

# Maximum-Principle-Satisfying and Positivity-Preserving High Order Discontinuous Galerkin Schemes for Conservation Laws on Triangular Meshes

Xiangxiong Zhang · Yinhua Xia · Chi-Wang Shu

Received: 14 May 2010 / Revised: 27 January 2011 / Accepted: 29 January 2011 /  
Published online: 19 February 2011  
© Springer Science+Business Media, LLC 2011

**Abstract** In Zhang and Shu (J. Comput. Phys. 229:3091–3120, 2010), two of the authors constructed uniformly high order accurate finite volume and discontinuous Galerkin (DG) schemes satisfying a strict maximum principle for scalar conservation laws on rectangular meshes. The technique is generalized to positivity preserving (of density and pressure) high order DG or finite volume schemes for compressible Euler equations in Zhang and Shu (J. Comput. Phys. 229:8918–8934, 2010). The extension of these schemes to triangular meshes is conceptually plausible but highly nontrivial. In this paper, we first introduce a special quadrature rule which is exact for two-variable polynomials over a triangle of a given degree and satisfy a few other conditions, by which we can construct high order maximum principle satisfying finite volume schemes (e.g. essentially non-oscillatory (ENO) or weighted ENO (WENO) schemes) or DG method solving two dimensional scalar conservation laws on triangular meshes. The same method can preserve the maximum principle for DG or finite volume schemes solving two-dimensional incompressible Euler equations in the vorticity stream-function formulation, or any passive convection equation with an incompressible velocity field. We also obtain positivity preserving (for density and pressure) high order DG or finite volume schemes solving compressible Euler equations on triangular meshes. Numerical tests for the third order Runge-Kutta DG (RKDG) method on unstructured meshes are reported.

---

Research supported by AFOSR grant FA9550-09-1-0126 and NSF grant DMS-0809086.

X. Zhang

Department of Mathematics, Brown University, Providence, RI 02912, USA  
e-mail: [zhangxx@dam.brown.edu](mailto:zhangxx@dam.brown.edu)

Y. Xia

Department of Mathematics, University of Science and Technology of China, Hefei, Anhui 230026, China  
e-mail: [yhxia@ustc.edu.cn](mailto:yhxia@ustc.edu.cn)

C.-W. Shu (✉)

Division of Applied Mathematics, Brown University, Providence, RI 02912, USA  
e-mail: [shu@dam.brown.edu](mailto:shu@dam.brown.edu)

**Keywords** Hyperbolic conservation laws · Finite volume scheme · Discontinuous Galerkin method · Essentially non-oscillatory scheme · Weighted essentially non-oscillatory scheme · Maximum principle · Positivity preserving · High order accuracy · Strong stability preserving time discretization · Passive convection equation · Incompressible flow · Compressible Euler equations · Triangular meshes

## 1 Introduction

We first consider numerical solutions of the scalar conservation law

$$u_t + \nabla \cdot \mathbf{F}(u) = 0, \quad u(\mathbf{x}, 0) = u_0(\mathbf{x}), \quad (1.1)$$

where  $u_0(\mathbf{x})$  is assumed to be a bounded variation function. The main difficulty in solving (1.1) is that the solution may contain discontinuities even if the initial condition is smooth, hence we must consider the physically relevant unique weak solution which is called the entropy solution. An important property of the entropy solution is that it satisfies a strict maximum principle (e.g. [4]), i.e., if

$$M = \max_{\mathbf{x}} u_0(\mathbf{x}), \quad m = \min_{\mathbf{x}} u_0(\mathbf{x}), \quad (1.2)$$

then  $u(\mathbf{x}, t) \in [m, M]$  for any  $\mathbf{x}$  and  $t$ .

Successful high order numerical schemes for solving (1.1) includes, among others, the Runge-Kutta discontinuous Galerkin (RKDG) method with a total variation bounded (TVB) limiter [2], the essentially non-oscillatory (ENO) finite volume and finite difference schemes [7, 19], and the weighted ENO (WENO) finite volume and finite difference schemes [8, 11]. Although these schemes are nonlinearly stable in numerical experiments and some of them can be proved to be total variation stable, they do not in general satisfy a strict maximum principle. It is very difficult to obtain a uniformly high order accurate scheme satisfying a strict maximum principle in the sense that the numerical solution never goes out of the range  $[m, M]$ , which is a desired property in some applications, for example, when  $u$  is a volume ratio which should not go outside the range of  $[0, 1]$ .

For hyperbolic conservation law systems, the entropy solution in general does not satisfy any maximum principle. In this paper we are mainly interested in the Euler equations for compressible flows. Physically, the density  $\rho$  and the pressure  $p$  should both be positive. We are interested in positivity-preserving high order schemes, which maintain the positivity of density and pressure at time level  $n + 1$ , provided that they are positive at time level  $n$ . We remark that failure of preserving positivity of density or pressure may cause blow-ups of the numerical algorithm, for example, for low density problems in computing blast waves, and low pressure problems in computing high Mach number astrophysical jets [6]. We also remark that most commonly used high order numerical schemes for solving Euler equations, for instance, the RKDG method with TVB limiter, ENO and WENO schemes do not in general satisfy the positivity property for Euler equations automatically.

Our objective is to design genuinely high order schemes (the order of accuracy of the scheme is uniformly high for all smooth solutions including at all kinds of extrema) which satisfy strict maximum principles for scalar conservation laws and positivity preserving property for compressible Euler systems in multispace dimensions. In [22], two of the authors established a general framework to construct a genuinely high order accurate maximum-principle-satisfying scheme for one dimensional and multidimensional scalar

conservation laws on rectangular meshes, in the sense that the numerical solution never goes out of the range  $[m, M]$  of the initial condition. In [23], the same authors constructed a genuinely high order accurate positivity preserving scheme for Euler systems by generalizing the techniques of limiters in [22] and the positivity-preserving schemes in [14]. As far as we know, this is the first time such objective is achieved, with a simple recipe which does not require the costly and unrealistic (in multidimension) “following the characteristics” procedure or exact time evolution.

In this paper, we first introduce an interesting quadrature rule for two variable polynomials on a triangle, by which we can obtain a sufficient condition for a finite volume or a DG method with Euler forward time discretization solving two dimensional scalar conservation laws on triangular meshes to satisfy a strict maximum principle. The same type of the linear scaling limiter as in [22] can enforce this condition without destroying accuracy and conservativity under suitable CFL condition. Strong stability preserving (SSP) high order time discretizations will keep the maximum principle. Then we show that the algorithm and conclusion are also valid for two dimensional incompressible Euler equations in the vorticity stream-function formulation, or for any passive convection equation with an incompressible velocity field. This framework is then used to extend the positivity preserving finite volume or DG schemes in [23] for compressible Euler equations from rectangular meshes to triangular meshes.

This paper is organized as follows: First we illustrate the ideas in [22, 23] by reviewing the one-dimensional schemes briefly in Sect. 2. Then we describe and prove the maximum principle for an arbitrarily high order scheme for scalar conservation laws on triangular meshes in two-dimensional space in Sect. 3. Section 4 is the application of the scheme to two dimensional incompressible Euler equations in the vorticity stream-function formulation, and to any passive convection equation with an incompressible velocity field. In Sect. 5, we generalize the idea in Sect. 3 to design positivity-preserving schemes for Euler systems. Numerical tests for the DG method will be shown in Sect. 6. Concluding remarks are given in Sect. 7.

## 2 Maximum-Principle-Satisfying and Positivity-Preserving High Order Schemes in One Dimension

In this section we briefly review the techniques introduced in [22, 23] for constructing maximum-principle-satisfying or positivity-preserving schemes in one space dimension.

### 2.1 Maximum-Principle-Satisfying High Order Schemes for Scalar Conservation Laws

We consider only the first order Euler forward time discretization in this subsection; higher order time discretization will be discussed later.

A finite volume scheme or a scheme satisfied by the cell averages of a DG method solving

$$u_t + f(u)_x = 0, \quad u(x, 0) = u_0(x),$$

can be written as

$$\bar{u}_j^{n+1} = \bar{u}_j^n - \lambda [h(u_{j+\frac{1}{2}}^-, u_{j+\frac{1}{2}}^+) - h(u_{j-\frac{1}{2}}^-, u_{j-\frac{1}{2}}^+)], \quad (2.1)$$

where  $n$  refers to the time step and  $j$  to the spatial cell (we assume uniform mesh size only for simplicity), and  $\lambda = \frac{\Delta t}{\Delta x}$  is the ratio of time and space mesh size.  $\bar{u}_j^n$  is the approximation

to the cell averages of  $u(x, t)$  in the cell  $I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$  at time level  $n$ , and  $u_{j+\frac{1}{2}}^-, u_{j+\frac{1}{2}}^+$  are the high order approximations of the nodal values  $u(x_{j+\frac{1}{2}}, t^n)$  within the cells  $I_j$  and  $I_{j+1}$  respectively. These values are either reconstructed from the cell averages  $\bar{u}_j^n$  in a finite volume method or read directly from the evolved polynomials in a DG method. We assume that there is a polynomial  $p_j(x)$  (either reconstructed in a finite volume method or evolved in a DG method) with degree  $k$ , defined on  $I_j$  such that  $\bar{u}_j^n$  is the cell average of  $p_j(x)$  on  $I_j$ ,  $u_{j-\frac{1}{2}}^+ = p_j(x_{j-\frac{1}{2}})$  and  $u_{j+\frac{1}{2}}^- = p_j(x_{j+\frac{1}{2}})$ . The numerical flux function  $h(\cdot, \cdot)$  is chosen to be a Lipschitz continuous monotone flux, i.e.,  $h(\cdot, \cdot)$  is nondecreasing in its first argument and nonincreasing in its second argument. For instance, the global Lax-Friedrichs flux defined by

$$h(u, v) = \frac{1}{2}[f(u) + f(v) - a(v - u)], \quad a = \max |f'(u)|, \quad (2.2)$$

where the maximum is taken over the whole region where  $u$  and  $v$  vary, is a monotone flux. In the rest of the paper, we will mainly use the global Lax-Friedrichs flux as an example to illustrate the idea, although any other monotone flux will also work.

