Stability analysis and a priori error estimates to the third order explicit Runge-Kutta discontinuous Galerkin Method for scalar conservation laws

Qiang Zhang * Chi-Wang Shu †

September 16, 2009

ABSTRACT. In this paper we present the analysis for the Runge-Kutta discontinuous Galerkin (RKDG) method to solve scalar conservation laws, where the time discretization is the third order explicit total variation diminishing Runge–Kutta (TVDRK3) method. We use an energy technique to present the L^2-norm stability for scalar linear conservation laws, and obtain a priori error estimates for smooth solutions of scalar nonlinear conservation laws. Quasi-optimal order is obtained for general numerical fluxes, and optimal order is given for upwind fluxes. The theoretical results are obtained for piecewise polynomials with any degree $k \geq 1$ under the standard temporal-spatial CFL condition $\tau \leq \gamma h$, where h and τ, respectively, are the element length and time step, and the positive constant γ is independent of h and τ.

Key words. Discontinuous Galerkin method, finite element, explicit Runge-Kutta method, stability analysis, error estimate
AMS subject classification. 65M60, 65M12, 65M15

1 Introduction

In this paper we continue our work in [18] to consider the stability analysis and error estimates for the Runge-Kutta discontinuous Galerkin (RKDG) method for scalar conservation laws

$$u_t + \nabla \cdot f(u) = 0, \quad (x,t) \in \Omega \times (0,T], \tag{1.1a}$$

$$u(x,0) = u_0(x), \qquad x \in \Omega, \tag{1.1b}$$

where $f(u) = (f_1(u), \ldots, f_d(u))$ is the given convection flux function; here $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ and $u(x,t)$ is the unknown solution. We do not pay attention to boundary conditions in this paper; hence the solution is considered to be either periodic or compactly supported.

*Department of Mathematics, Nanjing University, Nanjing 210093, Jiangsu, P.R. China. E-mail: qzh@nju.edu.cn. The research of this author is supported by NSFC grant 10871093.

†Division of Applied Mathematics, Brown University, Providence, RI 02912, USA. E-mail: shu@dam.brown.edu. The research of this author is supported by NSF grant DMS-0809086 and DOE grant DE-FG02-08ER25863.

For simplicity of presentation, in most cases we will only give detailed analysis for the one-dimensional case; i.e., $\Omega = I = (0, 1)$ is the unit interval. We will, however, point out any differences, both in the analysis and in the results, for the multidimensional cases.

The first version of the DG method was introduced in 1973 by Reed and Hill [16], in the framework of neutron linear transport. It was later developed into RKDG methods by Cockburn et al. [5–9] for nonlinear hyperbolic conservation laws, which use a DG discretization in space and combine it with an explicit total variation diminishing Runge-Kutta (TVDRK) time-marching algorithm [17]. Later, this method was developed to solve equations with higher order derivatives, see [10] and [11] for more details. It is well known that the DG method has strong stability and optimal accuracy to capture discontinuous jumps and/or sharp transient layers, and it combines the advantages of finite element and finite difference methods.

However, up until now there has been relatively few work on stability analysis and error estimates for the fully discrete RKDG methods with explicit TVDRK time marching. The method of line version (continuous in time) of the DG scheme for linear equations has been considered in [10, 13, 14], and has been proved to maintain good L^2-norm stability and optimal error estimates. For nonlinear equations, there exists the well-known local entropy inequality [12] for the semi-discrete DG scheme, as well as for the fully discrete DG scheme with some special time-discretizations such as the backward Euler and Crank-Nilson algorithms. Recently, RKDG method with explicit time-discretization for nonlinear conservation laws has been analyzed in [18, 19], where the (quasi)-optimal a priori error estimates are obtained for the second order explicit TVDRK (TVDRK2) time discretization. However, there are two limitations of the results in [18, 19]. The first is that the a priori error estimates for smooth solutions were obtained without a proof of stability for general (possibly nonsmooth) solutions. In fact, up to now there has been no stability result for general solutions, possibly non-smooth, for fully discrete RKDG methods with explicit Runge-Kutta time stepping without using limiters, except for an indirect proof of linear stability via Fourier analysis which applies only to linear partial differential equations (PDEs) with uniform meshes and periodic boundary conditions. The second is that the error estimates in [18, 19] were obtained only for $\mathbb{P}^1$ elements under the standard CFL condition $\tau \leq \gamma h$, where h and τ, respectively, are the element length and time step, with a positive constant γ independent of h and τ. For $\mathbb{P}^k$ elements with $k > 1$, the results in [18, 19] were obtained under the much stronger time step restriction $\tau = o(h)$, which is reasonable for second order Runge-Kutta since it is linearly unconditional unstable when coupled with DG discretization of $\mathbb{P}^k$ elements with $k > 1$ [11].

The main purpose of this paper is to overcome the two limitations above by considering RKDG methods with a third order explicit TVDRK (TVDRK3) time discretization. The third order explicit RKDG algorithm is popular in practice, because it provides better linear stability and higher order accuracy in time. However, it is highly non-trivial to prove stability for general solutions and to obtain error estimates for smooth solutions for the RKDG method with TVDRK3 time-marching, the techniques used in [18, 19] for the TVDRK2 time-marching must be significantly changed. We start our study with a L^2-norm stability analysis for linear conservation laws without restriction to the smoothness of the solution. This result is not only important for the following error estimates for smooth solutions, but also significant in its own right, as it provides stability assurance for the fully discrete RKDG scheme for possibly discontinuous solutions. We then proceed to prove quasi-optimal error estimates for general monotone fluxes and optimal error estimate for upwind numerical fluxes. All these results are obtained under the standard temporal-spatial restriction $\tau \leq \gamma h$ for the piecewise polynomials with arbitrary degree $k \geq 1$, where $\gamma > 0$ is again a constant independent of h and τ. These results are then consistent with the well-known practical numerical evidence and linear stability results obtained by Fourier analysis, albeit with much more generality for non-uniform and unstructured meshes, for linear conservation laws with variable coefficient fluxes $f(u) = \beta(x,t)u$, and for possible h-p adaptivity. The main technical tool used in this paper is an energy analysis. The error estimates that we obtain in this paper are for smooth solutions of nonlinear conservation laws. A few technical details in dealing with the nonlinearity of the flux $f(u)$ are not fully provided and are referred to [18].

It is well known for the semi-discrete DG scheme that the L^2-norm stability result contains also a stability term involving the jumps of the numerical solution u_h across element interfaces. This reflects the subtle built-in numerical dissipation mechanism of the DG methods involving these jumps at element interfaces which allows the DG methods to be more accurate than the standard Galerkin methods [3]. This extra stability is fully explored in our proof of L^2-norm stability of the fully discrete RKDG schemes. We will also point out the very different stability mechanisms for the RKDG method with TVDRK2 and TVDRK3 time-marchings, respectively. For both of them, the numerical viscosity provided by the DG spatial discretization is the foundation for the fully-discrete stability. However, the RKDG method with TVDRK3 has an additional numerical viscosity in the time direction, stated in terms of the L^2-norm of a certain combination of the numerical solution in different Runge-Kutta stages (see Section 4). This extra numerical viscosity is the key ingredient in our analysis for obtaining stability under the standard CFL condition for the RKDG method with arbitrary polynomial degree and TVDRK3 time-marching, while the lack of it explains why stability under the standard CFL condition for the RKDG method can only be obtained for piecewise linear polynomials with TVDRK2 time-marching.

An outline of this paper is as follows. In Section 2 we present, for the equation (1.1), the fully discrete RKDG method with the explicit TVDRK3 time discretization. In Section 3 we present some preliminaries about the discontinuous finite element space. In Section 4 we give some elementary properties of the DG spatial discretization, and prove the corresponding L^2-norm stability, for the linear flux $f(u) = \beta u$ with a constant β. Section 5 is devoted to the error estimates for smooth solutions of nonlinear conservation laws. The main analysis is given to obtain quasi-optimal error estimates for high order piecewise polynomials with a general numerical flux, and a brief description is given to explain how to obtain optimal error estimates for upwind numerical fluxes. Concluding remarks are given in Section 6, and Section 7 is an appendix in which we give a short analysis on the L^2-norm stability of RKDG method with the TVDRK2 time marching, mainly to explain the very different stability mechanisms of TVDRK2 and TVDRK3.

2 Fully discrete RKDG scheme with TVDRK3 time-marching

We follow [7] and define the RKDG method with TVDRK3 time-marching, for the problem (1.1) in one space dimension. The multidimensional case is similar. For each partition of the interval $I = (0,1)$, $\{x_{j+1/2}\}_{j=0}^{N}$, we set $I_j = (x_{j-1/2}, x_{j+1/2})$, and $h_j = x_{j+1/2} - x_{j-1/2}$ for $j = 1, \ldots, N$; and we denote the quantities

$$h = \max_{1 \le j \le N} h_j, \quad \text{and} \quad \rho = \min_{1 \le j \le N} h_j. \tag{2.1}$$

For simplicity of presentation we would like to assume the ratio of h over ρ is upper bounded by a fixed positive constant ν^{-1} when h goes to zero, namely, the computational triangulation is quasi-uniform with $\nu h \le \rho \le h$.

For a given time step τ (which could actually change from step to step but is taken as a constant with respect to the time level n for simplicity), the solution of the scheme is denoted by $u_h^n(x) = u_h(x, n\tau)$, which belongs to the finite element space

$$V_h = V_h^k = \{\, v \in L^2(0,1) : v|_{I_j} \in \mathbb{P}^k(I_j), j = 1, \ldots, N \,\}, \tag{2.2}$$

where $\mathbb{P}^k(I_j)$ denotes the space of polynomials in I_j of degree at most $k \ge 1$. Note that the functions in V_h are allowed to have discontinuities across element interfaces. In this paper we do not consider the piecewise constant, namely the $k = 0$ case, since in this case the considered DG method is just the monotone finite volume method.

In what follows, we will consider the standard L^2-projection of a function $p \in L^2(0,1)$ into the finite element space V_h, denoted by $\mathbb{P}_h p$, which is defined as the unique function in V_h such that

$$\int_0^1 (\mathbb{P}_h p(x) - p(x)) v_h(x) \,\mathrm{d}x = 0 \quad \forall v_h \in V_h. \tag{2.3}$$

Due to the discontinuity of the finite element space, this projection is locally defined on each element I_j.

For simplicity of presentation, we will use the following notations. Parallel to the discontinuous finite element space V_h, we denote the broken Sobolev space as follows

$$H^{1,h}(\mathcal{T}_h) = \{\phi \in L^2(I) : \phi|_{I_j} \in H^1(I_j), j = 1, 2, \ldots, N\,\}, \tag{2.4}$$

where $\mathcal{T}_h$ is the union of all cells I_j in the partition. For any function $p \in H^{1,h}(\mathcal{T}_h)$, at each element boundary point there are two limits from different directions, namely, the left-value p^- and the right-value p^+. Furthermore, the jump and the mean at the element boundary point, respectively, are denoted by $[\![p]\!] = p^+ - p^-$, and $\{\!\{p\}\!\} = \frac{1}{2}(p^+ + p^-)$.

Following [18], we would like to give an abbreviate notation for the DG spatial operator. For any functions p and q in $H^{1,h}(\mathcal{T}_h)$, on each element I_j we define

$$\mathcal{H}_j(p,q) = \int_{I_j} f(p) q_x(x) \,\mathrm{d}x - \hat{f}(p)_{j+\frac{1}{2}} q(x_{j+\frac{1}{2}}^-) + \hat{f}(p)_{j-\frac{1}{2}} q(x_{j-\frac{1}{2}}^+), \tag{2.5}$$

where $\hat{f}(p) \equiv \hat{f}(p^-, p^+)$ is a given monotone numerical flux that depends on the two values of the function p at the element interface point. The numerical flux $\hat{f}(a,b)$ is locally Lipschitz continuous with regard to both arguments, and is consistent with the flux $f(p)$, namely, $\hat{f}(p,p) = f(p)$. Furthermore, it is a nondecreasing function of its first argument and a nonincreasing function of its second argument. The best-known examples of monotone numerical fluxes are the Godunov flux, the Engquist-Osher flux, the Lax-Friedrichs flux, etc. Some of these monotone fluxes (e.g. the Godunov flux and the Engquist-Osher flux) are upwind fluxes, namely $\hat{f}(a,b) = f(a)$ if $f'(u) \geq 0$ when u is between a and b, and $\hat{f}(a,b) = f(b)$ if $f'(u) \leq 0$ when u is between a and b. For more details, see, for example, [15].

There are three steps in the construction of a RKDG scheme. First, we multiply a test function v_h on both sides of the equation (1.1a), and integrate them by parts in each element I_j. Next, we define the suitable numerical flux at the element interface point, and get the semi-discrete DG method. Finally we adopt the Runge-Kutta type time discretization. For more details, see, for example, [7, 11].

In this paper the considered scheme is the fully discrete RKDG method coupled with the explicit TVDRK3 time marching. First, we set the initial value $u_h^0 = \mathbb{P}_h u_0(x)$. Then for each $n \geq 0$, the approximate solution from the time $n\tau$ to the next time $(n+1)\tau$ is defined as follows: find $u_h^{n,1}, u_h^{n,2}$ and u_h^{n+1} in the finite element space V_h, such that for any $v_h \equiv v_h(x) \in V_h$ and $1 \leq j \leq N$, on each element I_j there hold

$$\int_{I_j} u_h^{n,1} v_h \,\mathrm{d}x = \int_{I_j} u_h^n v_h \,\mathrm{d}x + \tau \mathcal{H}_j(u_h^n, v_h), \tag{2.6a}$$

$$\int_{I_j} u_h^{n,2} v_h \,\mathrm{d}x = \frac{3}{4}\int_{I_j} u_h^n v_h \,\mathrm{d}x + \frac{1}{4}\int_{I_j} u_h^{n,1} v_h \,\mathrm{d}x + \frac{\tau}{4}\mathcal{H}_j(u_h^{n,1}, v_h), \tag{2.6b}$$

$$\int_{I_j} u_h^{n+1} v_h \,\mathrm{d}x = \frac{1}{3}\int_{I_j} u_h^n v_h \,\mathrm{d}x + \frac{2}{3}\int_{I_j} u_h^{n,2} v_h \,\mathrm{d}x + \frac{2\tau}{3}\mathcal{H}_j(u_h^{n,2}, v_h). \tag{2.6c}$$

This is an explicit time-marching method when a local orthogonal basis is chosen for polynomials on I_j or when a small local mass matrix on I_j is inverted.

3 Preliminaries

In this section we would like to present some properties of the discontinuous finite element space V_h, which will be used in the stability analysis and error estimates. In this paper we use C (possibly with subscripts) to denote a positive constant depending solely on the exact solution, which may have a different value in each occurrence.

The usual notation of norms in Sobolev spaces will be used. For any integer $s \geq 0$, let $H^s(\Omega)$ represent the well-known Sobolev space equipped with the norm $\|\cdot\|_s$, which consists of functions with (distributional) derivatives of order not greater than s in $L^2(\Omega)$. Next, let the scalar inner product on $L^2(\Omega)$ be denoted by $(\cdot,\cdot)_\Omega$, and the associated norm by $\|\cdot\|_\Omega$. Furthermore, let $\|\cdot\|_{\infty,\Omega}$ represent the norm on $L^\infty(\Omega)$. If $\Omega = I$ we omit this subscript. See Adams [1] for more details.

3.1 Projections to the finite element space

As a standard trick in DG analysis, two types of projections, denoted by $\mathbb{Q}_h p(\cdot,t)$, are used in this paper. One is the standard L^2-projection $\mathbb{P}_h$, and the other is the Gauss-Radau projection $\mathbb{R}_h^\pm$. They will be used to get error estimates of the quasi-optimal and optimal order, respectively.

The standard L^2-projection $\mathbb{P}_h$ has been defined in Section 2; see (2.3). In what follows we define two kinds of the Gauss-Radau projection $\mathbb{R}_h^\pm$, corresponding to the positive direction and the negative direction of the x-axis, respectively. For any function $p \in H^1(0,1)$,

$\mathbb{R}_h^{\pm} p$ is defined element by element as the unique function in V_h, such that for any $1 \leq j \leq N$,

$$\int_{I_j} \mathbb{R}_h^{\pm} p(x) v_h(x)\,\mathrm{d}x = \int_{I_j} p(x) v_h(x)\,\mathrm{d}x, \quad \forall\, v_h(x) \in \mathbb{P}^{k-1}(I_j), \tag{3.1}$$

and the exact collocation at one endpoint, namely $(\mathbb{R}_h^+ p)^-_{j+1/2} = p^-_{j+1/2}$ for the positive projection $\mathbb{R}_h^+$, and $(\mathbb{R}_h^- p)^+_{j-1/2} = p^+_{j-1/2}$ for the negative projection $\mathbb{R}_h^-$, respectively.

