,3})^\top$. In (4.41), V_p and V_b are related to the projection errors and the boundary terms in the following forms:

$$V_p = -\sqrt{d}\tau\boldsymbol{\eta}_{q,N}^{n\top}\mathbb{A}\boldsymbol{\xi}_{u,N}^n + \tau\int_\Omega \boldsymbol{\eta}_q^{n\top}\mathbb{A}\boldsymbol{\xi}_q^n\,\mathrm{d}x, \tag{4.42}$$

$$V_b = \widetilde{\mathcal{T}_{\mathrm{d}}}^{\mathrm{bry}} + \widetilde{\mathcal{T}_{\mathrm{d}}}^{\mathrm{bry}'}, \tag{4.43}$$

where $\boldsymbol{\eta}_{q,N}^n = ((\eta_{q,N}^{n,1})^-, (\eta_{q,N}^{n,2} - 2\eta_{q,N}^{n,1})^-, (\eta_{q,N}^{n,3})^-)^\top$ and $\boldsymbol{\eta}_q^n = (\eta_q^{n,1}, \eta_q^{n,2} - 2\eta_q^{n,1}, \eta_q^{n,3})^\top$, $\widetilde{\mathcal{T}}_{\mathrm{d}}^{\mathrm{bry}}$ and $\widetilde{\mathcal{T}}_{\mathrm{d}}^{\mathrm{bry}'}$ are replacing $(\boldsymbol{g}^{n,\ell}, u^{n,\ell}, q^{n,\ell})$ with $(\boldsymbol{\theta}^{n,\ell}, \xi_u^{n,\ell}, \xi_q^{n,\ell})$ in $\mathcal{T}_{\mathrm{d}}^{\mathrm{bry}}$ and $\mathcal{T}_{\mathrm{d}}^{\mathrm{bry}'}$. Using of the Cauchy-Schwarz inequality, the Young's inequality and (4.28) leads to

$$V_p \le \varepsilon\tau \left(\frac{d}{h}|\boldsymbol{\xi}_{u,N}^n|^2 + \|\boldsymbol{\xi}_q^n\|^2 \right) + C\tau \left(\|\boldsymbol{\eta}_q^n\|^2 + h|\boldsymbol{\eta}_{q,N}^n|^2 \right)$$

$$\le \varepsilon\tau \left(\frac{d}{h}|\boldsymbol{\xi}_{u,N}^n|^2 + \|\boldsymbol{\xi}_q^n\|^2 \right) + Ch^{2k+2}\tau,$$

for arbitrary $\varepsilon > 0$. Owing to Lemma 2.3 and the Young's inequality we can derive

$$V_b \le \varepsilon\tau \left(\frac{d}{h}|\boldsymbol{\xi}_{u,N}^n|^2 + \|\boldsymbol{\xi}_q^n\|^2 \right) + Cdh^{-1}\tau|\boldsymbol{\theta}^n|^2,$$

where $|\boldsymbol{\theta}^n|^2 = \sum_{\ell=1}^3 \left(|\theta_a^{n,\ell}|^2 + |\theta_b^{n,\ell}|^2 \right)$. As a consequence,

$$\widetilde{\mathcal{T}}_{\mathrm{d}} \le \left(2\varepsilon - \frac{\gamma}{4} \right) \tau \left(\frac{d}{h}|\boldsymbol{\xi}_{u,N}^n|^2 + \|\boldsymbol{\xi}_q^n\|^2 \right) + Cdh^{-1}\tau|\boldsymbol{\theta}^n|^2 + Ch^{2k+2}\tau. \tag{4.44}$$

We are now in the position of estimating $\widetilde{\mathcal{T}}_{\mathrm{c}}$. Noting that

$$\mathcal{H}(\boldsymbol{G};\boldsymbol{U},v) - \mathcal{H}(\boldsymbol{g};\boldsymbol{u},v) = \mathcal{H}_{\mathrm{int}}(\xi_u, v) + \mathcal{H}_{\mathrm{bry}}(\boldsymbol{\theta};v),$$

since $\mathcal{H}_{\mathrm{int}}(\eta_u, v) = 0$ by (4.27). Hence proceeding along the similar line as the estimate for $\mathcal{T}_{\mathrm{c}}$ in the previous section, by replacing $u^{n,\ell}$ with $\xi_u^{n,\ell}$ and $\boldsymbol{g}^{n,\ell}$ with $\boldsymbol{\theta}^{n,\ell}$, we get

$$\widetilde{\mathcal{T}}_{\mathrm{c}} \le \varepsilon\tau \left[\|\xi_u^n\|^2 + \|\boldsymbol{\xi}_q^n\|^2 + \frac{d}{h}|\boldsymbol{\xi}_{u,N}^n|^2 \right] + C\frac{c^2}{d}\tau\widetilde{\mathcal{S}} + C(c^2 + d)h^{-1}\tau|\boldsymbol{\theta}^n|^2 + Ch^{2k+2}\tau, \tag{4.45}$$

where the last term $Ch^{2k+2}\tau$ is obtained by adopting the relationship (4.31).