Consider the  $N$ -point Legendre Gauss-Lobatto quadrature rule on the interval  $I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$ , which is exact for the integral of polynomials of degree up to  $2N - 3$ . We denote these quadrature points on  $I_j$  as

$$S_j = \{x_{j-\frac{1}{2}} = \hat{x}_j^1, \hat{x}_j^2, \dots, \hat{x}_j^{N-1}, \hat{x}_j^N = x_{j+\frac{1}{2}}\}. \quad (2.3)$$

Define  $\hat{v}_\alpha = p_j(\hat{x}_j^\alpha)$  for  $\alpha = 1, \dots, N$ , and let  $\hat{w}_\alpha$  be the quadrature weights for the interval  $[-\frac{1}{2}, \frac{1}{2}]$  such that  $\sum_{\alpha=1}^N \hat{w}_\alpha = 1$ . Choose  $N$  to be the smallest integer satisfying  $2N - 3 \geq k$ , then

$$\bar{u}_j^n = \frac{1}{\Delta x} \int_{I_j} p_j(x) dx = \sum_{\alpha=1}^N \hat{w}_\alpha \hat{v}_\alpha = \sum_{\alpha=2}^{N-1} \hat{w}_\alpha \hat{v}_\alpha + \hat{w}_1 u_{j-\frac{1}{2}}^+ + \hat{w}_N u_{j+\frac{1}{2}}^-. \quad (2.4)$$

By adding and subtracting  $h(u_{j-\frac{1}{2}}^+, u_{j+\frac{1}{2}}^-)$ , and plugging (2.4) into the scheme (2.1), we obtain

$$\begin{aligned} \bar{u}_j^{n+1} &= \sum_{\alpha=2}^{N-1} \hat{w}_\alpha \hat{v}_\alpha + \hat{w}_1 u_{j-\frac{1}{2}}^+ + \hat{w}_N u_{j+\frac{1}{2}}^- - \lambda \left[ h(u_{j+\frac{1}{2}}^-, u_{j+\frac{1}{2}}^+) - h(u_{j-\frac{1}{2}}^+, u_{j+\frac{1}{2}}^-) \right] \\ &\quad - \lambda \left[ h(u_{j-\frac{1}{2}}^+, u_{j+\frac{1}{2}}^-) - h(u_{j-\frac{1}{2}}^-, u_{j-\frac{1}{2}}^+) \right] \\ &= \sum_{\alpha=2}^{N-1} \hat{w}_\alpha \hat{v}_\alpha + \hat{w}_N \left( u_{j+\frac{1}{2}}^- - \frac{\lambda}{\hat{w}_N} [h(u_{j+\frac{1}{2}}^-, u_{j+\frac{1}{2}}^+) - h(u_{j-\frac{1}{2}}^+, u_{j+\frac{1}{2}}^-)] \right) \\ &\quad + \hat{w}_1 \left( u_{j-\frac{1}{2}}^+ - \frac{\lambda}{\hat{w}_1} [h(u_{j-\frac{1}{2}}^+, u_{j+\frac{1}{2}}^-) - h(u_{j-\frac{1}{2}}^-, u_{j-\frac{1}{2}}^+)] \right) \\ &= \sum_{\alpha=2}^{N-1} \hat{w}_\alpha \hat{v}_\alpha + \hat{w}_1 H_1 + \hat{w}_N H_N \end{aligned}$$

where

$$H_1 = u_{j-\frac{1}{2}}^+ - \frac{\lambda}{\widehat{w}_1} [h(u_{j-\frac{1}{2}}^+, u_{j+\frac{1}{2}}^-) - h(u_{j-\frac{1}{2}}^-, u_{j-\frac{1}{2}}^+)], \quad (2.5)$$

$$H_N = u_{j+\frac{1}{2}}^- - \frac{\lambda}{\widehat{w}_N} [h(u_{j+\frac{1}{2}}^-, u_{j+\frac{1}{2}}^+) - h(u_{j-\frac{1}{2}}^+, u_{j+\frac{1}{2}}^-)]. \quad (2.6)$$

Assume all point values  $u_{j+\frac{1}{2}}^-, u_{j+\frac{1}{2}}^+, u_{j-\frac{1}{2}}^-, u_{j-\frac{1}{2}}^+, \widehat{v}_2, \widehat{v}_3, \dots, \widehat{v}_{N-1}$  are in the range  $[m, M]$ . Notice that (2.5) and (2.6) are two formal monotone schemes, and  $\widehat{w}_1 = \widehat{w}_N$ , therefore  $H_1, H_N \in [m, M]$  under the CFL condition  $\lambda a \leq \widehat{w}_1$  where  $a = \max |f'(u)|$ . Since  $\bar{u}_j^{n+1}$  is a convex combination of  $H_1, H_N$  and  $\widehat{v}_\alpha$  ( $\alpha = 2, \dots, N-1$ ), we get the maximum principle  $\bar{u}_j^{n+1} \in [m, M]$ . Thus a sufficient condition for (2.1) to satisfy  $\bar{u}_j^{n+1} \in [m, M]$  is  $p_j(\widehat{x}_j^\alpha) \in [m, M]$  for all  $j$  and  $\alpha$ , which could be enforced by a simple limiter without destroying accuracy and conservativity [22].

## 2.2 Positivity-Preserving High Order Schemes for Euler Equations

We consider only the first order Euler forward time discretization in this subsection; higher order time discretization will be discussed later.

The one dimensional Euler system for the perfect gas is given by

$$\mathbf{w}_t + \mathbf{f}(\mathbf{w})_x = 0, \quad t \geq 0, \quad x \in \mathbb{R}, \quad (2.7)$$

$$\mathbf{w} = \begin{pmatrix} \rho \\ m \\ E \end{pmatrix}, \quad \mathbf{f}(\mathbf{w}) = \begin{pmatrix} m \\ \rho u^2 + p \\ (E + p)u \end{pmatrix} \quad (2.8)$$

with

$$m = \rho u, \quad E = \frac{1}{2} \rho u^2 + \rho e, \quad p = (\gamma - 1) \rho e,$$

where  $\rho$  is the density,  $u$  is the velocity,  $m$  is the momentum,  $E$  is the total energy,  $p$  is the pressure,  $e$  is the internal energy, and  $\gamma > 1$  is a constant ( $\gamma = 1.4$  for the air). The speed of sound is given by  $c = \sqrt{\gamma p / \rho}$  and the three eigenvalues of the Jacobian  $\mathbf{f}'(\mathbf{w})$  are  $u - c$ ,  $u$  and  $u + c$ .

Let  $p(\mathbf{w}) = (\gamma - 1)(E - \frac{1}{2} \frac{m^2}{\rho})$  be the pressure function. It can be easily verified that  $p$  is a concave function of  $\mathbf{w} = (\rho, m, E)^T$  if  $\rho \geq 0$ . Define the set of admissible states by

$$G = \left\{ \mathbf{w} = \begin{pmatrix} \rho \\ m \\ E \end{pmatrix} \middle| \rho > 0 \text{ and } p = (\gamma - 1) \left( E - \frac{1}{2} \frac{m^2}{\rho} \right) > 0 \right\},$$

then  $G$  is a convex set. If the density or pressure becomes negative, the system (2.7) will be non-hyperbolic and thus the initial value problem will be ill-posed. Even though we discuss only the perfect gas in this paper, the technique applies to general gases provided the equation of state allows  $G$  to be a convex set.

We are interested in schemes for (2.7) producing the numerical solutions in the admissible set  $G$ . We start with a first order scheme

$$\mathbf{w}_j^{n+1} = \mathbf{w}_j^n - \lambda [\mathbf{h}(\mathbf{w}_j^n, \mathbf{w}_{j+1}^n) - \mathbf{h}(\mathbf{w}_{j-1}^n, \mathbf{w}_j^n)], \quad (2.9)$$

where  $\mathbf{h}(\cdot, \cdot)$  is a numerical flux. The scheme (2.9) and its numerical flux  $\mathbf{h}(\cdot, \cdot)$  are called positivity preserving, if the numerical solution  $\mathbf{w}_j^n$  being in the set  $G$  for all  $j$  implies the solution  $\mathbf{w}_j^{n+1}$  being also in the set  $G$ . This is usually achieved under a standard CFL condition

$$\lambda \|(|u| + c)\|_\infty \leq \alpha_0 \quad (2.10)$$

where  $\alpha_0$  is a constant related to the specific scheme. Examples of positivity preserving fluxes include the Godunov flux, the Lax-Friedrichs flux, the Boltzmann type flux, and the Harten-Lax-van Leer flux, see [14].