3.2 Projection properties

Denote by $\eta = p(x) - \mathbb{Q}_h p(x)$ the projection error. By a standard scaling argument [2, 14], it is easy to obtain, for both projections, that,

$$\|\eta\| + h\|\eta_x\| + h^{1/2}\|\eta\|_{\Gamma_h} \leq C h^{k+1} \tag{3.2}$$

if $p(x)$ is smooth enough, where C is a positive constant independent of h. This constant depends on $\|p\|_{k+1}$ for the standard L^2-projection, and on $\|p\|_{k+2}$ for the Gauss-Radau projection. Here Γ_h is the union of all element interface points, and the L^2-norm on Γ_h is defined by

$$\|\eta\|_{\Gamma_h} = \Big[\sum_{1\leq j\leq N} ((\eta^+_{j+\frac{1}{2}})^2 + (\eta^-_{j+\frac{1}{2}})^2) \Big]^{1/2}. \tag{3.3}$$

It is worthy to point out that on each element interface point the Gauss-Radau projection $\mathbb{R}_h^{\pm}$ has the estimate (3.2), and furthermore, it is the exact collocation along one of the two directions. Namely $\eta^-_{j+1/2} = 0$ for $\mathbb{R}_h^+$, and $\eta^+_{j+1/2} = 0$ for $\mathbb{R}_h^-$, respectively. This property is very helpful to obtain the optimal error estimates.

3.3 Inverse properties

Finally, we list some inverse properties of the finite element space V_h. For any $v_h \in V_h$, there exist positive constants $\mu_i, (i = 1, 2, 3)$, independent of v_h and h, such that

(i) $\|v_{h,x}\| \leq \mu_1 h^{-1}\|v_h\|$; (ii) $\|v_h\|_{\Gamma_h} \leq \mu_2 h^{-1/2}\|v_h\|$; (iii) $\|v_h\|_\infty \leq \mu_3 h^{-1/2}\|v_h\|$.

For more details of these inverse properties, we refer the reader to [2]. The inverse inequality (iii) will not be used in the analysis for linear conservation laws. In the following we will denote $\mu = \max\{\mu_1, (\mu_2)^2\}$, which increases with the degree of polynomials.

4 Stability analysis for the linear conservation law

In this section we are going to obtain the L^2-norm stability for the fully discrete RKDG method with TVDRK3 time-marching, where $f(u) = \beta u$ with the given constant β. The analysis in this section forms the foundation of all results in this paper.

4.1 DG spatial discretization for the linear flux

For the linear flux $f(u) = \beta u$, the DG spatial discrete operator is given as

$$\mathcal{H}_j(\phi, \psi) = \int_{I_j} \beta\phi\psi_x \mathrm{d}x - \hat{f}(\phi)_{j+1/2}\psi^-_{j+1/2} + \hat{f}(\phi)_{j-1/2}\psi^+_{j-1/2}, \tag{4.1}$$

for any ϕ and ψ in $H^{1,h}(\mathcal{T}_h)$. Note that in this case, all the monotone numerical fluxes mentioned in Section 2 and used in practice coincide and are equal to an upwind numerical flux

$$\hat{f}(\phi) = \hat{f}(\phi^-, \phi^+) = \begin{cases} \beta\phi^-, & \text{if } \beta > 0 \\ \beta\phi^+, & \text{if } \beta < 0. \end{cases} \tag{4.2}$$

It is convenient to consider the summation of (4.1) over all j, which is denoted by $\mathcal{H}(\phi, \psi)$. The periodic boundary condition yields that

$$\mathcal{H}(\phi, \psi) = \sum_{1\le j\le N} \mathcal{H}_j(\phi, \psi) = \sum_{1\le j\le N} \Big[\hat{f}(\phi)_{j+1/2}[\![\psi]\!]_{j+1/2} + \int_{I_j} \beta\phi\psi_x \mathrm{d}x\Big]. \tag{4.3}$$

Below we present some elementary properties of this bilinear operator.

The next lemma describes the continuity in the L^2-norm of this operator, with the bounding constant depending on the spatial mesh size h.

Lemma 4.1 *For any ψ and ϕ in V_h, we have $|\mathcal{H}(\phi, \psi)| \le (\sqrt{2} + 1)|\beta|\mu h^{-1}\|\phi\|\|\psi\|$.*

Proof. This proof is straightforward. First we point out that

$$\sum_{1\le j\le N} [\![\psi]\!]^2_{j+\frac{1}{2}} \le 2 \sum_{1\le j\le N} \Big[(\psi^+_{j+\frac{1}{2}})^2 + (\psi^+_{j+\frac{1}{2}})^2\Big] = 2\|\psi\|^2_{\Gamma_h}, \tag{4.4}$$

and $\sum_{1\le j\le N} |\hat{f}(\phi)_{j+\frac{1}{2}}|^2 \le |\beta|^2\|\phi\|^2_{\Gamma_h}$ owing to the definition (4.2). By using Schwartz inequality, together with the inverse properties (i) and (ii), we have

$$\begin{aligned}
|\mathcal{H}(\phi, \psi)| &\le \Big|\int_I \beta\phi\psi_x\,\mathrm{d}x\Big| + \sum_{1\le j\le N} |\hat{f}(\phi)_{j+\frac{1}{2}}| \cdot |[\![\psi]\!]_{j+\frac{1}{2}}| \\
&\le |\beta|\|\phi\|\|\psi_x\| + \sqrt{2}|\beta|\|\phi\|_{\Gamma_h}\|\psi\|_{\Gamma_h} \\
&\le \mu_1 h^{-1}|\beta|\|\phi\|\|\psi\| + \sqrt{2}(\mu_2)^2 h^{-1}|\beta|\|\phi\|\|\psi\| \\
&\le (\sqrt{2} + 1)|\beta|\mu h^{-1}\|\phi\|\|\psi\|,
\end{aligned}$$

since $\mu = \max\{\mu_1, (\mu_2)^2\}$. This completes the proof. $\square$

The next lemma describes the approximate anti-symmetry property of the operator $\mathcal{H}(\psi, \phi)$, which helps us to obtain the L^2-norm stability under the standard temporal-spatial CFL condition. It also implies the negative semi-definiteness of this operator.

Lemma 4.2 *For any ϕ and ψ in $H^{1,h}(\mathcal{T}_h)$, we have*

$$\mathcal{H}(\psi, \phi) + \mathcal{H}(\phi, \psi) = -\sum_{1\le j\le N} |\beta|[\![\phi]\!]_{j+\frac{1}{2}} \cdot [\![\psi]\!]_{j+\frac{1}{2}}, \tag{4.5a}$$

$$\mathcal{H}(\phi, \phi) = -\frac{1}{2}\sum_{1\le j\le N} |\beta|[\![\phi]\!]^2_{j+1/2}. \tag{4.5b}$$

Proof. Obviously, (4.5b) is a corollary of the identity (4.5a), if $\psi = \phi$. So we only prove (4.5a) below. In fact, from (4.3) we can easily work out the details as follows

$$
\begin{aligned}
&\mathcal{H}(\psi,\phi) + \mathcal{H}(\phi,\psi) \\
=& \sum_{1\le j\le N} \left[\hat{f}(\phi)_{j+1/2}[\![\psi]\!]_{j+1/2} + \hat{f}(\psi)_{j+1/2}[\![\phi]\!]_{j+1/2}\right] + \sum_{1\le j\le N}\int_{I_j} \beta(\phi\psi)_x \,\mathrm{d}x \\
=& \sum_{1\le j\le N} \left[\hat{f}(\phi)_{j+1/2}[\![\psi]\!]_{j+1/2} + \hat{f}(\psi)_{j+1/2}[\![\phi]\!]_{j+1/2} - \beta[\![\phi\psi]\!]_{j+1/2}\right] \\
=& \sum_{1\le j\le N} \left[(\hat{f}(\phi)_{j+1/2} - \beta\phi^+_{j+\frac12})[\![\psi]\!]_{j+1/2} + (\hat{f}(\psi)_{j+1/2} - \beta\psi^-_{j+\frac12})[\![\phi]\!]_{j+1/2}\right],
\end{aligned}
\tag{4.6}
$$

where we have used the periodic boundary condition for the second equality, and used $[\![\phi\psi]\!] = \phi^+[\![\psi]\!] + \psi^-[\![\phi]\!]$ for the last equality.

Noticing the definition of the numerical flux, (4.2), we get that

$$
\hat{f}(\phi) - \beta\phi^+ = \begin{cases} -\beta[\![\phi]\!], & \text{if } \beta > 0 \\ 0, & \text{if } \beta < 0. \end{cases}, \quad \hat{f}(\psi) - \beta\psi^- = \begin{cases} 0, & \text{if } \beta > 0 \\ \beta[\![\psi]\!], & \text{if } \beta < 0. \end{cases}
$$

Hence we clearly get (4.5a) from (4.6). □

4.2 Time-marching in TVDRK3

In the RKDG method with TVDRK3 time-marching, for linear conservation laws, the approximate time derivatives up to the third order at time $t = t^n$ are proportional to certain linear combinations of the numerical solution in different Runge-Kutta stages. To be more specific, we denote

$$
\mathbb{D}_1^n = u_h^{n,1} - u_h^n, \quad \mathbb{D}_2^n = 2u_h^{n,2} - u_h^{n,1} - u_h^n, \quad \mathbb{D}_3^n = u_h^{n+1} - 2u_h^{n,2} + u_h^n, \tag{4.7}
$$

and set up the following lemma, where the spatial DG operator $\mathcal{H}(\cdot,\cdot)$ can be considered as the discrete approximation of the time derivative because of the PDE.

Lemma 4.3 *For the fully discrete DG method* (2.6) *with the explicit TVDRK3 time marching, we have the following identities*

$$
(\mathbb{D}_1^n, v_h) = \tau\mathcal{H}(u_h^n, v_h), \quad \forall\, v_h \in V_h; \tag{4.8a}
$$

$$
(\mathbb{D}_2^n, v_h) = \frac{\tau}{2}\mathcal{H}(\mathbb{D}_1^n, v_h), \quad \forall\, v_h \in V_h; \tag{4.8b}
$$

$$
(\mathbb{D}_3^n, v_h) = \frac{\tau}{3}\mathcal{H}(\mathbb{D}_2^n, v_h), \quad \forall\, v_h \in V_h. \tag{4.8c}
$$

Proof. It is straightforward to get (4.8a) by summing (2.6a) over all elements. It is also straightforward to get (4.8b) by subtracting half of (2.6a) from twice of (2.6b), and summing over all elements.

To prove (4.8c), we first substitute (2.6a) into (2.6b). Summing it over all elements gives $(u_h^{n,2} - u_h^n, v_h) = \frac{\tau}{4}\mathcal{H}(u_h^{n,1} + u_h^n, v_h)$, for any $v_h \in V_h$. Subtracting $\frac43$ of this equality from

(2.6c) yields,

$$
\begin{aligned}
(\mathbb{D}_3^n, v_h) &= \big(u_h^{n+1} - \frac{1}{3}u_h^n - \frac{2}{3}u_h^{n,2}, v_h\big) - \frac{4}{3}\big(u_h^{n,2} - u_h^n, v_h\big)\\
&= \frac{2\tau}{3}\mathcal{H}(u_h^{n,2}, v_h) - \frac{\tau}{3}\mathcal{H}(u_h^{n,1} + u_h^n, v_h)\\
&= \frac{\tau}{3}\mathcal{H}(2u_h^{n,2} - u_h^{n,1} - u_h^n, v_h) = \frac{\tau}{3}\mathcal{H}(\mathbb{D}_2^n, v_h), \quad \forall\, v_h \in V_h,
\end{aligned}
$$

which is just (4.8c). So we have completed the proof of this lemma. □

4.3 Stability for linear conservation laws

Using the above elementary lemmas, we will obtain in this subsection, for linear conservation laws, the L^2-norm stability result for the RKDG method with TVDRK3 time-marching.

Theorem 4.1 (stability) *Assume $f(u) = \beta u$ is a linear flux. Let u_h be the numerical solution of the fully discrete scheme* (2.6) *with the explicit TVDRK3 time discretization, where the finite element space V_h is of piecewise polynomials with arbitrary degree $k \geq 1$, defined on any regular triangulations of $I = (0, 1)$. Then we have for any n, that*

$$
\|u_h^{n+1}\| \leq \|u_h^n\|, \tag{4.9}
$$

under the CFL condition

$$
\mu|\beta|\tau h^{-1} \leq \frac{1}{2}, \tag{4.10}
$$

where μ is the inverse constant defined in Section 3.3.

Proof. We take the test function v_h as u_h^n, $4u_h^{n,1}$ and $6u_h^{n,2}$ in the schemes (2.6a), (2.6b) and (2.6c), respectively, and sum them up to obtain the following identity

$$
\tau\Big[\mathcal{H}(u_h^n, u_h^n) + \mathcal{H}(u_h^{n,1}, u_h^{n,1}) + 4\mathcal{H}(u_h^{n,2}, u_h^{n,2})\Big] = \int_I \mathbb{T}\,\mathrm{d}x, \tag{4.11}
$$

in which $\mathbb{T}$ is given by

$$
\begin{aligned}
\mathbb{T} =& -2u_h^n u_h^{n,1} - (u_h^n)^2 + 4u_h^{n,1}u_h^{n,2} - (u_h^{n,1})^2 + 6u_h^{n,2}u_h^{n+1} - 2u_h^n u_h^{n,2} - 4(u_h^{n,2})^2\\
=& \,3\big[(u_h^{n+1})^2 - (u_h^n)^2\big] - (2u_h^{n,2} - u_h^{n,1} - u_h^n)^2 - 3(u_h^{n+1} - u_h^n)(u_h^{n+1} - 2u_h^{n,2} + u_h^n).
\end{aligned}
$$

It implies the following energy equality

$$
\begin{aligned}
3\|u_h^{n+1}\|^2 - 3\|u_h^n\|^2 =& \,\tau\Big[\mathcal{H}(u_h^n, u_h^n) + \mathcal{H}(u_h^{n,1}, u_h^{n,1}) + 4\mathcal{H}(u_h^{n,2}, u_h^{n,2})\Big]\\
&+ \|2u_h^{n,2} - u_h^{n,1} - u_h^n\|^2 + 3(u_h^{n+1} - u_h^n, u_h^{n+1} - 2u_h^{n,2} + u_h^n),
\end{aligned}
\tag{4.12}
$$

where each line on the right-hand side will be denoted by Π_1 and Π_2, respectively. This identity plays an important role for obtaining the L^2-norm stability.

By using Lemma 4.2, it is easy to express the first term Π_1 as

$$
\Pi_1 = -\frac{\tau|\beta|}{2}\sum_{1\leq j\leq N}\Big[[\![u_h^n]\!]_{j+\frac{1}{2}}^2 + [\![u_h^{n,1}]\!]_{j+\frac{1}{2}}^2 + 4[\![u_h^{n,2}]\!]_{j+\frac{1}{2}}^2\Big]. \tag{4.13}
$$

However, we have to cope with the second term Π_2 carefully, in order to show that Π_2 does not exceed the absolute value of Π_1 under a suitable CFL condition.

To do that, we express Π_2 by $\mathbb{D}_i^n, i = 1, 2, 3$. Since $u_h^{n+1} - u_h^n = \mathbb{D}_1^n + \mathbb{D}_2^n + \mathbb{D}_3^n$, there holds the following equivalent representation

$$\Pi_2 = (\mathbb{D}_2^n, \mathbb{D}_2^n) + 3(\mathbb{D}_3^n, \mathbb{D}_1^n) + 3(\mathbb{D}_3^n, \mathbb{D}_2^n) + 3(\mathbb{D}_3^n, \mathbb{D}_3^n) = \Lambda_1 + \Lambda_2 + \Lambda_3 + \Lambda_4. \tag{4.14}$$

Below we will estimate each term in (4.14) separately.