Taking ε small enough such that $3\varepsilon \le \frac{\gamma}{4}$, and letting $\tau \le \tau_0$ for some $\tau_0 \propto \frac{d}{c^2}$, then owing to (4.32), (4.37), (4.44) and (4.45) we have

$$\|\xi_u^{n+1}\|^2 - \|\xi_u^n\|^2 \le \tau\|\xi_u^n\|^2 + C(c^2 + d)h^{-1}\tau|\boldsymbol{\theta}^n|^2 + C(h^{2k+2}\tau + \tau^7). \tag{4.46}$$

Then by the aid of the discrete Gronwall's inequality, we can derive

$$\|\xi_u^n\|^2 \le e^{n\tau} \left[\|\xi_u^0\|^2 + C(c^2 + d)\tau \sum_{m=0}^{n-1} h^{-1}|\boldsymbol{\theta}^m|^2 + C(h^{2k+2} + \tau^6) \right]. \tag{4.47}$$

Noting that $\xi_u^0 = 0$, and owing to (4.28) and the triangle inequality, we get the main error estimate (4.23).

Remark 4.2. We would like to end this section with a comment on extending our results to multi-dimensional convection-diffusion problems with Dirichlet BCs. We expect the same stability and error estimates to hold on Cartesian meshes with $\mathcal{Q}^k$ elements, where $\mathcal{Q}^k$ means the tensor product of polynomials of degree k in each variables. Namely, the scheme is stable under the time step $\tau \le \tau_0$, since the similar relationship given in Lemma 2.4 also hold in multi-dimensional case, one can refer to [23] for the detailed proof. Also, if we do not consider the errors caused by boundary remedies, then optimal error estimates can be proved for $\mathcal{Q}^k$ elements on Cartesian meshes, we omit the details to save space.

5   Concluding remarks

We discuss a fully-discrete IMEX-RK3-LDG scheme for the one-dimensional convection-diffusion problems with Dirichlet boundary conditions in this paper, where the order reduction phenomenon of the third order IMEX RK method is discussed and a proper boundary treatment strategy at each intermediate stage is proposed to recover the third order accuracy of the scheme. In addition, by suitably defining numerical flux at boundary, we establish an important relationship between the gradient and interface jump of the numerical solution with the independent numerical solution of the gradient and the given boundary conditions, by which we get the unconditional stability of the corresponding scheme. The results of this paper can also be extended to multi-dimensional convection-diffusion problems with Dirichlet boundary conditions. We will consider other boundary conditions in future work.

References

[1] S. ABARBANEL, D. GOTTLIEB AND M. H. CARPENTER, *On the removal of boundary errors caused by Runge-Kutta integration of nonlinear partial differential equations*, SIAM J. Sci. Comput., 17 (1996), pp. 777-782.

[2] I. ALONSO-MALLO, *Runge-Kutta methods without order reduction for linear initial boundary value problems*, Numer. Math., 91 (2002), pp. 577-603.

[3] U. M. ASCHER, S. J. RUUTH AND R. J. SPITERI, *Implicit-explicit Runge-Kutta methods for time-dependent partial differential equations*, Appl. Numer. Math., 25 (1997), pp. 151–167.

[4] F. BASSI AND S. REBAY, *A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier-Stokes equations*, J. Comput. Phys. 131 (1997), pp. 267–279.

[5] M. P. CALVO, J. DE FRUTOS AND J. NOVO, *Linearly implicit Runge-Kutta methods for advection-reaction-diffusion equations*, Appl. Numer. Math., 37 (2001), pp. 535–549.

[6] M. P. CALVO, C. PALENCIA, *Avoiding the order reduction of Runge-Kutta methods for linear initial boundary value problems*, Math. Comput., 71 (2001), pp. 1529-1543.

[7] M. H. CARPENTER, D. GOTTLIEB, S. ABARBANEL AND W. S. DON, *The theoretical accuracy of Runge-Kutta time discretizations for the initial boundary value problem: a study of the boundary error*, SIAM J. Sci. Comput., 16 (1995), pp. 1241-1252.

[8] P. CASTILLO, B. COCKBURN, D. SCHÖTZAU AND C. SCHWAB, *Optimal a priori error estimates for the hp-version of the local discontinuous Galerkin method for convection-diffusion problems*, Math. Comp., 71 (2001), pp. 455–478.

[9] P. G. CIARLET, *The Finite Element Method for Elliptic Problems*, North–Holland, Amsterdam, New York, 1978.