We now consider a general high order finite volume scheme, or the scheme satisfied by the cell averages of a DG method solving (2.7), which has the following form

$$\bar{\mathbf{w}}_j^{n+1} = \bar{\mathbf{w}}_j^n - \lambda \left[ \mathbf{h} \left( \mathbf{w}_{j+\frac{1}{2}}^-, \mathbf{w}_{j+\frac{1}{2}}^+ \right) - \mathbf{h} \left( \mathbf{w}_{j-\frac{1}{2}}^-, \mathbf{w}_{j-\frac{1}{2}}^+ \right) \right], \quad (2.11)$$

where  $\mathbf{h}$  is a positivity preserving flux under the CFL condition (2.10),  $\bar{\mathbf{w}}_j^n$  is the approximation to the cell average of the exact solution  $\mathbf{v}(x, t)$  in the cell  $I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$  at time level  $n$ , and  $\mathbf{w}_{j+\frac{1}{2}}^-, \mathbf{w}_{j+\frac{1}{2}}^+$  are the high order approximations of the point values  $\mathbf{v}(x_{j+\frac{1}{2}}, t^n)$  within the cells  $I_j$  and  $I_{j+1}$  respectively. These values are either reconstructed from the cell averages  $\bar{\mathbf{w}}_j^n$  in a finite volume method or read directly from the evolved polynomials in a DG method. We assume that there is a polynomial vector  $\mathbf{q}_j(x) = (\rho_j(x), m_j(x), E_j(x))^T$  (either reconstructed in a finite volume method or evolved in a DG method) with degree  $k$ , where  $k \geq 1$ , defined on  $I_j$  such that  $\bar{\mathbf{w}}_j^n$  is the cell average of  $\mathbf{q}_j(x)$  on  $I_j$ ,  $\mathbf{w}_{j-\frac{1}{2}}^+ = \mathbf{q}_j(x_{j-\frac{1}{2}})$  and  $\mathbf{w}_{j+\frac{1}{2}}^- = \mathbf{q}_j(x_{j+\frac{1}{2}})$ . Next, we show a similar result as in the previous subsection, i.e., the sufficient condition for the scheme (2.11) to satisfy  $\bar{\mathbf{w}}_j^{n+1} \in G$ , is that  $\mathbf{q}_j(\hat{x}_j^\alpha) \in G$  for all  $j$  and  $\alpha$ .

The  $N$ -point Legendre Gauss-Lobatto implies

$$\bar{\mathbf{w}}_j^n = \frac{1}{\Delta x} \int_{I_j} \mathbf{q}_j(x) dx = \sum_{\alpha=1}^N \hat{w}_\alpha \mathbf{q}_j(\hat{x}_j^\alpha) = \sum_{\alpha=2}^{N-1} \hat{w}_\alpha \mathbf{q}_j(\hat{x}_j^\alpha) + \hat{w}_1 \mathbf{w}_{j-\frac{1}{2}}^+ + \hat{w}_N \mathbf{w}_{j+\frac{1}{2}}^-.$$

By adding and subtracting  $\mathbf{h}(\mathbf{w}_{j-\frac{1}{2}}^+, \mathbf{w}_{j+\frac{1}{2}}^-)$ , the scheme (2.11) becomes

$$\begin{aligned} \bar{\mathbf{w}}_j^{n+1} &= \sum_{\alpha=1}^N \hat{w}_\alpha \mathbf{q}_j(\hat{x}_j^\alpha) - \lambda \left[ \mathbf{h} \left( \mathbf{w}_{j+\frac{1}{2}}^-, \mathbf{w}_{j+\frac{1}{2}}^+ \right) - \mathbf{h} \left( \mathbf{w}_{j-\frac{1}{2}}^+, \mathbf{w}_{j+\frac{1}{2}}^- \right) \right. \\ &\quad \left. + \mathbf{h} \left( \mathbf{w}_{j-\frac{1}{2}}^+, \mathbf{w}_{j+\frac{1}{2}}^- \right) - \mathbf{h} \left( \mathbf{w}_{j-\frac{1}{2}}^-, \mathbf{w}_{j-\frac{1}{2}}^+ \right) \right] \\ &= \sum_{\alpha=2}^{N-1} \hat{w}_\alpha \mathbf{q}_j(\hat{x}_j^\alpha) + \hat{w}_N \left( \mathbf{w}_{j+\frac{1}{2}}^- - \frac{\lambda}{\hat{w}_N} \left[ \mathbf{h} \left( \mathbf{w}_{j+\frac{1}{2}}^-, \mathbf{w}_{j+\frac{1}{2}}^+ \right) - \mathbf{h} \left( \mathbf{w}_{j-\frac{1}{2}}^+, \mathbf{w}_{j+\frac{1}{2}}^- \right) \right] \right) \\ &\quad + \hat{w}_1 \left( \mathbf{w}_{j-\frac{1}{2}}^+ - \frac{\lambda}{\hat{w}_1} \left[ \mathbf{h} \left( \mathbf{w}_{j-\frac{1}{2}}^+, \mathbf{w}_{j+\frac{1}{2}}^- \right) - \mathbf{h} \left( \mathbf{w}_{j-\frac{1}{2}}^-, \mathbf{w}_{j-\frac{1}{2}}^+ \right) \right] \right) \\ &= \sum_{\alpha=2}^{N-1} \hat{w}_\alpha \mathbf{q}_j(\hat{x}_j^\alpha) + \hat{w}_N \mathbf{H}_N + \hat{w}_1 \mathbf{H}_1, \end{aligned}$$

where

$$\mathbf{H}_1 = \mathbf{w}_{j-\frac{1}{2}}^+ - \frac{\lambda}{\widehat{w}_1} \left[ \mathbf{h} \left( \mathbf{w}_{j-\frac{1}{2}}^+, \mathbf{w}_{j+\frac{1}{2}}^- \right) - \mathbf{h} \left( \mathbf{w}_{j-\frac{1}{2}}^-, \mathbf{w}_{j+\frac{1}{2}}^+ \right) \right], \quad (2.12)$$

$$\mathbf{H}_N = \mathbf{w}_{j+\frac{1}{2}}^- - \frac{\lambda}{\widehat{w}_N} \left[ \mathbf{h} \left( \mathbf{w}_{j+\frac{1}{2}}^-, \mathbf{w}_{j+\frac{1}{2}}^+ \right) - \mathbf{h} \left( \mathbf{w}_{j-\frac{1}{2}}^+, \mathbf{w}_{j+\frac{1}{2}}^- \right) \right]. \quad (2.13)$$

Notice that (2.12) and (2.13) are both of the type (2.9), and  $\widehat{w}_1 = \widehat{w}_N$ , therefore  $\mathbf{H}_1$  and  $\mathbf{H}_N$  are in the set  $G$  under the CFL condition

$$\lambda \|(|u| + c)\|_\infty \leq \widehat{w}_1 \alpha_0.$$

Now, it is easy to conclude that  $\overline{\mathbf{w}}_j^{n+1}$  is in  $G$ , since it is a convex combination of elements in  $G$ . The sufficient condition can be enforced by the linear scaling limiter [23].

### 2.3 Higher Order Time Discretization

We will use total variation diminishing (TVD), also referred to as strong stability preserving (SSP) high order time discretizations. For more details, see [5, 18, 19]. For example, the third order TVD Runge-Kutta method [19] (with the CFL coefficient  $c = 1$ ) is

$$\begin{aligned} u^{(1)} &= u^n + \Delta t F(u^n), \\ u^{(2)} &= \frac{3}{4}u^n + \frac{1}{4}(u^{(1)} + \Delta t F(u^{(1)})), \\ u^{n+1} &= \frac{1}{3}u^n + \frac{2}{3}(u^{(2)} + \Delta t F(u^{(2)})) \end{aligned} \quad (2.14)$$

where  $F(u)$  is the spatial operator, and the third order TVD multi-step method [18] (with the CFL coefficient  $c = \frac{1}{3}$ ) is

$$u^{n+1} = \frac{16}{27}(u^n + 3\Delta t F(u^n)) + \frac{11}{27}\left(u^{n-3} + \frac{12}{11}\Delta t F(u^{n-3})\right). \quad (2.15)$$

Here, the CFL coefficient  $c$  for a SSP time discretization refers to the fact that, if we assume the Euler forward time discretization for solving the equation  $u_t = F(u)$  is stable in a norm or a semi-norm under a time step restriction  $\Delta t \leq \Delta t_0$ , then the high order SSP time discretization is also stable in the same norm or semi-norm under the time step restriction  $\Delta t \leq c\Delta t_0$ .

Since a SSP high order time discretization is a convex combinations of Euler forward, the full scheme with a high order SSP time discretization will still satisfy the maximum principle and the positivity preserving property.