First we estimate the sum of Λ_1 and Λ_2, which will help us to get rid of the negative contribution from the integration in each element. The estimate reads

$$\begin{aligned}
\Lambda_1 + \Lambda_2 &= -(\mathbb{D}_2^n, \mathbb{D}_2^n) + 2(\mathbb{D}_2^n, \mathbb{D}_2^n) + 3(\mathbb{D}_3^n, \mathbb{D}_1^n) \\
&= -\|\mathbb{D}_2^n\|^2 + \tau\mathcal{H}(\mathbb{D}_1^n, \mathbb{D}_2^n) + \tau\mathcal{H}(\mathbb{D}_2^n, \mathbb{D}_1^n) \\
&= -\|\mathbb{D}_2^n\|^2 - \tau \sum_{1\le j\le N} |\beta| [\![\mathbb{D}_1^n]\!]_{j+\frac{1}{2}} [\![\mathbb{D}_2^n]\!]_{j+\frac{1}{2}} \\
&\le -\|\mathbb{D}_2^n\|^2 + \frac{\tau}{4} \sum_{1\le j\le N} |\beta| [\![\mathbb{D}_1^n]\!]^2_{j+\frac{1}{2}} + \tau \sum_{1\le j\le N} |\beta| [\![\mathbb{D}_2^n]\!]^2_{j+\frac{1}{2}}
\end{aligned} \tag{4.15}$$

where for the second equality we have used the identities (4.8b) and (4.8c) in Lemma 4.3, with different test functions $v_h = \mathbb{D}_2^n$ and $v_h = \mathbb{D}_1^n$, respectively; the third equality is a direct application of (4.5a) in Lemma 4.2; and the last inequality is a simple application of Young's inequality.

Identities (4.8c) and (4.5b), in Lemma 4.3 and Lemma 4.2, respectively, yield the estimate to the third term Λ_3. It reads

$$\Lambda_3 = 3(\mathbb{D}_3^n, \mathbb{D}_2^n) = \tau\mathcal{H}(\mathbb{D}_2^n, \mathbb{D}_2^n) = -\frac{\tau}{2} \sum_{1\le j\le N} |\beta| [\![\mathbb{D}_2^n]\!]^2_{j+\frac{1}{2}}. \tag{4.16}$$

To estimate the last term Λ_4, we use again identity (4.8c) in Lemma 4.3. It follows from Lemma 4.1 that

$$\|\mathbb{D}_3^n\|^2 = (\mathbb{D}_3^n, \mathbb{D}_3^n) = \frac{\tau}{3}\mathcal{H}(\mathbb{D}_2^n, \mathbb{D}_3^n) \le \frac{\sqrt{2}+1}{3} |\beta| \mu\tau h^{-1} \|\mathbb{D}_2^n\| \|\mathbb{D}_3^n\|.$$

Denote $\mathbb{a} = (2\sqrt{2}+3)/3$ and $\lambda_{\max} = \mu|\beta|\tau h^{-1}$. Consequently, the above inequality implies

$$\Lambda_4 = 3\|\mathbb{D}_3^n\|^2 \le \frac{2\sqrt{2}+3}{3} [\mu|\beta|\tau h^{-1}]^2 \|\mathbb{D}_2^n\|^2 \equiv \mathbb{a}\lambda_{\max}^2 \|\mathbb{D}_2^n\|^2. \tag{4.17}$$

Now we collect the estimates above from (4.15) to (4.17), and get that

$$\begin{aligned}
\Pi_2 &\le \frac{\tau}{4} \sum_{1\le j\le N} |\beta| [\![\mathbb{D}_1^n]\!]^2_{j+\frac{1}{2}} + \frac{\tau}{2} \sum_{1\le j\le N} |\beta| [\![\mathbb{D}_2^n]\!]^2_{j+\frac{1}{2}} - (1 - \mathbb{a}\lambda_{\max}^2) \|\mathbb{D}_2^n\|^2 \\
&\le \frac{\tau}{2} \sum_{1\le j\le N} |\beta| \Big[[\![u_h^n]\!]^2_{j+\frac{1}{2}} + [\![u_h^{n,1}]\!]^2_{j+\frac{1}{2}} \Big] + \tau|\beta| \|\mathbb{D}_2^n\|^2_{\Gamma_h} - (1 - \mathbb{a}\lambda_{\max}^2) \|\mathbb{D}_2^n\|^2 \\
&\le \frac{\tau}{2} \sum_{1\le j\le N} |\beta| \Big[[\![u_h^n]\!]^2_{j+\frac{1}{2}} + [\![u_h^{n,1}]\!]^2_{j+\frac{1}{2}} \Big] + |\beta|(\mu_2)^2 \tau h^{-1} \|\mathbb{D}_2^n\|^2 - (1 - \mathbb{a}\lambda_{\max}^2) \|\mathbb{D}_2^n\|^2 \\
&\le \frac{\tau}{2} \sum_{1\le j\le N} |\beta| \Big[[\![u_h^n]\!]^2_{j+\frac{1}{2}} + [\![u_h^{n,1}]\!]^2_{j+\frac{1}{2}} \Big] - (1 - \mathbb{a}\lambda_{\max}^2 - \lambda_{\max}) \|\mathbb{D}_2^n\|^2,
\end{aligned}$$

where we have used the inequality (4.4) for the second inequality, and the inverse inequality (ii) for the third inequality. The last inequality is obvious since $(\mu_2)^2 \leq \mu$, see Section 3.3.

We now substitute the estimates about Π_1 and Π_2 into the energy identity (4.12). Note that the temporal-spatial restriction (4.10) implies $1 - \mathbb{a}\lambda_{\max}^2 - \lambda_{\max} > 0$. Finally, under this CFL condition, we obtain the following inequality

$$3\|u_h^{n+1}\|^2 - 3\|u_h^n\|^2 + 2\tau \sum_{1\leq j\leq N} |\beta| [\![u_h^{n,2}]\!]_{j+\frac{1}{2}}^2 \leq 0. \tag{4.18}$$

This finishes the proof. □

Remark 4.1. We have proved that the fully discrete RKDG scheme with TVDRK3 time marching does not destroy the L^2-norm stability of the semi-discrete DG method. The proof depends strongly on the combination of numerical solutions in different stages, and the approximate anti-symmetry property (4.5a) in Lemma 4.2.

For the RKDG scheme, the TVDRK3 time marching has a different mechanism to ensure the L^2-norm stability, in comparison with TVDRK2. For both types of time-marching, the square of jumps over all element boundary points plays an important role. However, TVDRK3 has an additional numerical stability in terms of the L^2-norm of $\mathbb{D}_2^n$, which is lacking for TVDRK2. This is the reason that stability for RKDG with TVDRK2 can be proved only for piecewise linear polynomials, while with TVDRK3 stability holds for piecewise polynomials with arbitrary degree. In order to highlight this point, we provide the L^2-norm stability analysis for the RKDG method with TVDRK2 in the appendix.

Remark 4.2. Even though we proved Theorem 4.1 only for the specific third order TVD Runge-Kutta time discretization (2.6), the stability result actually holds for all three stage, third order Runge-Kutta time discretization. This is because for linear, constant coefficient ordinary differential equations, all three stage, third order Runge-Kutta methods are equivalent.

Remark 4.3. The analysis above and the stability results can be easily generalized to multi-dimensional linear conservation laws, including those with varying coefficients $\beta = \beta(x)$, on arbitrary regular triangulations with either periodic or other well-posed boundary conditions. However, it does not seem easy to extend the analysis to nonlinear flux functions $f(u)$. One technical difficulty is that the anti-symmetry property in Lemma 4.2 does not hold for the nonlinear case.

5 A priori error estimate

In this section we carry out an a priori error estimate for the fully-discrete RKDG scheme with the explicit TVDRK3 time marching for smooth solutions.

We assume the exact solution of the conservation law (1.1) is smooth enough, for example, $\|u\|_{k+1}$ and $\|u_t\|_{k+1}$ are bounded uniformly for any time $t \in [0, T]$. Moreover, u itself and its spatial derivatives up to the second order are all continuous in $I = (0, 1)$. For some results, additional smoothness properties may be needed. For example, to obtain optimal error estimates, we assume $\|u\|_{k+2}$ and $\|u_t\|_{k+2}$ are bounded uniformly for any time $t \in [0, T]$.

Furthermore, the nonlinear flux function $f(u)$ is assumed to be smooth enough, also $f(u)$ and its derivatives up to the third order are all bounded on $\mathbb{R}$. For a given initial condition, this assumption is reasonable with the original or a suitably modified flux $f(u)$, see [18] for more details. To emphasize the nonlinearity of $f(u)$, we use $C_\star$ to denote a nonnegative constant depending solely on the maximum of $|f''(u)|$. Remark that $C_\star = 0$ for a linear flux $f(u) = \beta u$.

We would like to present the error estimate result here, and prove it by the energy technique in the next subsections. The main work is to cope with the accumulation of the error from the time discretization. To deal with the nonlinearity of $f(u)$, Taylor expansion and an a priori assumption are used.

Theorem 5.1 (error estimate) *Let u_h be the numerical solution of the fully discrete scheme* (2.6) *with the explicit TVDRK3 time marching, where the finite element space V_h is of piecewise polynomials with arbitrary degree $k \geq 1$, defined on any regular triangulations of $I = (0, 1)$. Let u be the exact solution of problem* (1.1)*, where the flux $f(u)$ is smooth enough. If u is sufficiently smooth with bounded derivatives, then there holds the following error estimate*

$$\max_{n\tau\leq T} \|u(t^n) - u_h^n\| \leq C(h^{k+\sigma} + \tau^3), \tag{5.1}$$

under a CFL condition $\tau \leq \gamma h$ with a fixed CFL constant $\gamma > 0$, where $\sigma = \frac{1}{2}$ for a general monotone numerical flux, and $\sigma = 1$ for a upwind numerical flux. Here C is a positive constant independent of h, τ, and the approximate solution u_h.

5.1 Error representation

Following [18], reference functions are defined in parallel to the Runge-Kutta time discretization stages for the exact solution of the conservation law (1.1). Let $u^{(0)}(x, t) = u(x, t)$, and

$$u^{(1)}(x,t) = u^{(0)}(x,t) - \tau[f(u^{(0)}(x,t))]_x, \tag{5.2a}$$

$$u^{(2)}(x,t) = \frac{3}{4}u^{(0)}(x,t) + \frac{1}{4}u^{(1)}(x,t) - \frac{1}{4}\tau[f(u^{(1)}(x,t))]_x. \tag{5.2b}$$

Denote $u^{n,\sharp} = u^{(\sharp)}(x, t^n)$, for any time level n and $\sharp = 0, 1, 2$.

The error at each stage is denoted by $e^{n,\sharp} = u^{n,\sharp} - u_h^{n,\sharp}$, for any n and inner stage $\sharp = 0, 1, 2$, where $u_h^{n,0} = u_h^n$. As the usual treatment in a finite element analysis, we divide the error into two parts, $e^{n,\sharp} = \xi^{n,\sharp} - \eta^{n,\sharp}$, with

$$\xi^{n,\sharp} = \mathbb{Q}_h u^{n,\sharp} - u_h^{n,\sharp}, \quad \eta^{n,\sharp} = \mathbb{Q}_h u^{n,\sharp} - u^{n,\sharp}, \tag{5.3}$$

where $\mathbb{Q}_h$ is a given projection operator. Here $\xi^{n,\sharp} \in V_h$ would require a careful estimate in the next subsections, while $\eta^{n,\sharp}$ is the approximation error depending on the specific projection. For simplicity we denote $e^n = e^{n,0}$ and $\xi^n = \xi^{n,0}$ below.

The projection operator $\mathbb{Q}_h$ is taken for different purposes. To obtain quasi-optimal error estimates for general monotone fluxes, it is enough to take $\mathbb{Q}_h$ as the standard L^2-projection $\mathbb{P}_h$. However, to obtain an optimal error estimate we would like to follow a standard trick in DG analysis and take $\mathbb{Q}_h$ as the generalized Gauss-Radau projection $\mathbb{R}_h$. For any given function $p(\cdot, t^n)$ at the time level $t = t^n$, the projection operator $\mathbb{R}_h \equiv \mathbb{R}_h^n$ actually depends on the exact solution $u(x, t^n)$, which is defined in each element as

$$\mathbb{R}_h = \begin{cases} \mathbb{R}_h^+, & \text{if } f'(u^n) > 0 \text{ in the element } I_j, \\ \mathbb{R}_h^-, & \text{if } f'(u^n) < 0 \text{ in the element } I_j, \\ \mathbb{P}_h, & \text{if } f'(u^n) \text{ changes its sign on the element } \bar{I}_j. \end{cases} \tag{5.4}$$

For a linear flux $f(u) = \beta u$, this generalized Gauss-Radau projection $\mathbb{R}_h$ solely depends on the sign of β, namely, $\mathbb{R}_h = \mathbb{R}_h^+$ if $\beta > 0$, otherwise $\mathbb{R}_h = \mathbb{R}_h^-$ if $\beta < 0$.

At the end of this subsection we present some estimates for the projection error $\eta^{n,\sharp}$. Since u is assumed to be smooth enough, from (3.2) there holds

$$\|\eta^{n,\sharp}\| + h\|\eta_x^{n,\sharp}\| + h^{1/2}\|\eta^{n,\sharp}\|_{\Gamma_h} \le C_1 h^{k+1}, \quad \forall\, n\colon n\tau \le T; \tag{5.5a}$$

Then it follows from the Sobolev's inequality that

$$\|\eta^{n,\sharp}\|_\infty \le C_2 h^{k+\frac{1}{2}}, \quad \forall\, n\colon n\tau \le T. \tag{5.5b}$$

By $\mathbb{d}^n = d_3\eta^{n+1} + d_2\eta^{n,2} + d_1\eta^{n,1} + d_0\eta^n$, we denote a linear combination of errors in different stages, where $d_i, (i = 0, 1, 2, 3)$, are any four constants restricted by $d_0 + d_1 + d_2 + d_3 = 0$. Note $\mathbb{Q}_h$ is linear in time for both the standard L^2-projection and the generalized Gauss-Radau projection (5.4), since the considered exact solution is smooth and the characteristics will not intersect with each other; see [18] for more details. Thus we also have

$$\|\mathbb{d}^n\| + h^{1/2}\|\mathbb{d}^n\|_{\Gamma_h} \le C_3 h^{k+1}\tau, \quad \forall\, n\colon n\tau \le T; \tag{5.5c}$$

when u_t is smooth enough. Here C_1, C_2 and C_3 are positive constants independent of n, h, and τ.

5.2 The error equation and the energy equation

To obtain the error equation for different stages of the Runge-Kutta time-marching, we first present the local truncation error in time for the reference functions.

Lemma 5.1 *If $u(\cdot, t) \in C^4[0, T]$, then we have*

$$u(x, t+\tau) = \frac{1}{3}u^{(0)}(x, t) + \frac{2}{3}u^{(2)}(x, t) - \frac{2}{3}\tau[f(u^{(2)}(x, t))]_x + \mathcal{E}(x, t), \tag{5.6}$$

where $\mathcal{E}(x, t)$ is the local truncation error in time, and $\|\mathcal{E}(x, t)\| = \mathcal{O}(\tau^4)$ uniformly for any time $t \in [0, T]$.

Proof. Obviously this lemma is a direct application of the explicit TVDRK3 time-marching for the generalized "ordinary differential equation" $u_t = -[f(u)]_x$. □

We multiply the test function $v_h \in V_h$ on both sides of (5.2a), (5.2b) and (5.6), respectively. Let $t = t^n$ and integrate the above equations by parts. Since u is assumed to be smooth enough, the reference functions $u^{n,\sharp}$ are all continuous in $I = (0, 1)$. Thus the exact flux is equal to the numerical flux at each element boundary point. The process above then yields a set of equalities similar to the scheme (2.6). Subtracting these equalities from the scheme (2.6) gives the error equations for $\xi^{n,\sharp}$ as follows: for any test function $v_h \in V_h$ and $1 \le j \le N$, there holds

$$\int_{I_j} \xi^{n,1} v_h \,\mathrm{d}x = \int_{I_j} \xi^n v_h \,\mathrm{d}x + \tau \mathcal{J}_j(v_h), \tag{5.7a}$$

$$\int_{I_j} \xi^{n,2} v_h \,\mathrm{d}x = \frac{3}{4}\int_{I_j} \xi^n v_h \,\mathrm{d}x + \frac{1}{4}\int_{I_j} \xi^{n,1} v_h \,\mathrm{d}x + \frac{\tau}{4}\mathcal{K}_j(v_h), \tag{5.7b}$$

$$\int_{I_j} \xi^{n+1} v_h \,\mathrm{d}x = \frac{1}{3}\int_{I_j} \xi^n v_h \,\mathrm{d}x + \frac{2}{3}\int_{I_j} \xi^{n,2} v_h \,\mathrm{d}x + \frac{2\tau}{3}\mathcal{L}_j(v_h). \tag{5.7c}$$

These equalities have the same form as the scheme (2.6), with the spatial discrete DG operator $\mathcal{H}_j(u_h^{n,\sharp}, v_h), \sharp = 0, 1, 2$, respectively, replaced by three error operators

$$\mathcal{J}_j(v_h) = \int_{I_j} \frac{1}{\tau}(\eta^{n,1} - \eta^n) v_h \,\mathrm{d}x + \mathcal{D}_j(u^n, u_h^n, v_h), \tag{5.8a}$$

$$\mathcal{K}_j(v_h) = \int_{I_j} \frac{1}{\tau}(4\eta^{n,2} - 3\eta^n - \eta^{n,1}) v_h \,\mathrm{d}x + \mathcal{D}_j(u^{n,1}, u_h^{n,1}, v_h), \tag{5.8b}$$

$$\mathcal{L}_j(v_h) = \int_{I_j} \frac{1}{2\tau}\big[3\eta^{n+1} - \eta^n - 2\eta^{n,2} + 3\mathcal{E}(x, t^n)\big] v_h \,\mathrm{d}x + \mathcal{D}_j(u^{n,2}, u_h^{n,2}, v_h). \tag{5.8c}$$

The integrals in (5.8) are denoted by $\mathcal{J}_j^{tm}(v_h), \mathcal{K}_j^{tm}(v_h)$ and $\mathcal{L}_j^{tm}(v_h)$, respectively, and $\mathcal{D}_j(a, b, v_h) = \mathcal{H}_j(a, v_h) - \mathcal{H}_j(b, v_h)$. Furthermore, we would remove the subscript j to denote the sum of the operator over all elements.