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

[11] B. Dong and C.-W. Shu, *Analysis of a local discontinuous Galerkin methods for linear time-dependent fourth-order problems*, SIAM J. Numer. Anal., 47 (2009), pp. 3240–3268.

[12] D. Gottlieb, M. Gunzburger and E. Tadmor, *On numerical boundary treatment of hyperbolic systems for finite difference and finite element methods*, SIAM J. Numer. Anal., 19 (1982), pp. 671–682.

[13] S. Gottlieb, C.-W. Shu and E. Tadmor, *Strong stability-preserving high-order time discretization methods*, SIAM Rev., 43 (2001), pp. 89–112.

[14] T. Koto, *IMEX Runge-Kutta schemes for reaction-diffusion equations*, J. Comput. Appl. Math., 215 (2008), pp. 182-195.

[15] J. Lu, J. Fang, S. Tan, C.-W. Shu and M. Zhang, *Inverse Lax-Wendroff procedure for numerical boundary conditions of convection-diffusion equations*, J. Comput. Phys., 317 (2016), pp. 276-300.

[16] L. Pareschi, G. Russo, *Implicit-explicit Runge-Kutta schemes for stiff systems of differential equations*, In: Trigiante, D. (ed.) Recent Trends in Numerical Analysis, Advances Theory Computational Mathematics, vol. 3, pp. 269-288. Nova Sci. Publ., Huntington, NY (2001)

[17] D. Pathria, *The correct formulation of intermediate boundary conditions for Runge-Kutta time integration of initial boundary value problems*, SIAM J. Sci. Comput., 18 (1997), pp. 1255-1266.

[18] J. M. Sanz-Serna, J. G. Verwer and W. H. Hundsdorfer, *Convergence and order reduction of Runge-Kutta schemes applied to evolutionary problems in partial differential equations*, Numer. Math., 50 (1986), pp. 405-418.

[19] C.-W. Shu and S. Osher, *Efficient implementation of essentially non-oscillatory shock capturing schemes*, J. Comput. Phys., 77 (1988), pp. 439-471.

[20] J. G. Verwer, *Convergence and order reduction of diagonably implicit Runge-Kutta schemes in the method of lines*, in Proc. Dundee Numerical Analysis Conference 1985, Pitman Research Notes in Math. 140 (1986), pp. 220-237.

[21] H. J. Wang, C.-W. Shu and Q. Zhang, *Stability and error estimates of local discontinuous Galerkin methods with implicit-explicit time-marching for advection-diffusion problems,* SIAM J. Numer. Anal., 53 (2015), 206-227.

[22] H. J. Wang, C.-W. Shu and Q. Zhang, *Stability analysis and error estimates of local discontinuous Galerkin methods with implicit-explicit time-marching for nonlinear convection-diffusion problems,* Appl. Math. Comput., 272 (2016), 237-258.

[23] H. J. WANG, S. P. WANG, Q. ZHANG AND C.-W. SHU, *Local discontinuous Galerkin methods with implicit-explicit time-marching for multi-dimensional convection-diffusion problems.* ESAIM: M^2AN, 50 (2016), pp. 1083–1105.

[24] H. J. WANG AND Q. ZHANG, *Error estimate on a fully discrete local discontinuous Galerkin method for linear convection-diffusion problem*, J. Comput. Math., 31 (2013), pp. 283-307.

[25] Y. H. XIA, Y. XU AND C.-W. SHU, *Efficient time discretization for local discontinuous Galerkin methods*, Discrete Contin. Dyn. Syst. Ser. B, 8 (2007), pp. 677–693.

[26] Y. XU AND C.-W. SHU, *Local discontinuous Galerkin methods for high-order time-dependent partial differential equations*, Commun. Comput. Phys., 7 (2010), pp. 1–46.

[27] J. YAN AND C.-W. SHU, *A local discontinuous Galerkin method for KdV type equations*, SIAM J. Numer. Anal., 40 (2002), pp. 769–791.

[28] J. YAN AND C.-W. SHU, *Local discontinuous Galerkin methods for partial differential equations with higher order derivatives*, J. Sci. Comput., 17 (2002), pp. 27–47.

[29] Q. ZHANG, *Third order explicit Runge-Kutta discontinuous Galerkin method for linear conservation law with inflow boundary condition*, J. Sci. Comput., 46 (2011), pp. 294-313.

[30] Q. ZHANG AND C.-W. SHU, *Error estimates to smooth solution of Runge-Kutta discontinuous Galerkin methods for scalar conservation laws*, SIAM J. Numer. Anal., 42 (2004), pp. 641-666.

[31] Q. ZHANG AND C.-W. SHU, *Stability analysis and a priori error estimates to the third order explicit Runge-Kutta discontinuous Galerkin method for scalar conservation laws*, SIAM J. Numer. Anal., 48 (2010), pp. 1038-1063.