## 3 High Order Schemes Satisfying the Maximum Principle on Triangular Meshes

### 3.1 Preliminaries and the First Order Monotone Scheme

Consider two-dimensional scalar conservation laws (1.1) with  $\mathbf{F}(u) = \langle f(u), g(u) \rangle$  and the first order monotone schemes on a triangulation. We only discuss the global Lax-Friedrichs

scheme here, but similar results hold for other monotone schemes (e.g. local Lax-Friedrichs scheme). For each triangle  $K$  we denote by  $l_K^i$  ( $i = 1, 2, 3$ ) the length of its three edges  $e_K^i$  ( $i = 1, 2, 3$ ), with outward unit normal vector  $v^i$  ( $i = 1, 2, 3$ ).  $K(i)$  denotes the neighboring triangle along  $e_K^i$  and  $|K|$  is the area of the triangle  $K$ . The global Lax-Friedrichs flux is defined by

$$h(u, v, v) = \frac{1}{2}(\mathbf{F}(u) \cdot v + \mathbf{F}(v) \cdot v - a(v - u))$$

where

$$a = \max_{u,v} |\mathbf{F}'(u) \cdot v|.$$

It satisfies the conservativity and consistency

$$h(u, v, v) = -h(v, u, -v), \quad h(u, u, v) = \mathbf{F}(u) \cdot v. \quad (3.1)$$

The first order Lax-Friedrichs scheme can be written as

$$u_K^{n+1} = u_K^n - \frac{\Delta t}{|K|} \sum_{i=1}^3 h(u_K^n, u_{K(i)}^n, v^i) l_K^i = H(u_K^n, u_{K(1)}^n, u_{K(2)}^n, u_{K(3)}^n).$$

Then  $H(\cdot, \cdot, \cdot, \cdot)$  is a monotone increasing function with respect to each argument under the CFL condition

$$a \frac{\Delta t}{|K|} \sum_{i=1}^3 l_K^i \leq 1.$$

Now let us consider high order schemes. We only discuss Euler forward time discretization, see Sect. 2.3 for higher order time discretization. A finite volume scheme or a scheme satisfied by the cell averages of a DG method, with first order Euler forward time discretization, can be written as

$$\bar{u}_K^{n+1} = \bar{u}_K^n - \frac{\Delta t}{|K|} \sum_{i=1}^3 \int_{e_K^i} h(u_i^{int(K)}, u_i^{ext(K)}, v^i) ds,$$

where  $\bar{u}_K^n$  is the cell average over  $K$  of the numerical solution, and  $u_i^{int(K)}, u_i^{ext(K)}$  are the approximations to the values on the edge  $e_K^i$  obtained from the interior and the exterior of  $K$ . For simplicity, we only discuss the DG method from now on, but all the results also hold for the finite volume scheme (e.g. ENO and WENO). Assume the DG polynomial on the triangle  $K$  is  $p_K(x, y)$  of degree  $k$ , then in the DG method, the edge integral should be approximated by the  $(k+1)$ -point Gauss quadrature. The scheme becomes

$$\bar{u}_K^{n+1} = \bar{u}_K^n - \frac{\Delta t}{|K|} \sum_{i=1}^3 \sum_{\beta=1}^{k+1} h(u_{i,\beta}^{int(K)}, u_{i,\beta}^{ext(K)}, v^i) w_\beta l_K^i, \quad (3.2)$$

where  $w_\beta$  denote the  $(k+1)$ -point Gauss quadrature weights on the interval  $[-\frac{1}{2}, \frac{1}{2}]$ , so that  $\sum_{\beta=1}^{k+1} w_\beta = 1$ , and  $u_{i,\beta}^{int(K)}$  and  $u_{i,\beta}^{ext(K)}$  denote the values of  $u$  evaluated at the  $\beta$ -th Gauss quadrature point on the  $i$ -th edge from the interior and exterior of the element  $K$  respectively.

In the next subsections, we will rewrite the right hand side of (3.2) as a monotone increasing function of some point values of  $p_K(x, y)$  under a certain CFL condition.



### 3.2 Decomposition of the Cell Average

Motivated by the ideas in Sect. 2, the first step is to decompose the cell average  $\bar{u}_K^n$  as a convex combination of point values of the DG polynomial  $p_K(x, y)$ , which could be achieved by a quadrature. Therefore, we are interested in a quadrature rule satisfying:

- The quadrature rule is exact for integration of  $p_K(x, y)$  on  $K$ .
- All the quadrature weights should be positive.
- The quadrature points include all the Gauss quadrature points for each edge  $e_K^i$ . We would also need to find the quadrature weights for these points, since the CFL condition of our method will depend on them.

For convenience, we use the position vectors to denote the three vertices of  $K$ :  $\mathbf{V}_1, \mathbf{V}_2$  and  $\mathbf{V}_3$ . The position vector  $\mathbf{P}$  of an arbitrary point  $P$  in  $K$  can be specified by the barycentric coordinates  $(\xi_1, \xi_2, \xi_3)$ , where  $\mathbf{P} = \xi_1 \mathbf{V}_1 + \xi_2 \mathbf{V}_2 + \xi_3 \mathbf{V}_3$ .

First, consider the quadrature rule on the unit square  $R$  with vertices  $(-\frac{1}{2}, -\frac{1}{2}), (\frac{1}{2}, -\frac{1}{2}), (\frac{1}{2}, \frac{1}{2})$ , and  $(-\frac{1}{2}, \frac{1}{2})$  in the  $u$ - $v$  plane. Let  $N$  be the smallest integer such that  $2N - 3 \geq k$ , then the  $N$ -point Gauss-Lobatto quadrature rule is exact for a single variable polynomial of degree  $k$ . Let  $\{v^\beta : \beta = 1, \dots, k+1\}$  denote the Gauss quadrature points on  $[-\frac{1}{2}, \frac{1}{2}]$  with weights  $w_\beta$ , and  $\{\hat{u}^\alpha : \alpha = 1, \dots, N\}$  denote the Gauss-Lobatto quadrature points on  $[-\frac{1}{2}, \frac{1}{2}]$  with weights  $\hat{w}_\alpha$ . For a two-variable polynomial  $p(u, v)$ , we can use the tensor product of  $N$ -point Gauss-Lobatto for  $u$  and  $(k+1)$ -point Gauss for  $v$  as the quadrature rule on the square, then the quadrature points can be written as  $S_k = \{(\hat{u}^\alpha, v^\beta) : \alpha = 1, \dots, N; \beta = 1, \dots, k+1\}$ , see Fig. 1(a) for  $S_2$ . This quadrature is exact for a polynomial  $p(u, v)$  if the degree of  $p(u, v)$  with respect to  $u$  is not larger than  $k$  and the degree with respect to  $v$  is not larger than  $2k+1$ .

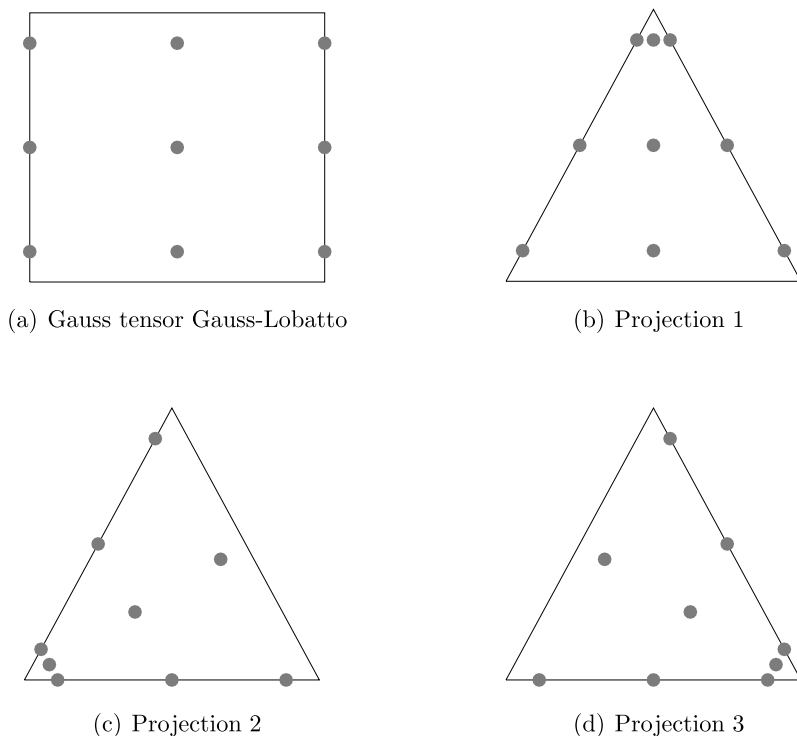
Without loss of the generality, we assume the orientation of the three vertices  $\mathbf{V}_1, \mathbf{V}_2$  and  $\mathbf{V}_3$  is clockwise in this subsection, then we define the following three functions:

$$\begin{aligned} \mathbf{g}_1(u, v) &= \left(\frac{1}{2} + v\right) \mathbf{V}_1 + \left(\frac{1}{2} + u\right) \left(\frac{1}{2} - v\right) \mathbf{V}_2 + \left(\frac{1}{2} - u\right) \left(\frac{1}{2} - v\right) \mathbf{V}_3, \\ \mathbf{g}_2(u, v) &= \left(\frac{1}{2} + v\right) \mathbf{V}_2 + \left(\frac{1}{2} + u\right) \left(\frac{1}{2} - v\right) \mathbf{V}_3 + \left(\frac{1}{2} - u\right) \left(\frac{1}{2} - v\right) \mathbf{V}_1, \\ \mathbf{g}_3(u, v) &= \left(\frac{1}{2} + v\right) \mathbf{V}_3 + \left(\frac{1}{2} + u\right) \left(\frac{1}{2} - v\right) \mathbf{V}_1 + \left(\frac{1}{2} - u\right) \left(\frac{1}{2} - v\right) \mathbf{V}_2. \end{aligned}$$