By taking the test function $v_h = \xi^n, 4\xi^{n,1}$ and $6\xi^{n,2}$ in the error equations (5.7a), (5.7b) and (5.7c), respectively, we get the energy equation for ξ^n in the form

$$\begin{aligned} 3\|\xi^{n+1}\|^2 - 3\|\xi^n\|^2 = {} & \tau\Big[\mathcal{J}(\xi^n) + \mathcal{K}(\xi^{n,1}) + 4\mathcal{L}(\xi^{n,2})\Big] \\ & + \|2\xi^{n,2} - \xi^{n,1} - \xi^n\|^2 + 3(\xi^{n+1} - \xi^n, \xi^{n+1} - 2\xi^{n,2} + \xi^n), \end{aligned} \tag{5.9}$$

where each line on the right-hand side will be denoted by Π_1' and Π_2', respectively.

In the next subsection we will estimate Π_1' and Π_2' separately. The analysis follows the same line as that in the stability analysis. When the modification from the stability analysis is trivial, we will only present the result without the detailed proof.

5.3 Estimate for the right-hand side of (5.9)

In this subsection we estimate Π_1' and Π_2' for a general monotone numerical flux, to obtain a quasi-optimal error estimate. It is then enough to use the standard L^2-projection, $\mathbb{Q}_h = \mathbb{P}_h$. However, this analysis can be suitably modified to treat a upwind numerical flux for an optimal error estimate; see Section 5.5.

5.3.1 An important quantity of viscosity

For a general monotone numerical flux $\hat{f}(u^-, u^+)$ consistent with $f(u)$, we follow [18] and introduce an important quantity $\alpha(\hat{f}; p)$ to measure the viscosity provided by the numerical flux. For any piecewise smooth function $p \in H^{1,h}(\mathcal{T}_h)$, on each element boundary point we define it as

$$\alpha(\hat{f};p) \equiv \alpha(\hat{f};p^-,p^+) \equiv \begin{cases} [\![p]\!]^{-1}(f(\{\!\{p\}\!\}) - \hat{f}(p^-,p^+)) & \text{if } [\![p]\!] \neq 0, \\ |f'(\{\!\{p\}\!\})| & \text{if } [\![p]\!] = 0. \end{cases} \tag{5.10}$$

For this quantity we have the following lemma.

Lemma 5.2 *$\alpha(\hat{f};p)$ is nonnegative, and is bounded for any $(p^-,p^+) \in \mathbb{R}^2$. Moreover, this quantity is equivalent to $\frac{1}{2}|f'(\{\!\{p\}\!\})|$, in the sense that there exists a constant $C_\star \geq 0$ depending solely on the maximum of $|f''(u)|$, such that*

$$\alpha(\hat{f};p) - C_\star|[\![p]\!]| \leq \frac{1}{2}\big|f'(\{\!\{p\}\!\})\big| \leq \alpha(\hat{f};p) + C_\star|[\![p]\!]|. \tag{5.11}$$

Proof. The conclusions are cited from [18], except the left inequality in (5.11). This inequality comes from the monotone property, with respect to each argument, of the numerical flux $\hat{f}(p^-,p^+)$. For simplicity, we assume $p^+ \geq p^-$ as an example. If $f'(\{\!\{p\}\!\}) \geq 0$, then we use the consistency property of $\hat{f}$ and a Taylor expansion to conclude

$$\alpha(\hat{f};p) \leq [\![p]\!]^{-1}(f(\{\!\{p\}\!\}) - \hat{f}(p^+,p^+)) = \frac{1}{2}f'(\{\!\{p\}\!\}) + \mathcal{O}(|[\![p]\!]|); \tag{5.12}$$

If $f'(\{\!\{p\}\!\}) < 0$, we also have

$$\alpha(\hat{f};p) \leq [\![p]\!]^{-1}(f(\{\!\{p\}\!\}) - \hat{f}(p^-,p^-)) = -\frac{1}{2}f'(\{\!\{p\}\!\}) + \mathcal{O}(|[\![p]\!]|), \tag{5.13}$$

where $-f'(\{\!\{p\}\!\}) = |f'(\{\!\{p\}\!\})|$. Remark that $\mathcal{O}(|[\![p]\!]|) \leq C_\star|[\![p]\!]|$, so we get the left inequality in (5.11). This completes the proof of this lemma. □

Below we will use some compact notations with regard to the quantity $\alpha(\hat{f};p)$, as we have done in [18]. For any functions q_1 and q_2, we denote

$$\alpha(\hat{f};p)[\![q_1]\!][\![q_2]\!] = \sum_{1\leq j\leq N} \alpha(\hat{f};p)_{j+\frac{1}{2}}[\![q_1]\!]_{j+\frac{1}{2}}[\![q_2]\!]_{j+\frac{1}{2}}. \tag{5.14}$$

If $q_1 = q_2$, we use the simplified notation $\alpha(\hat{f};p)[\![q_1]\!]^2$. Similarly we also use $|f'(p)|[\![q]\!]^2$ to denote the sum of itself over all element interface points.

5.3.2 Some basic estimates

The main term for the error operators (5.8) is $\mathcal{D}(u^{n,\sharp}, u_h^{n,\sharp}, v_h)$. Here and below we drop the superscripts and denote this term by $\mathcal{D}(u,u_h,v_h)$ with $u = u^{n,\sharp}$ and $u_h = u_h^{n,\sharp}$, if there is no confusion.

We present some basic estimates to $\mathcal{D}(u,u_h,v_h)$ for any test function v_h, and furthermore for the specific test function $\xi = \xi^{n,\sharp}$. For writing convenience, we denote

$$T_{int}(e) = f(u) - f(u_h) - f'(u)e, \tag{5.15a}$$

$$T_{bry}(e) = f(u) - f(\{\!\{u_h\}\!\}) - f'(u)\{\!\{e\}\!\}, \tag{5.15b}$$

to represent the nonlinear part of $f(u)$ together with the central numerical flux, where $\{\!\{u_h\}\!\}$ is referred as the reference value at each element interface. For a linear flux $f(u) = \beta u$, these terms in (5.15) are zero.

We separate the operator $\mathcal{D}(u, u_h, v_h)$ into three parts

$$\mathcal{D}(u, u_h, v_h) = \mathcal{H}^{lin}(f'(u); e, v_h) + \mathcal{H}^{nls}(e; v_h) + \mathcal{V}(u_h; v_h), \tag{5.16}$$

in which the three parts are termed as the linear part, the nonlinear part and the viscosity part, respectively. They are given for any w and v as follows:

$$\mathcal{H}^{lin}(f'(u); w, v) = \sum_{1\leq j\leq N} \Big[(f'(u)\{\!\{w\}\!\}[\![v]\!])_{j+\frac{1}{2}} + \int_{I_j} f'(u) w v_x \,\mathrm{d}x \Big], \tag{5.17a}$$

$$\mathcal{H}^{nls}(e; v) = \sum_{1\leq j\leq N} \Big[(T_{bry}(e)[\![v]\!])_{j+\frac{1}{2}} + \int_{I_j} T_{int}(e) v_x \,\mathrm{d}x \Big], \tag{5.17b}$$

$$\mathcal{V}(u_h; v) = \sum_{1\leq j\leq N} (f(\{\!\{u_h\}\!\}) - \hat{f}(u_h))_{j+\frac{1}{2}} [\![v]\!]_{j+\frac{1}{2}}. \tag{5.17c}$$

Note that only $\mathcal{H}^{lin}(f'(u); w, v)$ is a bilinear operator with respect to w and v. $\mathcal{H}^{nls}(e; v)$ and $\mathcal{V}(u_h; v)$ are linear operators with respect to v.

The next three lemmas are given for different parts. One can see that numerical stability is not provided by the linear part (with the central reference value), but by the viscosity part. The nonlinear part does not affect the order of the error, and it disappears for the linear flux $f(u) = \beta u$.

Lemma 5.3 *Denote $S_{\max} = \max_{s\in\mathbb{R}} |f'(s)|$. Then we have*

$$|\mathcal{H}^{lin}(f'(u); \xi, v_h)| \leq 2 S_{\max} \mu h^{-1} \|\xi\| \|v_h\|, \quad \forall\, v_h \in V_h; \tag{5.18a}$$

$$|\mathcal{H}^{lin}(f'(u); \xi, \xi)| \leq C_\star \|\xi\|^2, \tag{5.18b}$$

where $C_\star \geq 0$ is a constant independent of n, h, τ and u_h.

Let ε be any given positive constant. Then there exists a positive constant C independent of n, h, τ and u_h (but it may depend on ε), such that

$$|\mathcal{H}^{lin}(f'(u); \eta, v_h)| \leq \varepsilon |f'(u)| [\![v_h]\!]^2 + C h^{2k+1}, \quad \forall\, v_h \in V_h. \tag{5.18c}$$

Proof. For any test function $v_h \in V_h$, we use the inverse properties (i) and (ii) to get the first conclusion (5.18a). More specifically, from (5.17a) we have

$$\begin{aligned} |\mathcal{H}^{lin}(f'(u); \xi, v_h)| &\leq S_{\max} \sum_{1\leq j\leq N} \{\!\{\xi\}\!\}_{j+\frac{1}{2}} [\![v_h]\!]_{j+\frac{1}{2}} + S_{\max} \|\xi\| \|v_{h,x}\| \\ &\leq S_{\max} \|\xi\|_{\Gamma_h} \|v_h\|_{\Gamma_h} + S_{\max} \|\xi\| \|v_{h,x}\| \\ &\leq S_{\max} (\mu_2)^2 h^{-1} \|\xi\| \|v_h\| + S_{\max} \mu_1 h^{-1} \|\xi\| \|v_h\| \\ &\leq 2 S_{\max} \mu h^{-1} \|\xi\| \|v_h\|, \end{aligned} \tag{5.19}$$

where $\mu = \max\{\mu_1, (\mu_2)^2\}$; see subsection 3.3. In the above estimate, we have used the inequality (4.4) and the similar inequality $\sum_{1\leq j\leq N} \{\!\{\xi\}\!\}^2_{j+\frac{1}{2}} \leq \frac{1}{2} \|\xi\|^2_{\Gamma_h}$.

For a specific test function $v_h = \xi$, we are able to work out the integration in (5.17a). A simple manipulation indicates that

$$\begin{aligned}\mathcal{H}^{lin}(f'(u);\xi,\xi) &= \sum_{1\le j\le N}\Big[f'(u)\{\!\{\xi\}\!\}[\![\xi]\!] - \frac{1}{2}f'(u)[\![\xi^2]\!]\Big]_{j+\frac{1}{2}} - \int_I [f'(u)]_x\xi^2\,\mathrm{d}x\\ &= -\int_I [f'(u)]_x\xi^2\,\mathrm{d}x,\end{aligned} \tag{5.20}$$

which implies the second conclusion (5.18b).

By the definition of the standard L^2-projection $\mathbb{P}_h$ (or the generalized Gauss-Radau projection $\mathbb{R}_h$), we have $\int_{I_j}\eta v_{h,x}\,\mathrm{d}x = 0$ for any $v_h \in V_h$ and $j = 1, 2, \ldots, N$. Hence

$$\mathcal{H}^{lin}(f'(u);\eta,v_h) = \sum_{1\le j\le N}\Big[f'(u)\{\!\{\eta\}\!\}[\![v_h]\!] + \int_{I_j}[f'(u)-f'(u_j)]\eta v_{h,x}\,\mathrm{d}x\Big], \tag{5.21}$$

where $u_j = u(x_j)$; consequently $|f'(u)-f'(u_j)| = \mathcal{O}(h)$ in each element I_j. For each term in (5.21), we use the inverse property (i) and Schwartz inequality to estimate the integration term, and use Young's inequality to estimate the jump term. It gives

$$|\mathcal{H}^{lin}(f'(u);\eta,\xi)| \le \varepsilon|f'(u)|[\![\xi]\!]^2 + \frac{1}{4\varepsilon}|f'(u)|\{\!\{\eta\}\!\}^2 + C\|\xi\|^2 + C\|\eta\|^2, \tag{5.22}$$

where ε is an arbitrary positive constant. Finally, we use interpolation property (5.5a) again and get the last conclusion (5.18c) of this lemma. $\square$

Lemma 5.4 *There exists a constant $C_\star \ge 0$ independent of n, h, τ and u_h, such that*

$$|\mathcal{H}^{nls}(e;v_h)| \le C_\star\|v_h\|^2 + C_\star h^{-2}\|e\|_\infty^2\Big[\|\xi\|^2 + h^{2k+2}\Big], \quad \forall\, v_h \in V_h. \tag{5.23}$$

Proof. Using Taylor expansions up to the second order derivative terms, it is easy to see that $T_{int}(e) = -\frac{1}{2}f''e^2$ and $T_{bry}(e) = -\frac{1}{2}\tilde{f}''\{\!\{e\}\!\}^2$, where f'' and $\tilde{f}''$ are the mean values of the second order derivatives of f which are both bounded. Thus we have

$$\begin{aligned}|\mathcal{H}^{nls}(e;v_h)| &= \Big| -\frac{1}{2}\sum_{1\le j\le N}\Big[\tilde{f}''_u\{\!\{e\}\!\}^2_{j+1/2}[\![v_h]\!]_{j+1/2} + \int_{I_j} f''_u e^2 v_{h,x}\,\mathrm{d}x\Big]\Big|\\ &\le C_\star\|e\|_\infty\Big[\|v_h\|_{\Gamma_h}\|\xi-\eta\|_{\Gamma_h} + \|v_{h,x}\|\|\xi-\eta\|\Big]\\ &\le C_\star\|e\|_\infty\Big[\mu_2 h^{-\frac{1}{2}}\|v_h\|(\mu_2 h^{-\frac{1}{2}}\|\xi\| + \|\eta\|_{\Gamma_h}) + \mu_1 h^{-1}\|v_h\|(\|\xi\|+\|\eta\|)\Big]\\ &\le C_\star h^{-1}\|e\|_\infty\|v_h\|\Big[\|\xi\| + h^{k+1}\Big],\end{aligned}$$

where for the second step we have used inequality (4.4), for the third step we have used the inverse properties (i) and (ii), and for the last step we have used the approximation property (5.5a). Finally, we complete the proof of this lemma by a simple application of Young's inequality. $\square$

Lemma 5.5 *Denote* $\mathbb{S}_{\max} = \max_{(s^-,s^+)\in\mathbb{R}^2} \alpha(\hat{f}; s^-, s^+)$*, and let* ε *be any given positive constant. Then we have*

$$|\mathcal{V}(u_h; v_h)| \leq 2\mathbb{S}_{\max}\mu h^{-1}\Big[\|\xi\| + Ch^{k+1}\Big]\|v_h\|, \quad v_h \in V_h; \tag{5.24a}$$

$$\mathcal{V}(u_h; \xi) \leq Ch^{2k+1} - (1-\varepsilon)\alpha(\hat{f}; u_h)[\![\xi]\!]^2, \tag{5.24b}$$

where the positive constant C *is independent of* n, h, τ *and* u_h *but may depends on* ε*.*

Proof. Obviously $\mathbb{S}_{\max}$ is a finite number, by using Lemma 5.2. Since $u = u^{n,\sharp}$ is a continuous function, there holds $[\![u_h]\!] = -[\![u - u_h]\!] = [\![\eta]\!] - [\![\xi]\!]$ at each element interface point. Hence, it follows from the definition (5.10) that

$$\mathcal{V}(u_h; v_h) = \alpha(\hat{f}; u_h)[\![u_h]\!][\![v_h]\!] = \alpha(\hat{f}; u_h)[\![\eta]\!][\![v_h]\!] - \alpha(\hat{f}; u_h)[\![\xi]\!][\![v_h]\!]. \tag{5.25}$$

For any test function $v_h \in V_h$, we can get the first conclusion (5.24a) as follows

$$\begin{aligned}
|\mathcal{V}(u_h; v_h)| &\leq \mathbb{S}_{\max}\sqrt{2}\|v_h\|_{\Gamma_h}\Big[\sqrt{2}\|\xi\|_{\Gamma_h} + \sqrt{2}\|\eta\|_{\Gamma_h}\Big] \\
&\leq 2\mathbb{S}_{\max}\mu_2 h^{-\frac{1}{2}}\|v_h\|\Big[\mu_2 h^{-\frac{1}{2}}\|\xi\| + Ch^{k+\frac{1}{2}}\Big] \\
&\leq 2\mathbb{S}_{\max}\mu h^{-1}\Big[\|\xi\| + Ch^{k+1}\Big]\|v_h\|,
\end{aligned}$$

since $(\mu_2)^2 \leq \mu$. In the first step we have used the inequality (4.4), and in the second step we used the inverse property (ii) and the approximation property (5.5a).