Each of them is a projection from the square to  $K$ , mapping the top edge of the rectangle into one vertex and the other three edges to the edges of  $K$ . We will use  $\mathbf{g}_i$  ( $i = 1, 2, 3$ ) and  $S_k$  to construct our triangle quadrature. Let  $p_K(x, y)$  be the two-variable DG polynomial of degree  $k$  with cell average  $\bar{u}_K^n$  defined on the triangle  $K$ , then

$$\bar{u}_K^n = \frac{1}{|K|} \int \int_K p_K(x, y) dA(x, y) = \frac{1}{|K|} \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} p_K(\mathbf{g}_i(u, v)) \left| \frac{\partial \mathbf{g}_i(u, v)}{\partial(u, v)} \right| du dv, \\ i = 1, 2, 3.$$

**Lemma 3.1** *If the orientation of the three vertices  $\mathbf{V}_1, \mathbf{V}_2$  and  $\mathbf{V}_3$  is clockwise, then the Jacobian  $\left| \frac{\partial \mathbf{g}_i(u, v)}{\partial(u, v)} \right| = 2|K|(\frac{1}{2} - v)$ .*



**Fig. 1** Illustration of the three projections for  $k = 2$

*Proof* We only calculate the Jacobian of  $\mathbf{g}_1$  here. Notice that we have

$$\frac{\partial \mathbf{g}_1}{\partial u} = \left(\frac{1}{2} - v\right)(\mathbf{V}_2 - \mathbf{V}_3),$$

$$\frac{\partial \mathbf{g}_1}{\partial v} = \mathbf{V}_1 - \left(\frac{1}{2} + u\right)\mathbf{V}_2 - \left(\frac{1}{2} - u\right)\mathbf{V}_3 = (\mathbf{V}_1 - \mathbf{V}_3) - \left(\frac{1}{2} + u\right)(\mathbf{V}_2 - \mathbf{V}_3).$$

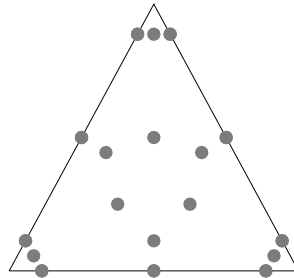
So  $|\frac{\partial \mathbf{g}_1(u,v)}{\partial(u,v)}| = \frac{\partial \mathbf{g}_1}{\partial u} \times \frac{\partial \mathbf{g}_1}{\partial v} = \left(\frac{1}{2} - v\right)(\mathbf{V}_2 - \mathbf{V}_3) \times (\mathbf{V}_1 - \mathbf{V}_3) = 2|K|(\frac{1}{2} - v)$ .  $\square$

Define  $\hat{p}_K^i(u, v) = p_K(\mathbf{g}_i(u, v))|\frac{\partial \mathbf{g}_i(u,v)}{\partial(u,v)}|$ , then it is still a polynomial of  $u$  and  $v$ . Moreover, the degree of this polynomial with respect to  $u$  and  $v$  is  $k$  and  $k + 2$ , therefore the double integral in  $u$  and  $v$  is equal to the quadrature of  $S_k$ . Thus,

$$\bar{u}_K^n = \frac{1}{|K|} \sum_{\alpha=1}^N \sum_{\beta=1}^{k+1} \hat{p}_K^i(\hat{u}^\alpha, v^\beta) w_\alpha \hat{w}_\beta = \sum_{\alpha=1}^N \sum_{\beta=1}^{k+1} p_K(\mathbf{g}_i(\hat{u}^\alpha, v^\beta)) 2\left(\frac{1}{2} - v^\beta\right) w_\alpha \hat{w}_\beta.$$

So we have three different quadrature rules for  $p_K^i(x, y)$  over  $K$ , and the quadrature points are  $\mathbf{g}_i(S_k)$  ( $i = 1, 2, 3$ ) with positive quadrature weights, see Fig. 1(b), (c), (d) for  $k = 2$ . By combining the points of the three quadrature rules, we obtain a  $3(N - 1)(k + 1)$ -point quadrature rule which includes all the Gauss quadrature points for three edges, see Fig. 2

**Fig. 2** The quadrature points on a triangle for  $k = 2$



for the quadrature points  $S_k^K = \mathbf{g}_1(S_k) \cup \mathbf{g}_2(S_k) \cup \mathbf{g}_3(S_k)$  with  $k = 2$ . In the barycentric coordinates, the set  $S_k^K$  can be written as

$$S_k^K = \left\{ \left( \frac{1}{2} + v^\beta, \left( \frac{1}{2} + \widehat{u}^\alpha \right) \left( \frac{1}{2} - v^\beta \right), \left( \frac{1}{2} - \widehat{u}^\alpha \right) \left( \frac{1}{2} - v^\beta \right) \right), \right. \\ \left( \left( \frac{1}{2} - \widehat{u}^\alpha \right) \left( \frac{1}{2} - v^\beta \right), \frac{1}{2} + v^\beta, \left( \frac{1}{2} + \widehat{u}^\alpha \right) \left( \frac{1}{2} - v^\beta \right) \right), \\ \left( \left( \frac{1}{2} + \widehat{u}^\alpha \right) \left( \frac{1}{2} - v^\beta \right), \left( \frac{1}{2} - \widehat{u}^\alpha \right) \left( \frac{1}{2} - v^\beta \right), \frac{1}{2} + v^\beta \right) : \\ \left. \alpha = 1, \dots, N; \beta = 1, \dots, k+1 \right\}. \quad (3.3)$$

The cell average can now be decomposed as:

$$\bar{u}_K^n = \frac{1}{3} \sum_{i=1}^3 \bar{u}_K^n = \sum_{i=1}^3 \sum_{\alpha=1}^N \sum_{\beta=1}^{k+1} p_K(\mathbf{g}_i(\widehat{u}^\alpha, v^\beta)) \frac{2}{3} \left( \frac{1}{2} - v^\beta \right) w_\alpha \widehat{w}_\beta = \sum_{\mathbf{x} \in S_k^K} p_K(\mathbf{x}) w_{\mathbf{x}}. \quad (3.4)$$

Next, we would like to find the quadrature weights  $w_{\mathbf{x}}$  in (3.4) for the quadrature points lying on the edge  $e_K^1$  which are  $(0, \frac{1}{2} + v^\beta, \frac{1}{2} - v^\beta)$ , the points on  $e_K^2$  which are  $(\frac{1}{2} - v^\beta, 0, \frac{1}{2} + v^\beta)$ , and the points on  $e_K^3$  which are  $(\frac{1}{2} + v^\beta, \frac{1}{2} - v^\beta, 0)$ . Notice that  $\mathbf{g}_2(\frac{1}{2}, v^\beta)$  and  $\mathbf{g}_3(-\frac{1}{2}, -v^\beta)$  are the same point  $(0, \frac{1}{2} + v^\beta, \frac{1}{2} - v^\beta)$ . Since  $\widehat{w}_1 = \widehat{w}_N$ , the weight of  $(0, \frac{1}{2} + v^\beta, \frac{1}{2} - v^\beta)$  is

$$\frac{2}{3} \left( \frac{1}{2} + v^\beta \right) w_\beta \widehat{w}_1 + \frac{2}{3} \left( \frac{1}{2} - v^\beta \right) w_\beta \widehat{w}_N = \frac{2}{3} w_\beta \widehat{w}_1.$$

Similarly, we get that the weights for  $(\frac{1}{2} - v^\beta, 0, \frac{1}{2} + v^\beta)$  and  $(\frac{1}{2} + v^\beta, \frac{1}{2} - v^\beta, 0)$  are  $\frac{2}{3} w_\beta \widehat{w}_1$ .

The values of  $p_K(x, y)$  at the quadrature points on the edge  $e_K^i$  can be denoted as  $u_{i,\beta}^{int(K)}$ . There are  $L = 3(N-2)(k+1)$  quadrature points lying in the interior of  $K$ . We denote the values of  $p_K(x, y)$  at these points as  $u_\gamma^{int}$  ( $\gamma = 1, \dots, L$ ) with the coefficients  $\tilde{w}_\gamma$  in the decomposition. Then we can rewrite (3.4) as a convex combination of  $u_{i,\beta}^{int(K)}$  and  $u_\gamma^{int}$ :

$$\bar{u}_K^n = \sum_{i=1}^3 \sum_{\beta=1}^{k+1} \frac{2}{3} w_\beta \widehat{w}_1 u_{i,\beta}^{int(K)} + \sum_{\gamma=1}^L \tilde{w}_\gamma u_\gamma^{int}. \quad (3.5)$$

### 3.3 Decomposition of High Order Schemes

**Theorem 3.2** For the scheme (3.2) with the polynomial  $p_K(x, y)$  (either reconstruction or DG polynomial) of degree  $k$  to satisfy the maximum principle

$$m \leq \bar{u}_K^{n+1} \leq M,$$

a sufficient condition is that each  $p_K(x, y)$  satisfies  $p_K(x, y) \in [m, M], \forall (x, y) \in S_K^K$  where  $S_K^K$  is defined in (3.3), under the CFL condition

$$a \frac{\Delta t}{|K|} \sum_{i=1}^3 l_K^i \leq \frac{2}{3} \hat{w}_1. \quad (3.6)$$

Here  $\hat{w}_1$  is the quadrature weight of the  $N$ -point Gauss-Lobatto rule on  $[-\frac{1}{2}, \frac{1}{2}]$  for the first quadrature point. For  $k = 2, 3$ ,  $\hat{w}_1 = \frac{1}{6}$  and for  $k = 4, 5$ ,  $\hat{w}_1 = \frac{1}{12}$ .

*Proof* Rewrite the scheme (3.2) as

$$\begin{aligned} \bar{u}_K^{n+1} &= \bar{u}_K^n - \frac{\Delta t}{|K|} \sum_{i=1}^3 \sum_{\beta=1}^{k+1} h(u_{i,\beta}^{int(K)}, u_{i,\beta}^{ext(K)}, v^i) w_\beta l_K^i \\ &= \bar{u}_K^n - \frac{\Delta t}{|K|} \sum_{\beta=1}^{k+1} w_\beta \left( \sum_{i=1}^3 h(u_{i,\beta}^{int(K)}, u_{i,\beta}^{ext(K)}, v^i) l_K^i \right). \end{aligned} \quad (3.7)$$

Then decompose the flux term inside the bracket. Let

$$\begin{aligned} &\sum_{i=1}^3 h(u_{i,\beta}^{int(K)}, u_{i,\beta}^{ext(K)}, v^i) l_K^i \\ &= h(u_{1,\beta}^{int(K)}, u_{1,\beta}^{ext(K)}, v^1) l_K^1 + h(u_{2,\beta}^{int(K)}, u_{2,\beta}^{ext(K)}, v^2) l_K^2 + h(u_{3,\beta}^{int(K)}, u_{3,\beta}^{ext(K)}, v^3) l_K^3 \\ &= h(u_{1,\beta}^{int(K)}, u_{1,\beta}^{ext(K)}, v^1) l_K^1 + h(u_{1,\beta}^{int(K)}, u_{2,\beta}^{int(K)}, -v^1) l_K^1 \\ &\quad + h(u_{2,\beta}^{int(K)}, u_{1,\beta}^{int(K)}, v^1) l_K^1 + h(u_{2,\beta}^{int(K)}, u_{2,\beta}^{ext(K)}, v^2) l_K^2 + h(u_{2,\beta}^{int(K)}, u_{3,\beta}^{int(K)}, v^3) l_K^3 \\ &\quad + h(u_{3,\beta}^{int(K)}, u_{2,\beta}^{int(K)}, -v^3) l_K^3 + h(u_{3,\beta}^{int(K)}, u_{3,\beta}^{ext(K)}, v^3) l_K^3, \end{aligned} \quad (3.8)$$

where we have used the conservativity of the flux (3.1).

Plugging (3.5) and (3.8) into (3.7), we get the monotone form

$$\begin{aligned} \bar{u}_K^{n+1} &= \sum_{i=1}^3 \sum_{\beta=1}^{k+1} \frac{2}{3} w_\beta \hat{w}_1 u_{i,\beta}^{int(K)} + \sum_{\gamma=1}^L \tilde{w}_\gamma u_\gamma^{int} - \frac{\Delta t}{|K|} \sum_{\beta=1}^{k+1} w_\beta \left( \sum_{i=1}^3 h(u_{i,\beta}^{int(K)}, u_{i,\beta}^{ext(K)}, v^i) l_K^i \right) \\ &= \sum_{\gamma=1}^L \tilde{w}_\gamma u_\gamma^{int} + \sum_{\beta=1}^{k+1} \frac{2}{3} w_\beta \hat{w}_1 [H_{1,\beta} + H_{2,\beta} + H_{3,\beta}] \end{aligned}$$

where  $H_{1,\beta}$ ,  $H_{2,\beta}$ , and  $H_{3,\beta}$  are

$$H_{1,\beta} = u_{1,\beta}^{int(K)} - \frac{3\Delta t}{2\hat{w}_1|K|} \left[ h(u_{1,\beta}^{int(K)}, u_{1,\beta}^{ext(K)}, v^1) l_K^1 + h(u_{1,\beta}^{int(K)}, u_{2,\beta}^{int(K)}, -v^1) l_K^1 \right],$$

$$\begin{aligned}
H_{2,\beta} &= u_{2,\beta}^{int(K)} - \frac{3\Delta t}{2\widehat{w}_1|K|} \left[ h(u_{2,\beta}^{int(K)}, u_{1,\beta}^{int(K)}, v^1)l_K^1 + h(u_{2,\beta}^{int(K)}, u_{2,\beta}^{ext(K)}, v^2)l_K^2 \right. \\
&\quad \left. + h(u_{2,\beta}^{int(K)}, u_{3,\beta}^{int(K)}, v^3)l_K^3 \right], \\
H_{3,\beta} &= u_{3,\beta}^{int(K)} - \frac{3\Delta t}{2\widehat{w}_1|K|} \left[ h(u_{3,\beta}^{int(K)}, u_{2,\beta}^{int(K)}, -v^3)l_K^3 + h(u_{3,\beta}^{int(K)}, u_{3,\beta}^{ext(K)}, v^3)l_K^3 \right].
\end{aligned}$$

Under the CFL condition (3.6),  $H_{1,\beta}$  is a monotone increasing function of  $u_{1,\beta}^{int(K)}$ ,  $u_{1,\beta}^{ext(K)}$  and  $u_{2,\beta}^{int(K)}$ .  $H_{2,\beta}$  is a monotone increasing function of  $u_{2,\beta}^{int(K)}$ ,  $u_{1,\beta}^{int(K)}$ ,  $u_{2,\beta}^{ext(K)}$  and  $u_{3,\beta}^{int(K)}$ .  $H_{3,\beta}$  is a monotone increasing function of  $u_{3,\beta}^{int(K)}$ ,  $u_{2,\beta}^{int(K)}$  and  $u_{3,\beta}^{ext(K)}$ . In other words,  $H_{i,\beta}$  ( $i = 1, 2, 3$ ) are three formal monotone schemes.

Writing the right-hand side of the scheme (3.2) as a function  $H$  of  $u_{i,\beta}^{int(K)}$ ,  $u_{i,\beta}^{ext(K)}$  and  $u_\gamma^{int}$ , then

$$\bar{u}_K^{n+1} = H\left(u_{1,1}^{int(K)}, u_{1,2}^{int(K)}, \dots, u_{3,k+1}^{int(K)}, u_{1,1}^{ext(K)}, \dots, u_{3,k+1}^{ext(K)}, u_1^{int}, \dots, u_L^{int}\right),$$

and this function  $H$  is monotone increasing with respect to each argument.

Therefore, if all the point values involved here,  $u_{i,\beta}^{int(K)}$ ,  $u_{i,\beta}^{ext(K)}$  and  $u_\gamma^{int}$  are in the range  $[m, M]$ , which is equivalent to say that each  $p_K(x, y)$  satisfies  $p_K(x, y) \in [m, M]$ ,  $\forall(x, y) \in S_k^K$ , then we have the maximum principle:

$$m = H(m, m, \dots, m) \leq \bar{u}_K^{n+1} \leq H(M, M, \dots, M) = M. \quad \square$$

*Remark 1* Similar result holds for other monotone fluxes.

### 3.4 Limiter and Its Implementation for the DG Method

To enforce the sufficient condition in the previous theorem, we need to modify  $p_K(x, y)$  such that  $p_K(x, y) \in [m, M]$  for all  $(x, y) \in S_k^K$ . For all  $K$ , assume  $\bar{u}_K^n \in [m, M]$ , we use the following modified polynomial  $\tilde{p}_K(x, y)$  instead of  $p_K(x, y)$  by a linear scaling limiter:

$$\tilde{p}_K(x, y) = \theta(p_K(x, y) - \bar{u}_K^n) + \bar{u}_K^n, \quad \theta = \min \left\{ \left| \frac{M - \bar{u}_K^n}{M_K - \bar{u}_K^n} \right|, \left| \frac{m - \bar{u}_K^n}{m_K - \bar{u}_K^n} \right|, 1 \right\}, \quad (3.9)$$

with

$$M_K = \max_{(x,y) \in S_k^K} p_K(x, y), \quad m_K = \min_{(x,y) \in S_k^K} p_K(x, y). \quad (3.10)$$

Let  $\Delta x$  denote the mesh size, then  $\tilde{p}_K(x, y)$  defined in (3.9) satisfies the following properties:

- Accuracy: for smooth solutions,  $\tilde{p}_K(x, y) - p_K(x, y) = O(\Delta x^{k+1})$ ,  $\forall(x, y) \in K$ .
- Boundedness:  $\tilde{p}_K(x, y) \in [m, M]$ ,  $\forall(x, y) \in S_k^K$ .
- Conservativity:  $\frac{1}{|K|} \int_K \tilde{p}_K(x, y) dx dy = \bar{u}_K^n$ .

The boundedness and conservativity are straightforward. The accuracy can be proven by following the same line as in [22].

At time level  $n$ , assuming the DG polynomial on the triangle  $K$  is  $p_K(x, y)$  and the cell average of  $p_K(x, y)$  is  $\bar{u}_K^n \in [m, M]$ , the algorithm flowchart of our method for the Euler forward time discretization is

- Evaluate the point values of  $p_K(x, y)$  in (3.3) to get  $m_K$  and  $M_K$  in (3.10).
- Compute  $\tilde{p}_K(x, y)$  in (3.9).
- Replace  $p_K(x, y)$  by  $\tilde{p}_K(x, y)$  in the DG scheme with Euler forward in time under the CFL condition (3.6).

For SSP high order time discretizations, we need to use the limiter (3.9) and (3.10) in each stage for a Runge-Kutta method or in each step for a multistep method.