As for the last conclusion (5.24b), it is trivial to get by using Young's inequality for the first term in (5.25). More specifically, from (5.25) we have

$$\mathcal{V}(u_h; \xi) \leq (\varepsilon - 1)\alpha(\hat{f}; u_h)[\![\xi]\!]^2 + \frac{1}{4\varepsilon}\alpha(\hat{f}; u_h)[\![\eta]\!]^2 \leq Ch^{2k+1} - (1-\varepsilon)\alpha(\hat{f}; u_h)[\![\xi]\!]^2,$$

since $\alpha(\hat{f}; u_h)$ is bounded by $\mathbb{S}_{\max}$. Here the approximation property (5.5a) is used again. Now we complete the proof of this lemma. □

At the end of this subsection, we would like to give a crude estimate for the L^2-norm of the error $\xi^{n,\sharp}$ in each stage of the explicit TVDRK3 time-marching. It is used to control the error at intermediate stages by the error at the time level $t = t^n$.

Lemma 5.6 *If the time step satisfies* $\tau = \mathcal{O}(h)$*, then there exists a positive constant* C *independent of* n, h, τ *and* u_h*, such that*

$$\|\xi^{n,1}\|^2 \leq C\|\xi^n\|^2 + Ch^{2k+2}, \tag{5.26a}$$

$$\|\xi^{n,2}\|^2 \leq C\|\xi^n\|^2 + C\|\xi^{n,1}\|^2 + Ch^{2k+2}. \tag{5.26b}$$

Proof. To prove the first conclusion (5.26a), we take the test function $v_h = \xi^{n,1}$ in the error equation (5.7a), and get

$$\|\xi^{n,1}\|^2 \leq \|\xi^n\|\|\xi^{n,1}\| + |\tau\mathcal{D}(u^n, u_h^n, \xi^{n,1})| + |\tau\mathcal{J}^{tm}(\xi^{n,1})|. \tag{5.27}$$

For the linear part and the nonlinear part in $\mathcal{D}(u^n, u_h^n, \xi^{n,1})$, we can use Lemma 5.3 and Lemma 5.4 to directly bound them, by taking $v_h = \xi^{n,1}$ in (5.18a) and (5.23). However, in this proof we consider the sum of the linear part and the nonlinear part

$$\mathcal{W} = \sum_{1\le j\le N} \left[f(u^n) - f(\{\!\{u_h^n\}\!\})\right]_{j+\frac{1}{2}} [\![\xi^{n,1}]\!]_{j+\frac{1}{2}} + \sum_{1\le j\le N} \int_{I_j} \Big[f(u^n) - f(u_h^n)\Big] \xi_x^{n,1}\,\mathrm{d}x.$$

The inverse properties (i) and (ii), together with the approximation property (5.5a), yield

$$|\mathcal{W}| \le C\|e^n\|_{\Gamma_h}\|\xi^{n,1}\|_{\Gamma_h} + C\|e^n\|\|\xi_x^{n,1}\| \le Ch^{-1}[\|\xi^n\| + h^{k+1}]\|\xi^{n,1}\|, \tag{5.28}$$

since $f(u)$ is globally Lipschitz continuous. Furthermore, the viscosity part is estimated by using (5.24a) in Lemma 5.5, and the last term is estimated by the approximation property (5.5c). The estimate reads

$$|\mathcal{V}(u_h^n, \xi^{n,1})| + |\mathcal{J}^{tm}(\xi^{n,1})| \le Ch^{-1}\Big[\|\xi^n\| + h^{k+1}\Big]\|\xi^{n,1}\| + Ch^{k+1}\|\xi^{n,1}\|. \tag{5.29}$$

Finally we collect estimates (5.28) and (5.29) into (5.27), to obtain

$$\|\xi^{n,1}\| \le \|\xi^n\| + C\tau h^{-1}[\|\xi^n\| + h^{k+1}] + Ch^{k+1}\tau \le C\|\xi^n\| + h^{k+1},$$

since $\tau = \mathcal{O}(h) < 1$. The square of this inequality implies the first conclusion (5.26a).

The next conclusion (5.26b) can be obtained similarly, by taking $v_h = \xi^{n,2}$ in (5.7b) and repeating the above process. The detail is therefore omitted. □

5.3.3 The estimate to Π_1'

Below we will use some compact notations with regard to $f(u)$. For any function p, we denote the bounded constant in the error estimates by a short form $C(p) = C + C_\star h^{-2}\|p\|_\infty^2$, where C and $C_\star$ are positive constants independent of h, τ and p.

The estimate to Π_1' is given by the next lemma.

Lemma 5.7 *For any function $v_h \in V_h$, we have that*

$$\tau\mathcal{J}(\xi^n) \le C(e^n)\Big[\|\xi^n\|^2\tau + h^{2k+1}\tau\Big] - \frac{3\tau}{4}\alpha(\hat{f}; u_h^n)[\![\xi^n]\!]^2, \tag{5.30a}$$

$$\tau\mathcal{K}(\xi^{n,1}) \le C(e^{n,1})\Big[\|\xi^{n,1}\|^2\tau + h^{2k+1}\tau\Big] - \frac{3\tau}{4}\alpha(\hat{f}; u_h^{n,1})[\![\xi^{n,1}]\!]^2, \tag{5.30b}$$

$$\tau\mathcal{L}(\xi^{n,2}) \le C(e^{n,2})\Big[\|\xi^{n,2}\|^2\tau + h^{2k+1}\tau\Big] + C\tau^7 - \frac{3\tau}{4}\alpha(\hat{f}; u_h^{n,2})[\![\xi^{n,2}]\!]^2, \tag{5.30c}$$

where the constants C and $C_\star$ are independent of n, h, τ, and u_h.

Proof. The proof is straightforward, since the test function is taken at the same time level as the error operator is defined at. Below we take the first conclusion as an example, since the proofs for all three conclusions are similar.

We can obtain (5.30a) by using (5.18b) and (5.18c) in Lemmas 5.3, (5.23) in Lemma 5.4, and (5.24b) in Lemma 5.5, together with

$$|\mathcal{J}^{tm}(\xi^n)| \le Ch^{2k+2} + C\|\xi^n\|^2$$

implied by the approximation property (5.5c). Here the positive constant ε which emerges in both (5.18c) and (5.24b), respectively, is taken small enough, for example, $\varepsilon = 1/8$. □

5.3.4 The estimate to Π_2'

The term Π_2' results from the time discretization, so the temporal-spatial condition is needed. We would like to obtain an error estimate under $\tau \leq \gamma h$ with a suitable constant $\gamma > 0$, which is referred as the CFL condition. Below we assume there is a fixed constant δ such that the time step τ satisfies

$$\max\{S_{\max}, \mathbb{S}_{\max}\}\mu\tau h^{-1} \leq \delta. \tag{5.31}$$

For example, $\delta = 1/60$ is enough for our error estimate. Here $S_{\max}$ and $\mathbb{S}_{\max}$ have been defined in Lemma 5.3 and Lemma 5.5, respectively. Lemma 5.2 implies that these quantities are almost the same.

Parallel to $\mathbb{D}_i^n$ in the stability analysis, we introduce the following notations

$$\mathbb{G}_1^n = \xi^{n,1} - \xi^n, \quad \mathbb{G}_2^n = 2\xi^{n,2} - \xi^{n,1} - \xi^n, \quad \mathbb{G}_3^n = \xi^{n+1} - 2\xi^{n,2} + \xi^n. \tag{5.32}$$

Similarly to Lemma 4.3, we first build up the relationships among these combinations of errors. The proof follows the same lines as that for Lemma 4.3 and is therefore omitted.

Lemma 5.8 *For the fully discrete RKDG method (2.6) with the explicit TVDRK3 time marching, we have the following identities*

$$(\mathbb{G}_1^n, v_h) = \tau\mathcal{J}(e^n, v_h) \equiv \tau\mathcal{J}_{RK}(v_h), \tag{5.33a}$$

$$(\mathbb{G}_2^n, v_h) = \frac{\tau}{2}\Big[\mathcal{K}(e^{n,1}, v_h) - \mathcal{J}(e^n, v_h)\Big] \equiv \frac{\tau}{2}\mathcal{K}_{RK}(v_h), \tag{5.33b}$$

$$(\mathbb{G}_3^n, v_h) = \frac{\tau}{3}\Big[2\mathcal{L}(e^{n,2}, v_h) - \mathcal{K}(e^{n,1}, v_h) - \mathcal{J}(e^n, v_h)\Big] \equiv \frac{\tau}{3}\mathcal{L}_{RK}(v_h). \tag{5.33c}$$

for any test function $v_h \in V_h$.

The estimate to Π_2' is given in the next lemma, where $-\frac{1}{2}\|\mathbb{G}_2^n\|^2$ plays an important role to obtain error estimates under the standard CFL condition.

Lemma 5.9 *Assume the temporal-spatial condition (5.31) holds. Then we have*

$$\begin{aligned}\Pi_2' \leq & -\frac{1}{2}\|\mathbb{G}_2^n\|^2 + \frac{1}{2}\sum_{\sharp=0,1,2}\tau\alpha(\hat{f}, u^{n,\sharp})[\![\xi^{n,\sharp}]\!]^2 + \sum_{\sharp=0,1,2}C(e^{n,\sharp})\Big[\|\xi^{n,\sharp}\|^2 + h^{2k+1}\Big]\tau \\ & + \sum_{\sharp=0,1}C_\star h^{-1}\|\mathbb{z}^{n,\sharp}\|_\infty^2\Big[\|\xi^{n,\sharp}\|^2 + h^{2k+1}\Big]\tau + C\tau^7,\end{aligned} \tag{5.34}$$

where $C > 0$ and $C_\star \geq 0$ are constants independent of n, h, τ and u_h.

Note that $\xi^{n+1} - \xi^n = \mathbb{G}_1^n + \mathbb{G}_2^n + \mathbb{G}_3^n$, and there holds the equivalent form

$$\Pi_2' = (\mathbb{G}_2^n, \mathbb{G}_2^n) + 3(\mathbb{G}_1^n, \mathbb{G}_3^n) + 3(\mathbb{G}_2^n, \mathbb{G}_3^n) + 3(\mathbb{G}_3^n, \mathbb{G}_3^n) = \Theta_1 + \Theta_2 + \Theta_3 + \Theta_4. \tag{5.35}$$

Below we will prove Lemma 5.9 by estimating each term in (5.35) separately. The bridge is the formulas (5.33b) and (5.33c) in Lemma 5.8, and the analysis depends on the properties of $\mathcal{K}_{RK}$ and $\mathcal{L}_{RK}$. This process includes three steps.

Step 1: To discuss clearly, we divide $\mathcal{K}_{RK}$ and $\mathcal{L}_{RK}$, respectively, into four parts, namely, the linear part, the nonlinear part, the viscosity part and the time-marching part. It reads

$$\mathcal{K}_{RK}(v_h) = \mathcal{K}_{RK}^{lin}(v_h) + \mathcal{K}_{RK}^{nls}(v_h) + \mathcal{K}_{RK}^{vis}(v_h) + \mathcal{K}_{RK}^{tm}(v_h), \tag{5.36a}$$

$$\mathcal{L}_{RK}(v_h) = \mathcal{L}_{RK}^{lin}(v_h) + \mathcal{L}_{RK}^{nls}(v_h) + \mathcal{L}_{RK}^{vis}(v_h) + \mathcal{L}_{RK}^{tm}(v_h). \tag{5.36b}$$

Each part is made up by a combination of same parts in $\mathcal{K}$ and $\mathcal{L}$. For example, the linear part is given as

$$\mathcal{K}_{RK}^{lin}(v_h) = \mathcal{H}^{lin}(f'(u^{n,1}); e^{n,1}, v_h) - \mathcal{H}^{lin}(f'(u^n); e^n, v_h), \tag{5.36c}$$

$$\mathcal{L}_{RK}^{lin}(v_h) = 2\mathcal{H}^{lin}(f'(u^{n,2}); e^{n,2}, v_h) - \mathcal{H}^{lin}(f'(u^{n,1}); e^{n,1}, v_h) - \mathcal{H}^{lin}(f'(u^n); e^n, v_h). \tag{5.36d}$$

The definitions for the other parts are obvious and similar. We omit the details here to save space.

The estimates for the nonlinear part, the viscosity part and time-marching part are easily obtained from Lemma 5.4, Lemma 5.5 and the approximation property (5.5c). However, the estimate for the linear part needs a careful analysis, especially for the following combination.

Lemma 5.10 *If the time step satisfies $\tau = \mathcal{O}(h)$, then we have*

$$|\mathcal{K}_{RK}^{lin}(\mathbb{G}_2^n) + \mathcal{L}_{RK}^{lin}(\mathbb{G}_1^n)| \le C\|\xi^n\|^2 + C\|\xi^{n,1}\|^2 + C\|\xi^{n,2}\|^2 + Ch^{2k+2}, \tag{5.37}$$

where the positive constant C is independent of n, h, τ and u_h.

Proof. The proof is based on the simple separation $\mathcal{K}_{RK}^{lin}(\mathbb{G}_2^n) + \mathcal{L}_{RK}^{lin}(\mathbb{G}_1^n) = \mathcal{R}(\xi) - \mathcal{R}(\eta)$, where $\mathcal{R}(p)$ is defined for any function p, in the form

$$\begin{aligned}\mathcal{R}(p) = &\mathcal{H}^{lin}(f'(u^n); p^{n,1} - p^n, \mathbb{G}_2^n) + \mathcal{H}^{lin}(f'(u^n); 2p^{n,2} - p^{n,1} - p^n, \mathbb{G}_1^n)\\ &+ \mathcal{H}^{lin}(z^{n,1}; p^{n,1}, \mathbb{G}_2^n) + 2\mathcal{H}^{lin}(z^{n,2}; p^{n,2}, \mathbb{G}_1^n) - \mathcal{H}^{lin}(z^{n,1}; p^{n,1}, \mathbb{G}_1^n).\end{aligned}$$

Here $z^{n,\sharp} = f'(u^{n,\sharp}) - f'(u^n), \sharp = 0, 1$, denotes the difference of the flow speed in each stage of the TVDRK3 time-marching. Each term above is denoted by $\mathcal{R}_i(p)$ in the natural order, for $i = 1, 2, 3, 4, 5$.