## 4 Application to Two Dimensional Incompressible Flows

### 4.1 Preliminaries

We are interested in solving the two dimensional incompressible Euler equations in the vorticity stream-function formulation:

$$\omega_t + (u\omega)_x + (v\omega)_y = 0, \quad (4.1)$$

$$\Delta\psi = \omega, \quad \langle u, v \rangle = \langle -\psi_y, \psi_x \rangle, \quad (4.2)$$

$$\omega(x, y, 0) = \omega_0(x, y), \quad \langle u, v \rangle \cdot \nu \text{ given on } \partial\Omega,$$

where  $\nu$  denotes the unit outward normal vector. The definition of  $\langle u, v \rangle$  in (4.2) gives us the divergence-free condition  $u_x + v_y = 0$ , which implies (4.1) is equivalent to the non-conservative form

$$\omega_t + u\omega_x + v\omega_y = 0. \quad (4.3)$$

The exact solution of (4.3) satisfies the maximum principle  $\omega(x, y, t) \in [m, M]$ , for all  $(x, y, t)$ , where  $m = \min_{x,y} \omega_0(x, y)$  and  $M = \max_{x,y} \omega_0(x, y)$ . For discontinuous solutions or solutions containing sharp gradient regions, it is preferable to solve the conservative form (4.1) rather than the nonconservative form (4.3). However, without the incompressibility condition  $u_x + v_y = 0$ , the conservative form (4.1) itself does not imply the maximum principle  $\omega(x, y, t) \in [m, M]$  for all  $(x, y, t)$ . This is the main difficulty to get a maximum-principle-satisfying scheme solving the conservative form (4.1) directly. In [22], we proved that the  $(k+1)$ -th order accurate (for any  $k$ ) DG scheme in [10] on a rectangular mesh with the linear scaling limiter under suitable CFL condition satisfied a global maximum principle. Here, we extend this result to schemes on a triangular mesh.

In [10], Liu and Shu introduced a high order discontinuous Galerkin method solving (4.1). We will first recall the method in [10] briefly. First, solve (4.2) by a standard Poisson solver for the stream-function  $\psi$  using continuous finite elements, then take  $u = -\psi_y$ ,  $v = \psi_x$ . Notice that on the boundary of each cell,  $\langle u, v \rangle \cdot \nu = \langle -\psi_y, \psi_x \rangle \cdot \nu = \frac{\partial\psi}{\partial\tau}$ , which is the tangential derivative. Thus  $\langle u, v \rangle \cdot \nu$  is continuous across the cell boundary since the tangential derivative of  $\psi$  along each edge is continuous. Therefore, the DG scheme for (4.1) can be defined as follows: start with a triangulation  $T_h$  of the domain  $\Omega$ , consisting of polygons of maximum size  $h$ , and the two approximation spaces

$$V_h^k = \{v : v|_K \in P^k(K), \forall K \in T_h\}, \quad W_{0,h}^k = V_h^k \cap C_0(\Omega),$$

where  $P^k(k)$  is the set of all polynomials of degree at most  $k$  on the cell  $K$ . For given  $\psi_h \in W_{0,h}^k$ , find the  $\omega_h \in V_h^k$  such that

$$\int_K \partial_t \omega_h v dx dy - \int_K \omega_h \mathbf{u}_h \cdot \nabla v dx dy + \sum_{e \in \partial K} \int_e \mathbf{u}_h \cdot \nu \widehat{\omega_h} v^- ds = 0, \quad \forall v \in V_h^k,$$

where

$$\mathbf{u}_h = \langle u_h, v_h \rangle = \left\langle -\frac{\partial \psi_h}{\partial y}, \frac{\partial \psi_h}{\partial x} \right\rangle.$$

Since the normal velocity  $\mathbf{u}_h \cdot \nu$  is continuous across any element boundary  $e$ , we can define the Lax-Friedrichs flux:

$$\mathbf{u}_h \cdot \nu \widehat{\omega}_h = h(\omega_h^-, \omega_h^+, \mathbf{u}_h \cdot \nu) = \frac{1}{2}[\mathbf{u}_h \cdot \nu(\omega_h^+ + \omega_h^-) - a(\omega_h^+ - \omega_h^-)], \quad (4.4)$$

where  $a$  is the maximum of  $|\mathbf{u}_h \cdot \nu|$  either locally or globally and  $h(\cdot, \cdot, \cdot)$  denotes the numerical flux. Upwind flux can also be used here.

#### 4.2 The Main Result

Consider a triangle  $K$  with the same notations as in the previous section. Assume the stream-function  $\psi$  is obtained with  $P^k$  elements, where  $P^k$  refers to the space of two-variable polynomials of degree  $k$ , and the DG method uses  $P^k$  elements. At time level  $n$ , in the triangle  $K$ , let  $\omega_i^{int(K)}(x, y)$  denote the traces of the DG polynomial  $\omega_K(x, y)$  on the edge  $e_K^i$ .  $\nu^i$  denotes the unit outer normal vector of the edge  $e_K^i$  and  $\mathbf{u} = \langle u, v \rangle$  denotes the velocity.

The cell average scheme with Euler forward in time of the DG method in [10] is

$$\overline{\omega}_K^{n+1}(x, y) = \overline{\omega}_K^n(x, y) - \frac{\Delta t}{|K|} \sum_{i=1}^3 \int_{e_K^i} h\left(\omega_i^{int(K)}(x, y), \omega_i^{ext(K)}(x, y), \mathbf{u} \cdot \nu^i\right) ds, \quad (4.5)$$

where  $h(\cdot, \cdot, \cdot)$  is defined in (4.4). The integrals in (4.5) are assumed to be computed exactly. Notice that the integrands are equivalent to single variable polynomials of degree  $2k - 1$ , therefore the integral is equal to the  $(k + 1)$ -point Gauss quadrature. Substituting the integrals by the  $(k + 1)$ -point Gauss quadrature in (4.5), we obtain the mathematically equivalent expression

$$\overline{\omega}_K^{n+1}(x, y) = \overline{\omega}_K^n(x, y) - \frac{\Delta t}{|K|} \sum_{i=1}^3 \sum_{\beta=1}^{k+1} h\left(\omega_{i,\beta}^{int(K)}, \omega_{i,\beta}^{ext(K)}, \mathbf{u}_\beta \cdot \nu^i\right) w_{\beta} l_K^i. \quad (4.6)$$

Now let us assume  $\overline{\omega}_K^n \in [m, M]$  and the DG polynomial  $\omega_K(x, y)$  is already processed by the limiter (3.9) and (3.10). In particular,  $\omega_K(x, y) \in [m, M], \forall (x, y) \in S_K^k$  with  $S_K^k$  defined in (3.3). Then we will show the scheme (4.6) satisfies the maximum principle  $\overline{\omega}_K^{n+1} \in [m, M]$ .

Plugging the decomposition of the cell average (3.5) into (4.6), we get the monotone form

$$\begin{aligned} \overline{\omega}_K^{n+1} &= \sum_{i=1}^3 \sum_{\beta=1}^{k+1} \frac{2}{3} w_\beta \widehat{w}_1 \omega_{i,\beta}^{int(K)} + \sum_{\gamma=1}^L \widetilde{w}_\gamma \omega_\gamma^{int} - \frac{\Delta t}{|K|} \sum_{\beta=1}^{k+1} w_\beta \\ &\quad \times \left( \sum_{i=1}^3 h(\omega_{i,\beta}^{int(K)}, \omega_{i,\beta}^{ext(K)}, \mathbf{u}_\beta \cdot \nu^i) l_K^i \right) \\ &= \sum_{\gamma=1}^L \widetilde{w}_\gamma \omega_\gamma^{int} + \sum_{\beta=1}^{k+1} \frac{2}{3} w_\beta \widehat{w}_1 [H_{1,\beta} + H_{2,\beta} + H_{3,\beta}] \end{aligned} \quad (4.7)$$

where  $H_{1,\beta}$ ,  $H_{2,\beta}$ , and  $H_{3,\beta}$  are

$$\begin{aligned} H_{1,\beta} &= \omega_{1,\beta}^{int(K)} - \frac{3\Delta t}{2\widehat{w}_1|K|} \left[ h(\omega_{1,\beta}^{int(K)}, \omega_{1,\beta}^{ext(K)}, \mathbf{u}_\beta \cdot \mathbf{v}^1) l_K^1 + h(\omega_{1,\beta}^{int(K)}, \omega_{2,\beta}^{int(K)}, -\mathbf{u}_\beta \cdot \mathbf{v}^1) l_K^1 \right], \\ H_{2,\beta} &= \omega_{2,\beta}^{int(K)} - \frac{3\Delta t}{2\widehat{w}_1|K|} \left[ h(\omega_{2,\beta}^{int(K)}, \omega_{1,\beta}^{int(K)}, \mathbf{u}_\beta \cdot \mathbf{v}^1) l_K^1 + h(\omega_{2,\beta}^{int(K)}, \omega_{2,\beta}^{ext(K)}, \mathbf{u}_\beta \cdot \mathbf{v}^2) l_K^2 \right. \\ &\quad \left. + h(\omega_{2,\beta}^{int(K)}, \omega_{3,\beta}^{int(K)}, \mathbf{u}_\beta \cdot \mathbf{v}^3) l_K^3 \right], \\ H_{3,\beta} &= \omega_{3,\beta}^{int(K)} - \frac{3\Delta t}{2\widehat{w}_1|K|} \left[ h(\omega_{3,\beta}^{int(K)}, \omega_{2,\beta}^{int(K)}, -\mathbf{u}_\beta \cdot \mathbf{v}^3) l_K^3 + h(\omega_{3,\beta}^{int(K)}, \omega_{3,\beta}^{ext(K)}, \mathbf{u}_\beta \cdot \mathbf{v}^3) l_K^3 \right]. \end{aligned}$$