We first estimate each term in $\mathcal{R}(\xi)$. Along the same line as in Lemma 4.2, we can work out the equivalent expression for $\mathcal{R}_1(\xi) + \mathcal{R}_2(\xi)$. Noticing the periodic boundary condition, a series of manipulations yields

$$\begin{aligned}&\mathcal{R}_1(\xi) + \mathcal{R}_2(\xi) + \int_I [f'(u^n)]_x \mathbb{G}_1^n \mathbb{G}_2^n \,\mathrm{d}x\\ =&\sum_{1\le j\le N} \Big[\int_{I_j} [f'(u^n)\mathbb{G}_1^n\mathbb{G}_2^n]_x \,\mathrm{d}x + f'(u^n)\{\!\{\mathbb{G}_1^n\}\!\}[\![\mathbb{G}_2^n]\!] + f'(u^n)\{\!\{\mathbb{G}_2^n\}\!\}[\![\mathbb{G}_1^n]\!]\Big]_{j+\frac{1}{2}}\\ =&\sum_{1\le j\le N} \Big[- f'(u^n)[\![\mathbb{G}_1^n\mathbb{G}_2^n]\!] + f'(u^n)\{\!\{\mathbb{G}_1^n\}\!\}[\![\mathbb{G}_2^n]\!] + f'(u^n)\{\!\{\mathbb{G}_2^n\}\!\}[\![\mathbb{G}_1^n]\!]\Big]_{j+\frac{1}{2}}\\ =&\,0,\end{aligned}$$

where we have used the basic fact, for any a and b, that

$$-(a^+b^+ - a^-b^-) + \frac{a^+ + a^-}{2}(b^+ - b^-) + \frac{b^+ + b^-}{2}(a^+ - a^-) = 0.$$

Consequently, $|\mathcal{R}_1(\xi) + \mathcal{R}_2(\xi)| \le C_\star \|\mathbb{G}_1^n\| \|\mathbb{G}_2^n\|$. Since $|z^{n,1}| = \mathcal{O}(\tau) = \mathcal{O}(h)$, it follows from the inverse properties (i) and (ii) that $|\mathcal{R}_3(\xi)| \le C_\star \|\xi^{n,1}\| \|\mathbb{G}_2^n\|$. Similarly, we also have $|\mathcal{R}_4(\xi)| + |\mathcal{R}_5(\xi)| \le C_\star[\|\xi^n\| + \|\xi^{n,1}\|] \|\mathbb{G}_1^n\|$. Thus, summing up the above inequalities yields

$$|\mathcal{R}(\xi)| \le C_\star \Big[\|\mathbb{G}_1^n\|^2 + \|\mathbb{G}_2^n\|^2 + \|\xi^n\|^2 + \|\xi^{n,1}\|^2 \Big]. \tag{5.38}$$

It is easy to estimate each term in $\mathcal{R}(\eta)$ by using the inverse properties (i) and (ii), together with the approximation properties (5.5a) and (5.5c). For example, the first term $\mathcal{R}_1(\eta) = \mathcal{H}^{lin}(f'(u^n); \eta^{n,1} - \eta^n, \mathbb{G}_2^n)$ is almost the same as (5.21), except that η is changed to $\eta^{n,1} - \eta^n$. Hence we can use the inverse inequalities (i) and (ii), to get

$$|\mathcal{R}_1(\eta)| = \Big[S_{\max}\mu_2 h^{-\frac{1}{2}} \|\eta^{n,1} - \eta^n\|_{\Gamma_h} + C_\star \|\eta^{n,1} - \eta^n\| \Big] \|\mathbb{G}_2^n\| \le C\|\mathbb{G}_2^n\|^2 + Ch^{2k}\tau^2;$$

Similarly we can estimate the second term as $|\mathcal{R}_2(\eta)| \le C\|\mathbb{G}_1^n\|^2 + Ch^{2k}\tau^2$. By noticing again $z^{n,\sharp} = \mathcal{O}(\tau) = \mathcal{O}(h)$, it is easy to estimate the remaining three terms in form $|\mathcal{R}_3(\eta)| + |\mathcal{R}_4(\eta)| + |\mathcal{R}_5(\eta)| \le C\|\mathbb{G}_1^n\|^2 + C\|\mathbb{G}_2^n\|^2 + Ch^{2k+2}$. Since $\tau = \mathcal{O}(h)$, finally we get

$$\mathcal{R}(\eta) \le C\|\mathbb{G}_1^n\|^2 + C\|\mathbb{G}_2^n\|^2 + Ch^{2k+2}. \tag{5.39}$$

Now we collect estimates (5.38) and (5.39) and complete the proof of this lemma, since $\|\mathbb{G}_1^n\|^2 + \|\mathbb{G}_2^n\|^2 \le C(\|\xi^n\|^2 + \|\xi^{n,1}\|^2 + \|\xi^{n,2}\|^2)$. □

Remark 5.1. The conclusion and proof of this step demonstrate the well-known fact that a Runge-Kutta algorithm obtains high order accuracy through a combination of Euler forward time-marchings. If we estimate each of the stages separately, the integration in each cell will prevent us from obtaining quasi-optimal and/or optimal error estimates.

Step 2: Next we turn to estimate the sum of the first two terms Θ_1 and Θ_2. By taking different test functions in the identities (5.33b) and (5.33c) of Lemma 5.8, we have that

$$\begin{aligned}
\Theta_1 + \Theta_2 &= -(\mathbb{G}_2^n, \mathbb{G}_2^n) + 2(\mathbb{G}_2^n, \mathbb{G}_2^n) + 3(\mathbb{G}_3^n, \mathbb{G}_1^n) \\
&= -\|\mathbb{G}_2^n\|^2 + \tau\mathcal{K}_{RK}(\mathbb{G}_2^n) + \tau\mathcal{L}_{RK}(\mathbb{G}_1^n) \\
&= -\|\mathbb{G}_2^n\|^2 + \tau\mathcal{K}_{RK}^{lin}(\mathbb{G}_2^n) + \tau\mathcal{L}_{RK}^{lin}(\mathbb{G}_1^n) + \tau\mathcal{K}_{RK}^{nls}(\mathbb{G}_2^n) + \tau\mathcal{L}_{RK}^{nls}(\mathbb{G}_1^n) \\
&\qquad + \tau\mathcal{K}_{RK}^{vis}(\mathbb{G}_2^n) + \tau\mathcal{L}_{RK}^{vis}(\mathbb{G}_1^n) + \tau\mathcal{K}_{RK}^{tm}(\mathbb{G}_2^n) + \tau\mathcal{L}_{RK}^{tm}(\mathbb{G}_1^n) \\
&= -\|\mathbb{G}_2^n\|^2 + \mathcal{Q}_1 + \mathcal{Q}_2 + \mathcal{Q}_3 + \mathcal{Q}_4,
\end{aligned} \tag{5.40}$$

where $\mathcal{Q}_i, i = 1, 2, 3, 4$, consecutively represent the sums of two adjacent terms. We will estimate each of them separately below.

The term $\mathcal{Q}_1$ is controlled by Lemma 5.10. By Lemma 5.4, the estimate to $\mathcal{Q}_2$ is obtained as

$$|\mathcal{Q}_2| \le C\|\mathbb{G}_1^n\|^2\tau + C\|\mathbb{G}_2^n\|^2\tau + \sum_{\sharp=0,1,2} C_\star h^{-2} \|e^{n,\sharp}\|_\infty^2 \Big[\|\xi^{n,\sharp}\|^2 + h^{2k+2} \Big]\tau. \tag{5.41}$$

By the approximation property (5.5c), it is easy to estimate $\mathcal{Q}_4$ in form

$$|\mathcal{Q}_4| \le C(\|\mathbb{G}_1^n\|^2\tau + \|\mathbb{G}_2^n\|^2\tau + h^{2k+2}\tau + \tau^7). \tag{5.42}$$

We now come to the estimate of $\mathcal{Q}_3$. This estimate would be easy if the quantities $\alpha(\hat{f}, u^{n,\sharp})$ in every terms which form the sum $\mathcal{Q}_3$ are the same. If they are not, we denote the gap of the viscosity in different stages by

$$\mathbb{z}^{n,0} = \alpha(\hat{f}; u_h^{n,2}) - \alpha(\hat{f}; u_h^n), \quad \mathbb{z}^{n,1} = \alpha(\hat{f}; u_h^{n,2}) - \alpha(\hat{f}; u_h^{n,1}), \tag{5.43}$$

which are parallel to $z^{n,\sharp}$ defined in the proof of Lemma 5.10. The bound for these quantities will be given in Section 5.5.

Then we split $\mathcal{Q}_3$ into three terms $\mathcal{Q}_3 = \mathcal{Q}_3^{(1)} + \mathcal{Q}_3^{(2)} + \mathcal{Q}_3^{(3)}$, where

$$\begin{aligned}
\mathcal{Q}_3^{(1)} &= \tau\alpha(\hat{f}, u_h^n)[\![\xi^n]\!][\![\mathbb{G}_2^n]\!] - \tau\alpha(\hat{f}, u_h^{n,1})[\![\xi^{n,1}]\!][\![\mathbb{G}_2^n]\!] - \tau\alpha(\hat{f}; u_h^{n,2})[\![\mathbb{G}_2^n]\!][\![\mathbb{G}_1^n]\!],\\
\mathcal{Q}_3^{(2)} &= \tau\alpha(\hat{f}; u_h^{n,1})[\![\eta^{n,1}]\!][\![\mathbb{G}_2^n]\!] - \tau\alpha(\hat{f}; u_h^n)[\![\eta^n]\!][\![\mathbb{G}_2^n]\!] + \tau\alpha(\hat{f}; u_h^{n,2})[\![\tilde{\mathbb{G}}_2^n]\!][\![\mathbb{G}_1^n]\!],\\
\mathcal{Q}_3^{(3)} &= -\,\mathbb{z}^{n,0}[\![\xi^n - \eta^n]\!][\![\mathbb{G}_1^n]\!]\tau - \mathbb{z}^{n,1}[\![\xi^{n,1} - \eta^{n,1}]\!][\![\mathbb{G}_1^n]\!]\tau,
\end{aligned}$$

here $\tilde{\mathbb{G}}_2^n = 2\eta^{n,2} - \eta^{n,1} - \eta^n$. Notice that all these terms should be understood as the sum over all element boundary points.

Keeping in mind that $\alpha(\hat{f}; u_h)$ is bounded by $\mathbb{S}_{\max}$, we estimate the above terms one by one. The inverse property (ii) and Young's inequality yield

$$\begin{aligned}
|\mathcal{Q}_3^{(1)}| \le &\varepsilon_1\tau\alpha(\hat{f}; u_h^n)[\![\xi^n]\!]^2 + \frac{1}{4\varepsilon_1}\tau\alpha(\hat{f}; u_h^n)[\![\mathbb{G}_2^n]\!]^2 + \varepsilon_1\tau\alpha(\hat{f}; u_h^{n,1})[\![\xi^{n,1}]\!]^2\\
&+ \frac{1}{4\varepsilon_1}\tau\alpha(\hat{f}; u_h^{n,1})[\![\mathbb{G}_2^n]\!]^2 + \varepsilon_1\tau\alpha(\hat{f}; u_h^{n,1})[\![\mathbb{G}_1^n]\!]^2 + \frac{1}{4\varepsilon_1}\tau\alpha(\hat{f}; u_h^{n,1})[\![\mathbb{G}_2^n]\!]^2\\
\le &\varepsilon_1\sum_{\sharp=0,1}\tau\alpha(\hat{f}; u_h^{n,\sharp})[\![\xi^{n,\sharp}]\!]^2 + \varepsilon_1\tau\alpha(\hat{f}; u_h^{n,2})[\![\mathbb{G}_1^n]\!]^2 + \frac{3}{2\varepsilon_1}\mathbb{S}_{\max}\mu\tau h^{-1}\|\mathbb{G}_2^n\|,
\end{aligned} \tag{5.44a}$$

where we have used the inequality (4.4). Similarly, by the approximation property (5.5a) and Young's inequality, we have that

$$\begin{aligned}
|\mathcal{Q}_3^{(2)}| \le &\varepsilon_1\tau\alpha(\hat{f}; u_h^{n,2})[\![\mathbb{G}_1^n]\!]^2 + \frac{1}{4\varepsilon_1}\tau\alpha(\hat{f}; u_h^{n,2})[\![\tilde{\mathbb{G}}_2^n]\!]^2 + \frac{1}{4}\tau\alpha(\hat{f}; u_h^n)[\![\mathbb{G}_2^n]\!]^2\\
&+ \tau\alpha(\hat{f}; u_h^n)[\![\eta^n]\!]^2 + \frac{1}{4}\tau\alpha(\hat{f}; u_h^{n,1})[\![\mathbb{G}_2^n]\!]^2 + \tau\alpha(\hat{f}; u_h^{n,1})[\![\eta^{n,1}]\!]^2\\
\le &\varepsilon_1\tau\alpha(\hat{f}; u_h^{n,2})[\![\mathbb{G}_1^n]\!]^2 + \mathbb{S}_{\max}\mu\tau h^{-1}\|\mathbb{G}_2^n\|^2 + Ch^{2k+1}\tau,
\end{aligned} \tag{5.44b}$$

Here ε_1 in (5.44a) and (5.44b) is a suitably small positive constant, which now is taken as $\varepsilon_1 = 1/16$. As for the last term, it is easy to get that

$$|\mathcal{Q}_3^{(3)}| \le C\|\mathbb{G}_1^n\|^2\tau + \sum_{\sharp=0,1} Ch^{-2}\|\mathbb{z}^{n,\sharp}\|_\infty^2\Big[\|\xi^{n,\sharp}\|^2 + h^{2k+1}\Big]\tau. \tag{5.44c}$$

by the approximation property (5.5a) and Young's inequality again.

By the definition of $\mathbb{z}^{n,\sharp}$ and the inequality $2ab \le a^2+b^2$, we have

$$\alpha(\hat{f}, u_h^{n,2})[\![\mathbb{G}_1^n]\!]^2 \le 2\sum_{\sharp=0,1} \alpha(\hat{f}, u_h^{n,\sharp})[\![\xi^{n,\sharp}]\!]^2 + 2\sum_{\sharp=0,1} \|\mathbb{z}^{n,\sharp}\|_\infty [\![\xi^{n,\sharp}]\!]^2.$$

Noticing the CFL condition (5.31), we collect the estimates above about $\mathcal{Q}_i, i=1,2,3,4$, to obtain the following estimate

$$\begin{aligned}|\Theta_1+\Theta_2| \le &-\frac{7}{12}\|\mathbb{G}_2^n\|^2 + \frac{5}{16}\sum_{\sharp=0,1}\tau\alpha(\hat{f}; u_h^{n,\sharp})[\![\xi^{n,\sharp}]\!]^2 + \sum_{\sharp=0,1,2} C(e^{n,\sharp})\Big[\|\xi^{n,\sharp}\|^2 + h^{2k+1}\Big]\tau \\ &+ \sum_{\sharp=0,1} C_\star h^{-2}\|\mathbb{z}^{n,\sharp}\|_\infty^2\Big[\|\xi^n\|^2 + h^{2k+1}\Big]\tau + C\tau^7. \end{aligned} \tag{5.45}$$

Step 3: Next we turn to estimate the last two terms Θ_3 and Θ_4, separately. It follows from the identity (5.33c) in Lemma 5.8, that

$$\Theta_3 = 3(\mathbb{G}_3^n, \mathbb{G}_2^n) = \tau\mathcal{L}_{RK}(\mathbb{G}_2^n), \quad \Theta_4 = 3\|\mathbb{G}_3^n\|^2 = \tau\mathcal{L}_{RK}(\mathbb{G}_3^n). \tag{5.46}$$

Thus we have to estimate $\mathcal{L}_{RK}(v_h)$ for different test functions. From (5.36b), each part can be estimated easily, except the linear part which needs a careful treatment to dig out $\mathbb{G}_2^n$.

Along the same line as in Lemma 5.10, we take the flow speed to be the same and rewrite the linear part $\mathcal{L}_{RK}^{lin}(v_h)$ in the equivalent form

$$\mathcal{H}^{lin}(f'(u^n); \mathbb{G}_2^n, v_h) - \mathcal{H}^{lin}(f'(u^n); \tilde{\mathbb{G}}_2^n, v_h) + 2\mathcal{H}^{lin}(z^{n,2}; e^{n,2}, v_h) - \mathcal{H}^{lin}(z^{n,1}; e^{n,1}, v_h),$$

where $\tilde{\mathbb{G}}_2^n = 2\eta^{n,2} - \eta^{n,1} - \eta^n$. Here, the first term is estimated by (5.18a) and the inequality $2ab \le a^2+b^2$, and the second term is estimated by (5.18c) with $\varepsilon = 1/2$, together with the inverse property (ii) and the inequality (4.4); see Lemma 5.3. More specifically, the sum of first two terms, denoted by $\mathcal{Z}$, is bounded by

$$\begin{aligned}|\mathcal{Z}| &\le 2S_{\max}\mu h^{-1}\|\mathbb{G}_2^n\|\|v_h\| + \frac{1}{2}|f'(u^n)|[\![v_h]\!]^2 + Ch^{2k+1} \\ &\le S_{\max}\mu h^{-1}\|\mathbb{G}_2^n\|^2 + 2S_{\max}\mu h^{-1}\|v_h\|^2 + Ch^{2k+1}.\end{aligned}$$

The other terms are easily bounded by the fact $z^{n,\sharp} = \mathcal{O}(\tau) = \mathcal{O}(h)$, and the inverse properties (i) and (ii), together with approximation property (5.5a). Finally, the result reads

$$\begin{aligned}\tau\mathcal{L}_{RK}^{lin}(v_h) \le\, &S_{\max}\mu\tau h^{-1}\|\mathbb{G}_2^n\|^2 + 2S_{\max}\mu\tau h^{-1}\|v_h\|^2 \\ &+ C_\star\|v_h\|^2\tau + C_\star\|\xi^{n,1}\|^2\tau + C_\star\|\xi^{n,2}\|^2\tau + Ch^{2k+1}\tau.\end{aligned} \tag{5.47a}$$

We use Lemma 5.4 for the nonlinear part, and the approximation property (5.5c) for the time-marching part. The estimates read

$$\tau\mathcal{L}_{RK}^{nls}(v_h) \le C_\star\|v_h\|^2\tau + \sum_{\sharp=0,1,2} C_\star h^{-2}\|e^{n,\sharp}\|_\infty^2\Big[\|\xi^{n,\sharp}\|^2 + h^{2k+1}\Big]\tau, \tag{5.47b}$$

$$\tau\mathcal{L}_{RK}^{tm}(v_h) \le C\|v_h\|^2\tau + C\big[h^{2k+2} + \tau^6\big]\tau. \tag{5.47c}$$

Furthermore, we use the inverse property (ii), the approximation property (5.5c), and Young's inequality to estimate the viscosity part $\mathcal{L}_{RK}^{vis}(v_h) = 2\mathcal{V}(u_h^{n,2}; v_h) - \mathcal{V}(u_h^{n,1}; v_h) - \mathcal{V}(u_h^n; v_h)$. ¿From (5.25) and (4.4), it is easy to get

$$\begin{aligned}|\mathcal{V}(u_h^{n,\sharp}; v_h)| \le& \varepsilon_2^{n,\sharp}\alpha(\hat{f}; u_h^{n,\sharp})[\![\xi^{n,\sharp}]\!]^2 + \Big[\frac{1}{4\varepsilon_2^{n,\sharp}} + \varepsilon_3^{n,\sharp}\Big]\alpha(\hat{f}; u_h^{n,\sharp})[\![v_h]\!]^2 + \frac{1}{4\varepsilon_3^{n,\sharp}}\alpha(\hat{f}; u_h^{n,\sharp})[\![\eta^{n,\sharp}]\!]^2\\ \le& \varepsilon_2^{n,\sharp}\alpha(\hat{f}; u_h^{n,\sharp})[\![\xi^{n,\sharp}]\!]^2 + 2\Big[\frac{1}{4\varepsilon_2^{n,\sharp}} + \varepsilon_3^{n,\sharp}\Big]\mathbb{S}_{\max}\mu h^{-1}\|v_h\|^2 + Ch^{2k+1},\end{aligned}$$

where $\varepsilon_2^{n,\sharp}$ and $\varepsilon_3^{n,\sharp}$ are suitable small positive constants. Taking $\varepsilon_2^{n,0} = \varepsilon_2^{n,1} = 2\varepsilon_2^{n,2} = 1/8$ and $\varepsilon_3^{n,\sharp} = 1/6$ for $\sharp = 0, 1, 2$, then we have

$$\tau\mathcal{L}_{RK}^{vis}(v_h) \le 25\mathbb{S}_{\max}\mu\tau h^{-1}\|v_h\|^2 + \frac{1}{8}\sum_{\sharp=0,1,2}\tau\alpha(\hat{f}, u_h^{n,\sharp})[\![\xi^{n,\sharp}]\!]^2 + Ch^{2k+1}\tau. \tag{5.47d}$$

The above estimates (5.47), together with the identity (5.46), give a bound for Θ_3 and Θ_4, respectively, under the CFL condition (5.31). Finally, the result reads

$$\begin{aligned}|\Theta_3| + |\Theta_4| \le& \Big[28\delta + \frac{3\delta}{3-27\delta}\Big]\|\mathbb{G}_2^n\|^2 + \frac{1}{8}\Big[1 + \frac{3}{3-27\delta}\Big]\sum_{\sharp=0,1,2}\tau\alpha(\hat{f}, u^{n,\sharp})[\![\xi^{n,\sharp}]\!]^2\\ &+ \sum_{\sharp=0,1,2} C(e^{n,\sharp})\Big[\|\xi^{n,\sharp}\|^2 + h^{2k+1}\Big]\tau + C\tau^7.\end{aligned} \tag{5.48}$$

Since we have taken $\delta = 1/60$ in (5.31), a simple manipulation shows the factors, contained in the square brackets of the first two terms in (5.48), are not greater than 1/2, and 3/8, respectively, since $3 - 27\delta > 3/2$.

Finally we collect the estimates (5.45) and (5.48), and complete the proof for Lemma 5.9.

5.4 Proof of Theorem 5.1

In this subsection we will present the proof of Theorem 5.1 for a general monotone flux. To deal with the nonlinearity of the flux $f(u)$, we would like to adopt the following *a priori* assumption for m, if $m\tau \le T$, that

$$\|e^{n,\sharp}\|_\infty \le \mathbb{C}h, \quad \text{for } n \le m, \sharp = 0, 1, 2, \tag{5.49}$$

which holds for h small enough, where $\mathbb{C}$ is a fixed positive constant independent of m, h, τ and u_h. For a linear flux $f(u) = \beta u$, this assumption is not necessary. Later we will verify the reasonableness of (5.49) for high order piecewise polynomials of degree $k \ge 2$. For the piecewise linear polynomials, see Remark 5.1.

This assumption (5.49) implies that $C(e^{n,\sharp}) \le C$ for any $n \le m$ and $\sharp = 0, 1, 2$, where C is a positive constant independent of m, n, h and τ.

By (5.11), the property of $\alpha(\hat{f}, \cdot)$ yields the estimate on each element boundary point

$$\begin{aligned}|\mathbb{z}^{n,0}| \le& |f'(\{\!\{u_h^{n,2}\}\!\}) - f'(\{\!\{u_h^n\}\!\})| + C_\star\Big[|[\![u_h^{n,2}]\!]| + |[\![u_h^{n,0}]\!]|\Big]\\ \le& C_\star\Big[|\{\!\{u^{n,2} - u^{n,0}\}\!\}| + |\{\!\{e^{n,2}\}\!\}| + |\{\!\{e^n\}\!\}| + |[\![e^{n,2}]\!]| + |[\![e^{n,0}]\!]|\Big]\\ \le& C_\star\Big[C\tau + \|e^n\|_\infty + \|e^{n,2}\|_\infty\Big].\end{aligned}$$

Since $\tau = \mathcal{O}(h)$, the a priori assumption (5.49) also implies $\|\mathbb{z}^{n,0}\|_\infty \le Ch$ for any $n \le m$ and $\sharp = 0, 1, 2$, where C is a positive constant independent of m, n, h and τ. The same is true for $\|\mathbb{z}^{n,1}\|_\infty$.

We now collect the estimates in Lemma 5.7 and Lemma 5.9, into the energy identity (5.9). Noticing Lemma 5.6, under the a priori assumption (5.49) we finally obtain, for any $n \le m$, the following inequality

$$
\begin{aligned}
&3\|\xi^{n+1}\|^2 - 3\|\xi^n\|^2 + \frac{1}{12}\|\mathbb{G}_2^n\|^2 + \frac{1}{16}\tau \sum_{\sharp=0,1,2} \alpha(\hat{f}; u_h^{n,\sharp})[\![\xi^{n,\sharp}]\!]^2 \\
\le& C\|\xi^n\|^2\tau + C\|\xi^{n,1}\|^2\tau + C\|\xi^{n,2}\|^2\tau + Ch^{2k+1}\tau + C\tau^7 \\
\le& C\|\xi^n\|^2\tau + Ch^{2k+1}\tau + C\tau^7,
\end{aligned}
$$

where C is a positive constant independent of m, n, h and τ. Then an application of the discrete Gronwall lemma yields the error estimate for the fully-discrete DG scheme with TVDRK3 time-marching, in the form

$$
\|\xi^{n+1}\|^2 \le Ch^{2k+1} + C\tau^6, \quad n \le m, \tag{5.50}
$$

where C is also a positive constant independent of m, n, h and τ. It together with the approximation property (5.5a) yields that $\|e^{n+1}\|^2 \le Ch^{2k+1} + C\tau^6$.

Before we complete the proof of Theorem 5.1, we need to verify the reasonableness of the a priori assumption (5.49). Since $\xi^0 = 0$, the approximation property (5.5b) implies $\|e^0\|_\infty \le \mathbb{C}h$, and the positive constant $\mathbb{C}$ is solely determined by the initial solution $u_0(x)$. Furthermore, Lemma 5.6 shows $\|\xi^{0,\sharp}\| \le Ch^{k+1}$ for any $\sharp = 1, 2$, and consequently

$$
\|e^{0,\sharp}\|_\infty \le \mu_3 h^{-\frac{1}{2}}\|\xi^{0,\sharp}\| + \|\eta^{0,\sharp}\|_\infty \le \mu_3 Ch^{k+\frac{1}{2}} + Ch^{k+\frac{1}{2}} \le \mathbb{C}h, \tag{5.51}
$$

for small enough h, since $k \ge 2$. Here the inverse property (iii) and the approximation property (5.5b) have been used. Suppose (5.49) holds for m, we can show this assumption is also true for $m+1$. Inequality (5.50) as well as Lemma 5.6 imply that $\|\xi^{m+1,\sharp}\| \le C(h^{k+1/2} + \tau^3)$, for any $\sharp = 0, 1, 2$. Since $k \ge 2$, we can get

$$
\|e^{m+1,\sharp}\|_\infty \le \mu h^{-\frac{1}{2}}[Ch^{k+1/2} + \tau^3] + Ch^{k+1/2} \le C_0 h^2, \tag{5.52}
$$

where the positive constant C_0 depends on T in general, which is independent of m, n, h and τ. Certainly there exists a constant $h_0 > 0$ such that $C_0 h \le \mathbb{C}$, and consequently $\|e^{m+1,\sharp}\|_\infty \le \mathbb{C}h$ if $h \le h_0$.

Thus the assumption (5.49) is reasonable and hence the above error estimate holds for any $m\colon m\tau \le T$. Now we complete the proof of the theorem for a general numerical flux, when piecewise polynomials of degree $k > 1$ is used.

Remark 5.1. The proof above does not work for piecewise linear polynomials, which is usually not used in the RKDG method with TVDRK3 time-marching anyway. The reason is a loss of the a priori assumption (5.49) for piecewise linear polynomials. A more involved analysis shows that the quasi-optimal error estimate also holds for piecewise linear polynomials. Since the proof is much more technical and the RKDG method with TVDRK3 time-marching is rarely used, we will not present the proof here and refer to a similar proof in [18], where the TVDRK2 time-marching is discussed.

Remark 5.2. It is easy to generalize the above error estimates from one-dimension to multi-dimensions. The quasi-optimal error estimates hold for any triangulation. For linear conservation laws, the error estimates hold for any piecewise polynomials, if the exact solution is smooth enough. However, for nonlinear conservation laws, the error estimates only hold for piecewise polynomials with degree $k > d/2+1$, where d is the spatial dimension. This restriction is solely used to ensure the a priori assumption (5.49).

5.5 Extension to upwind numerical fluxes

In this subsection we consider the optimal error estimate when a upwind numerical flux is used. In this case, the a priori assumption (5.49) is also enough for piecewise polynomials with any degree $k \geq 1$. Error estimates start from the energy equation (5.9), and the analysis follows the same line as before.

To obtain an optimal error estimate, the main change is twofold. One is the generalized Gauss-Radau projection (5.4), which is defined in Section 5.1. The other is that, on each element boundary point, the reference value $\{\!\{u_h\}\!\}$ in (5.17) is replaced with the limit in the generalized upwind direction, denoted by u_h^{ex}.

The reference value $u_h^{ex} = u_h^{n,\sharp,ex}$ for any $\sharp = 0,1,2$, is determined by the flow direction of the exact solution u^n, in the adjacent elements. More specifically, it is defined as

$$u_{h,j+\frac{1}{2}}^{n,\sharp,ex} = \begin{cases} (u_h^{n,\sharp})_{j+\frac{1}{2}}^-, & \text{if } f'(u^n) > 0 \text{ on } I_j \cup x_{j+\frac{1}{2}} \cup I_{j+1}, \\ (u_h^{n,\sharp})_{j+\frac{1}{2}}^+, & \text{if } f'(u^n) < 0 \text{ on } I_j \cup x_{j+\frac{1}{2}} \cup I_{j+1}, \\ \{\!\{u_h^{n,\sharp}\}\!\}_{j+\frac{1}{2}}, & \text{otherwise.} \end{cases} \tag{5.53}$$

This leads into some modifications to the analysis for the linear part and the viscosity part. The estimate to the nonlinear part is still the same, see Lemma 5.4.

The modification for the linear part consists of the next lemma.

Lemma 5.11 *If the time step satisfies $\tau = \mathcal{O}(h)$, then for any n and $\sharp = 0,1,2$, there exist two constants $C > 0$ and $C_\star \geq 0$ independent of n, h, τ and u_h, such that*

$$|\mathcal{H}^{lin}(f'(u^{n,\sharp}); e^{n,\sharp}, \xi^{n,\sharp})| \leq -\frac{1}{2}|f'(u^n)|[\![\xi^{n,\sharp}]\!]^2 + C_\star\|\xi^{n,\sharp}\|^2 + Ch^{2k+2}. \tag{5.54}$$

Proof. We first estimate $\mathcal{H}^{lin}(f'(u^n); e^{n,\sharp}, \xi^{n,\sharp})$ along the same line as in Lemma 5.3, corresponding to difference cases in (5.53).

For the first two cases, on each element interface point the upwind reference value provides the viscosity $\frac{1}{2}|f'(u^n)|[\![\xi^{n,\sharp}]\!]^2$ (ref. (4.5b) in Lemma 4.2), and only the approximation error $\eta^{n,\sharp}$ appears in the upwind direction, which is zero since the Gauss-Radau projection is used. For the last case, there is no longer viscosity; however, owing to $f'(u_{j+\frac{1}{2}}) = \mathcal{O}(h)$, the inverse property (ii) yields $\frac{1}{2}|f'(u^n)|[\![\xi^{n,\sharp}]\!]^2 \leq C_\star\|\xi^{n,\sharp}\|^2$, and the lost half order on each element boundary point is therefore recovered. Thus $\mathcal{H}^{lin}(f'(u^n); e^{n,\sharp}, \xi^{n,\sharp})$ is bounded by the right-hand side of (5.54).

Finally, we can complete the proof of this lemma by noticing, for any $\sharp = 0, 1, 2$, that $|\mathcal{H}^{lin}(f'(u^{n,\sharp}) - f'(u^n); e^{n,\sharp}, \xi^{n,\sharp})| \le C_\star(\|\xi^{n,\sharp}\|^2 + Ch^{2k+2})$, since $f'(u^{n,\sharp}) - f'(u^n) = \mathcal{O}(\tau) = \mathcal{O}(h)$. $\square$

The modification for the viscosity part consists of the redefinition of the quantity $\alpha(\hat{f}, p)$ for the upwind numerical flux, which will be denoted by the same notation. It is defined on each element boundary point, by replacing the reference value $\{\!\{p\}\!\}$ in (5.10) by p^{num}, the limit in the generalized upwind direction. More specifically, it reads

$$p^{num}_{j+\frac{1}{2}} = \begin{cases} p^-_{j+\frac{1}{2}}, & \text{if } f'(q) > 0 \text{ for any } q \text{ between } p^-_{j+\frac{1}{2}} \text{ and } p^+_{j+\frac{1}{2}}, \\ p^+_{j+\frac{1}{2}}, & \text{if } f'(q) < 0 \text{ for any } q \text{ between } p^-_{j+\frac{1}{2}} \text{ and } p^+_{j+\frac{1}{2}}, \\ \{\!\{p\}\!\}_{j+\frac{1}{2}}, & \text{otherwise.} \end{cases} \tag{5.55}$$

Now the viscosity part is given as

$$\mathcal{V}(u_h; v_h) = \sum_{1\le j\le N} (f(u_h^{ex}) - f(u_h^{num}))_{j+\frac{1}{2}} [\![v_h]\!]_{j+\frac{1}{2}} + \alpha(\hat{f}; u_h)[\![u_h]\!][\![v_h]\!], \tag{5.56}$$

and can be estimated by the next lemma. Here and below, we drop the superscript n and $\sharp$ for simplicity.

Lemma 5.12 *There exists a constant $C_\star \ge 0$ solely depending on the maximum of $|f''(u)|$, such that*

$$|\mathcal{V}(u_h; v_h)| \le C_\star\Big[1 + h^{-1}\|e\|_\infty\Big]\Big[\|\xi\| + h^{k+1}\Big]\|v_h\|, \quad \forall\, v_h \in V_h. \tag{5.57}$$

Proof. We can get the bound for $\alpha(\hat{f}; u_h)$ for the three different cases in (5.53). For the last case, there must exist a zero point of $f'(p)$ between u_h^- and u_h^+, so $|f'(\{\!\{u_h\}\!\})| \le C_\star|[\![u_h]\!]|$. Now $\alpha(\hat{f}; u_h)$ is defined in the same way as (5.10) and it satisfies Lemma 5.2. Hence, the inequality (5.11) yields on each element boundary point

$$0 \le \alpha(\hat{f}; u_h) \le C_\star|[\![u_h]\!]| \le C_\star\|e\|_\infty, \tag{5.58}$$

since u is continuous. Obviously, for the other cases, if $f'(s)$ does not changing its sign, $\alpha(\hat{f}; u_h) = 0$ and hence also satisfies (5.58).

Similarly, we can get the bound for $f(u_h^{ex}) - f(u_h^{num})$ by combining the three cases in (5.53) and the three cases in (5.55). If the derivatives are in the same sign, then $f(u_h^{ex}) - f(u_h^{num}) = 0$. Otherwise, there is a zero point of $f'(p)$ in the interval covered by the two limits of the numerical solution $u^{\pm}_{h,j+1/2}$, and the exact solution u in the adjacent elements $I_j \cup I_{j+1}$. Taking $u_{j+1/2}$ as the starting point, one can see the length of this interval is not greater than $Ch + \|e\|_\infty$. Hence $|f'(u_h^{ex})| \le C_\star[Ch + \|e\|_\infty]$. Since $|u_h^{ex} - u_h^{num}| \le |[\![u_h]\!]| = |[\![e]\!]|$, we have

$$\begin{aligned} \|f(u_h^{ex}) - f(u_h^{num})\|_{\Gamma_h} \le& |f'(u_h^{ex})|\|u_h^{ex} - u_h^{num}\|_{\Gamma_h} + C_\star\|u_h^{ex} - u_h^{num}\|^2_{\Gamma_h} \\ \le& C_\star\Big[h + \|e\|_\infty\Big]\|\xi - \eta\|_{\Gamma_h}. \end{aligned} \tag{5.59}$$

It is now straightforward to prove this lemma by using the inverse property (ii), the approximation property (5.5a), and the above bound for $\alpha(\hat{f}; u_h)$ and $f(u_h^{ex}) - f(u_h^{num})$. $\square$

Remark 5.3. For multi-dimensions, the optimal error estimates for the upwind fluxes hold for the tensor product polynomials and meshes, and for special triangulations [4] on which similar Gauss-Radau projections can be defined. The proof is similar to that for the one-dimensional case.

6 Concluding remarks

In this paper we analyze RKDG schemes with TVDRK3 time discretization and present L^2-norm stability for linear conservation laws, and obtain quasi-optimal L^2 error estimates for general monotone fluxes and optimal L^2 error estimates for upwind monotone fluxes for smooth solutions of nonlinear conservation laws. The main technique used in this paper is the energy analysis, which does not require a uniform mesh and can be easily generalized to arbitrary triangulation in multi-dimensions and for linear equations with variable coefficients, as well as to non-periodic boundary conditions.

The error estimates for nonlinear conservation laws in this paper are obtained using stability for the linear case and the smoothness of the exact solution. It is not clear if stability holds for the nonlinear conservation laws with general, non-smooth solutions. Such a stability proof for the fully discrete RKDG schemes is challenging and constitute our ongoing work.

7 Appendix

In this appendix we would like to present a proof for the L^2-norm stability, for linear conservation laws, of the RKDG method with explicit TVDRK2 time-marching. From this analysis, we would like to demonstrate the very different stability mechanisms between the TVDRK2 and TVDRK3 time-marchings, and explain why the stability for TVDRK2 holds only for piecewise linear polynomials.

Let $f(u) = \beta u$ with a constant β. The corresponding fully-discrete RKDG scheme is given as follows: find successively the numerical solution $u_h^{n,1}$ and u_h^{n+1} in the finite element space V_h (see Section 2), such that there holds the formulae

$$\int_{I_j} u_h^{n,1} v_h \,\mathrm{d}x = \int_{I_j} u_h^n v_h \,\mathrm{d}x + \tau \mathcal{H}_j(u_h^n, v_h), \tag{7.1a}$$

$$\int_{I_j} u_h^{n+1} v_h \,\mathrm{d}x = \frac{1}{2}\int_{I_j} u_h^n v_h \,\mathrm{d}x + \frac{1}{2}\int_{I_j} u_h^{n,1} v_h \,\mathrm{d}x + \frac{\tau}{2}\mathcal{H}_j(u_h^{n,1}, v_h), \tag{7.1b}$$

for any test function $v_h \in V_h$ and $j = 1, 2, \ldots, N$, with the initial solution $u_h^0 = \mathbb{P}_h u_0(x)$. The operator $\mathcal{H}_j$ as well as its sum $\mathcal{H}$ over j has been defined in Section 2.

Take the test function $v_h = u_h^n$ and $v_h = u_h^{n,1}$ in (7.1a) and (7.1b), respectively. A simple manipulation yields the energy equation

$$\|u_h^{n+1}\|^2 - \|u_h^n\|^2 = \tau\mathcal{H}(u_h^n, u_h^n) + \tau\mathcal{H}(u_h^{n,1}, u_h^{n,1}) + \|u_h^{n+1} - u_h^{n,1}\|^2. \tag{7.2}$$

This is the starting point of the L^2-norm stability analysis. Below we assume the finite element space V_h consists of piecewise linear polynomials, since the RKDG method with TVDRK2 time-marching is stable in the L^2-norm only for this case.

The first two terms on the right hand side of (7.2) provide the necessary numerical viscosity for stability, same as that for the semi-discrete DG spatial discretization. By (4.5b) in Lemma 4.5a, it reads

$$\tau\mathcal{H}(u_h^n,u_h^n)+\tau\mathcal{H}(u_h^{n,1},u_h^{n,1})=-\frac{\tau}{2}\sum_{1\le j\le N}\Big[|\beta|[\![u_h^n]\!]_{j+\frac12}^2+|\beta|[\![u_h^{n,1}]\!]_{j+\frac12}^2\Big]. \tag{7.3}$$

To obtain the L^2-norm stability, we must control the time discretization contribution $\|u_h^{n+1}-u_h^{n,1}\|^2$ by the numerical viscosity in (7.3). This can be achieved for piecewise linear polynomials under a standard CFL condition.

As is done in Section 4, we denote the combinations of the numerical solution in different stages by $\mathbb{B}_1^n=u_h^{n,1}-u_h^n$ and $\mathbb{B}_2^n=u_h^{n+1}-u_h^{n,1}$, respectively. By subtracting (7.1b) from (7.1a), we get

$$(\mathbb{B}_2^n,v_h)=\frac{\tau}{2}\mathcal{H}(\mathbb{B}_1^n,v_h),\qquad\forall v_h\in V_h. \tag{7.4}$$

We would like to take the test function $v_h=\mathbb{B}_2^n$ in (7.4) to estimate $\|\mathbb{B}_2^n\|$. For piecewise linear polynomials, we do not directly use Lemma 4.1 but make the following treatment to derive a sharper estimate for $\|\mathbb{B}_2^n\|$.

First we take $\phi=\mathbb{B}_1^n$ and $\psi=\mathbb{B}_2^n$ in (4.5a) of Lemma 4.2. Noticing (4.3), we have

$$\begin{aligned}(\mathbb{B}_2^n,\mathbb{B}_2^n)=&\frac{\tau}{2}\mathcal{H}(\mathbb{B}_1^n,\mathbb{B}_2^n)=-\frac{\tau}{2}\Big[\mathcal{H}(\mathbb{B}_2^n,\mathbb{B}_1^n)+\sum_{1\le j\le N}|\beta|[\![\mathbb{B}_1^n]\!]_{j+\frac12}[\![\mathbb{B}_2^n]\!]_{j+\frac12}\Big]\\=&-\frac{\tau}{2}\sum_{1\le j\le N}\Big[\big(|\beta|[\![\mathbb{B}_2^n]\!]+\hat f(\mathbb{B}_2^n)\big)_{j+1/2}[\![\mathbb{B}_1^n]\!]_{j+1/2}+\int_{I_j}\beta\mathbb{B}_2^n\mathbb{B}_{1,x}^n\,\mathrm{d}x\Big].\end{aligned}$$

From (4.2), the definition of $\hat f(v_h)$, it implies $|\beta|[\![\mathbb{B}_2^n]\!]+\hat f(\mathbb{B}_2^n)$ is either equal to $\beta\mathbb{B}_2^{n,-}$ if $\beta>0$, or equal to $\beta\mathbb{B}_2^{n,+}$ if $\beta<0$. Together with the inverse property (ii), we have

$$\begin{aligned}\|\mathbb{B}_2^n\|^2\le&\Big[\frac{\tau}{2}|\beta|\mu_2h^{-\frac12}\sum_{1\le j\le N}[\![\mathbb{B}_1^n]\!]_{j+\frac12}+\frac{\tau}{2}\|\mathbb{B}_{1,x}^n\|\Big]^2\\\le&\frac12|\beta|^2\mu_2^2\tau^2h^{-1}\sum_{1\le j\le N}[\![\mathbb{B}_1^n]\!]_{j+\frac12}^2+\frac12|\beta|^2\tau^2\|\mathbb{B}_{1,x}^n\|^2.\end{aligned} \tag{7.5}$$

Next we estimate $\|\mathbb{B}_{1,x}^n\|$. Denote by $\overline{\mathbb{B}_1^n}$ a piecewise constant which in each element I_j is defined as the cell average of $\mathbb{B}_1^n$. Let $\mathbb{b}_1=\mathbb{B}_1^n-\overline{\mathbb{B}_1^n}$, which also belongs to the finite element space V_h. Taking $v_h=\mathbb{b}_1$ in (7.1a) and using Lemma 4.2, after some manipulation we will get

$$\begin{aligned}\|\mathbb{b}_1\|^2=&(\mathbb{B}_1^n,\mathbb{b}_1)=\tau\mathcal{H}(u_h^n,\mathbb{b}_1)=-\tau\Big[\mathcal{H}(\mathbb{b}_1,u_h^n)+\sum_{1\le j\le N}|\beta|[\![u_h^n]\!]_{j+\frac12}[\![\mathbb{b}_1]\!]_{j+\frac12}\Big]\\=&-\tau\sum_{1\le j\le N}\Big[\big(|\beta|[\![\mathbb{b}_1]\!]+\hat f(\mathbb{b}_1)\big)_{j+1/2}[\![u_h^n]\!]_{j+1/2}+\int_{I_j}\beta\mathbb{b}_1u_{h,x}^n\,\mathrm{d}x\Big],\\=&-\tau\sum_{1\le j\le N}\big(|\beta|[\![\mathbb{b}_1]\!]+\hat f(\mathbb{b}_1)\big)_{j+1/2}[\![u_h]\!]_{j+\frac12}.\end{aligned} \tag{7.6}$$

Note that in the last step we have used the fact that $(u_h^n)_x$ is a piecewise constant, since u_h^n is a piecewise linear polynomial. Similarly, the inverse inequality (ii) gives the important estimate for piecewise linear polynomials

$$\|\mathbb{b}_1\|^2 \leq |\beta|^2\mu_2^2\tau^2h^{-1}\sum_{1\leq j\leq N}[\![u_h^n]\!]_{j+\frac{1}{2}}^2. \tag{7.7}$$

This is a stronger continuity property for the DG spatial discretization with piecewise linear polynomials, in comparison with Lemma 4.1.

Denote $\lambda_{\max} = |\beta|\mu\tau h^{-1}$ as before, where $\mu = \max\{\mu_1, \mu_2^2\}$. Since the inverse inequality (i) implies that $\|\mathbb{B}_{1,x}^n\| = \|g_{h,x}\| \leq \mu_1 h^{-1}\|\mathbb{b}_1\|$, it follows from inequalities (7.5) and (7.7) that there holds for piecewise linear polynomials

$$\begin{aligned}
\|\mathbb{B}_2^n\|^2 &\leq \frac{1}{2}|\beta|^2\mu_2^2\tau^2h^{-1}\sum_{1\leq j\leq N}[\![u_h^{n,1}-u_h^n]\!]_{j+\frac{1}{2}}^2 + |\beta|^4\mu_1^2\mu_2^2\tau^4h^{-3}\sum_{1\leq j\leq N}[\![u_h^n]\!]_{j+\frac{1}{2}}^2\\
&\leq \Big[\lambda_{\max}+\lambda_{\max}^3\Big]\tau\sum_{1\leq j\leq N}|\beta|[\![u_h^n]\!]_{j+\frac{1}{2}}^2 + \lambda_{\max}\tau\sum_{1\leq j\leq N}|\beta|[\![u_h^{n,1}]\!]_{j+\frac{1}{2}}^2\\
&\leq \frac{\tau}{4}\sum_{1\leq j\leq N}|\beta|\Big[[\![u_h^{n,1}]\!]_{j+\frac{1}{2}}^2 + [\![u_h^n]\!]_{j+\frac{1}{2}}^2\Big],
\end{aligned} \tag{7.8}$$

if $\lambda_{\max}$ is less than the maximum positive root of $\lambda + \lambda^3 = 1/4$.

Finally, we collect estimates (7.3) and (7.8) into the energy equation (7.2), to obtain

$$\|u_h^{n,2}\|^2 - \|u_h^n\|^2 + \frac{\tau}{4}\sum_{1\leq j\leq N}\Big[|\beta|[\![u_h^n]\!]_{j+\frac{1}{2}}^2 + |\beta|[\![u_h^{n,1}]\!]_{j+\frac{1}{2}}^2\Big] \leq 0, \tag{7.9}$$

under a suitable CFL condition for piecewise linear polynomials. It implies that the L^2-norm of approximation solution does not increase, so the scheme is stable in the L^2-norm.

We point out that, for higher order piecewise polynomials, we can not obtain the sharp estimate (7.8), because the integration in (7.6) needs a stronger temporal-spatial restriction, for example $\tau = \mathcal{O}(h^{4/3})$. We will not give the detailed analysis here.

References

[1] R. A. Adams, *Sobolev Spaces*, New York, Academic Press, 1975.

[2] P. G. Ciarlet, *Finite Element Method for Elliptic Problems*, North–Holland, Amsterdam, 1978.

[3] B. Cockburn, *An introduction to the discontinuous Galerkin method for convection-dominated problems*, in Advanced Numerical Approximation of Nonlinear Hyperbolic Equations, B. Cockburn, C. Johnson, C.-W. Shu and E. Tadmor (Editor: A. Quarteroni), Lecture Notes in Math. 1697, Springer, Berlin, 1998, pp. 151–268.

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

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

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

[7] B. Cockburn and C.-W. Shu, *TVB Runge-Kutta local projection discontinuous Galerkin finite element method for scalar conservation laws* II: *General framework*, Math. Comp., 52 (1989), pp. 411–435.

[8] B. Cockburn and C.-W. Shu, *The Runge-Kutta local projection* $\mathbb{P}^1$*-discontinuous Galerkin method for scalar conservation laws*, RAIRO Modél. Math. Anal. Numér., 25 (1991), pp. 337–361.

[9] B. Cockburn and C.-W. Shu, *The Runge-Kutta discontinuous Galerkin finite element method for conservation laws* V: *Multidimensional systems*, J. Comput. Phys., 141 (1998), pp. 199–224.

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

[12] G.-S. Jiang and C.-W. Shu, *On cell entropy inequality for discontinuous Galerkin methods*, Math. Comp., 62 (1994), pp. 531–538.

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

[14] P. Lesaint and P. A. Raviart, *On a finite element method for solving the neutron transport equation*, in Mathematical Aspects of Finite Elements in Partial Differential Equations, C. de Boor, ed., Academic Press, New York, 1974, pp. 89–145.

[15] S. Osher, *Riemann solvers, the entropy condition, and difference approximations*, SIAM. J. Numer. Anal., 21 (1984), pp. 217–235.

[16] W. H. Reed and T. R. Hill, *Triangular mesh methods for the neutron transport equation*, Los Alamos Scientific Laboratory report LA-UR-73-479, Los Alamos, NM, 1973

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

[18] 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.

[19] Q. Zhang and C.-W Shu, *Error estimates to smooth solution of Runge-Kutta discontinuous Galerkin methods for symmetrizable system of conservation laws*, SIAM. J. Numer. Anal., 44 (2006), pp. 1702-1720.