Under the CFL condition

$$a \frac{\Delta t}{|K|} \sum_{i=1}^3 l_K^i \leq \frac{2}{3} \widehat{w}_1, \quad (4.8)$$

$H_{1,\beta}$  is a monotone increasing function of  $\omega_{1,\beta}^{int(K)}$ ,  $\omega_{1,\beta}^{ext(K)}$  and  $\omega_{2,\beta}^{int(K)}$ .  $H_{2,\beta}$  is a monotone increasing function of  $\omega_{2,\beta}^{int(K)}$ ,  $\omega_{1,\beta}^{int(K)}$ ,  $\omega_{2,\beta}^{ext(K)}$  and  $\omega_{3,\beta}^{int(K)}$ .  $H_{3,\beta}$  is a monotone increasing function of  $\omega_{3,\beta}^{int(K)}$ ,  $\omega_{2,\beta}^{int(K)}$  and  $\omega_{3,\beta}^{ext(K)}$ . We write the right-hand side of the scheme (4.7) as a function  $H$  of  $\omega_{i,\beta}^{int(K)}$ ,  $\omega_{i,\beta}^{ext(K)}$  and  $\omega_\gamma^{int}$  to obtain

$$\bar{\omega}_K^{n+1} = H \left( \omega_{1,1}^{int(K)}, \omega_{1,2}^{int(K)}, \dots, \omega_{3,k+1}^{int(K)}, \omega_{1,1}^{ext(K)}, \dots, \omega_{3,k+1}^{ext(K)}, \omega_1^{int}, \dots, \omega_L^{int} \right),$$

and this function  $H$  is monotone increasing with respect to each argument.

To prove the maximum principle, it suffices to show the consistency of  $H$ .

**Lemma 4.1** *The function  $H$  defined above is consistent:  $H(M, M, \dots, M) = M$ .*

*Proof* Let  $\omega_{i,\beta}^{int(K)}$ ,  $\omega_{i,\beta}^{ext(K)}$  and  $\omega_\gamma^{int}$  be equal to a constant  $M$ , then

$$\begin{aligned} H_{1,\beta} &= M - \frac{3\Delta t}{2\widehat{w}_1|K|} \left[ h(M, M, \mathbf{u}_\beta \cdot \mathbf{v}^1) l_K^1 + h(M, M, -\mathbf{u}_\beta \cdot \mathbf{v}^1) l_K^1 \right] \\ &= M - \frac{3\Delta t}{2\widehat{w}_1|K|} \left[ M \mathbf{u}_\beta \cdot \mathbf{v}^1 l_K^1 - M \mathbf{u}_\beta \cdot \mathbf{v}^1 l_K^1 \right] \\ &= M, \\ H_{2,\beta} &= M - \frac{3\Delta t}{2\widehat{w}_1|K|} \left[ h(M, M, \mathbf{u}_\beta \cdot \mathbf{v}^1) l_K^1 + h(M, M, \mathbf{u}_\beta \cdot \mathbf{v}^2) l_K^2 + h(M, M, \mathbf{u}_\beta \cdot \mathbf{v}^3) l_K^3 \right] \\ &= M - \frac{3\Delta t}{2\widehat{w}_1|K|} \left[ M \mathbf{u}_\beta \cdot \mathbf{v}^1 l_K^1 + M \mathbf{u}_\beta \cdot \mathbf{v}^2 l_K^2 + M \mathbf{u}_\beta \cdot \mathbf{v}^3 l_K^3 \right] \\ &= M - \frac{3\Delta t}{2\widehat{w}_1|K|} M \left[ \mathbf{u}_\beta \cdot \mathbf{v}^1 l_K^1 + \mathbf{u}_\beta \cdot \mathbf{v}^2 l_K^2 + \mathbf{u}_\beta \cdot \mathbf{v}^3 l_K^3 \right], \\ H_{3,\beta} &= M - \frac{3\Delta t}{2\widehat{w}_1|K|} \left[ h(M, M, -\mathbf{u}_\beta \cdot \mathbf{v}^3) l_K^3 + h(M, M, \mathbf{u}_\beta \cdot \mathbf{v}^3) l_K^3 \right] \\ &= M - \frac{3\Delta t}{2\widehat{w}_1|K|} \left[ -M \mathbf{u}_\beta \cdot \mathbf{v}^3 l_K^3 + M \mathbf{u}_\beta \cdot \mathbf{v}^3 l_K^3 \right] = M. \end{aligned}$$



Thus, by (4.7), we get

$$\begin{aligned}
 H(M, M, \dots, M) &= \sum_{\gamma=1}^L \tilde{w}_{\gamma} \omega_{\gamma}^{int} + \sum_{\beta=1}^{k+1} \frac{2}{3} w_{\beta} \widehat{w}_1 [H_{1,\beta} + H_{2,\beta} + H_{3,\beta}] \\
 &= \sum_{\gamma=1}^L \tilde{w}_{\gamma} M + \sum_{\beta=1}^{k+1} \frac{2}{3} w_{\beta} \widehat{w}_1 \left[ 3M - \frac{3\Delta t}{2\widehat{w}_1 |K|} M (\mathbf{u}_{\beta} \cdot \mathbf{v}^1 l_K^1 + \mathbf{u}_{\beta} \cdot \mathbf{v}^2 l_K^2 + \mathbf{u}_{\beta} \cdot \mathbf{v}^3 l_K^3) \right] \\
 &= \left( \sum_{\gamma=1}^L \tilde{w}_{\gamma} M + \sum_{\beta=1}^{k+1} 2w_{\beta} \widehat{w}_1 M \right) - M \frac{\Delta t}{|K|} \sum_{\beta=1}^{k+1} (\mathbf{u}_{\beta} \cdot \mathbf{v}^1 l_K^1 + \mathbf{u}_{\beta} \cdot \mathbf{v}^2 l_K^2 + \mathbf{u}_{\beta} \cdot \mathbf{v}^3 l_K^3) w_{\beta}.
 \end{aligned}$$

By the quadrature rule (3.5), we get  $(\sum_{\gamma=1}^L \tilde{w}_{\gamma} M + \sum_{\beta=1}^{k+1} 2w_{\beta} \widehat{w}_1 M) = M$ , therefore

$$\begin{aligned}
 H(M, M, \dots, M) &= M - M \frac{\Delta t}{|K|} \sum_{\beta=1}^{k+1} (\mathbf{u}_{\beta} \cdot \mathbf{v}^1 l_K^1 + \mathbf{u}_{\beta} \cdot \mathbf{v}^2 l_K^2 + \mathbf{u}_{\beta} \cdot \mathbf{v}^3 l_K^3) w_{\beta} \\
 &= M - M \frac{\Delta t}{|K|} \sum_{i=1}^3 \sum_{\beta=1}^{k+1} \mathbf{u}_{\beta} \cdot \mathbf{v}^i l_K^i w_{\beta}, \\
 (\text{Gauss quadrature}) &= M - M \frac{\Delta t}{|K|} \sum_{i=1}^3 \int_{e_K^i} \mathbf{u} \cdot \mathbf{v}^i ds, \\
 (\text{Divergence Theorem}) &= M - M \frac{\Delta t}{|K|} \iint_K (u_x + v_y) dA, \\
 (u_x + v_y = 0) &= M.
 \end{aligned}$$

□

Therefore, we have the maximum principle

$$m = H(m, m, \dots, m) \leq \overline{\omega}_K^{n+1} \leq H(M, M, \dots, M) = M.$$

*Remark 1* We can prove the same results for the upwind flux defined in [10] following the same lines.

*Remark 2* The result holds also for any passive convection linear equations with divergence-free velocity coefficients, namely (4.1) in which  $u$  and  $v$  are given functions satisfying  $u_x + v_y = 0$ , as long as the quadratures are exact for the integrands in the scheme. This can be easily achieved if we pre-process the divergence-free velocity field so that it is piecewise polynomial of the right degree for accuracy, continuous in the normal component across cell boundaries, and divergence-free.

## 5 High Order Schemes Preserving Positivity for Compressible Euler Equations on Triangular Meshes

### 5.1 Preliminaries and First Order Schemes

Consider two dimensional Euler equations

$$\mathbf{w}_t + \mathbf{f}(\mathbf{w})_x + \mathbf{g}(\mathbf{w})_y = 0, \quad t \geq 0, (x, y) \in \mathbb{R}^2, \quad (5.1)$$

$$\mathbf{w} = \begin{pmatrix} \rho \\ m \\ n \\ E \end{pmatrix}, \quad \mathbf{f}(\mathbf{w}) = \begin{pmatrix} m \\ \rho u^2 + p \\ \rho uv \\ (E + p)u \end{pmatrix}, \quad \mathbf{g}(\mathbf{w}) = \begin{pmatrix} n \\ \rho uv \\ \rho v^2 + p \\ (E + p)v \end{pmatrix} \quad (5.2)$$

with

$$m = \rho u, \quad n = \rho v, \quad E = \frac{1}{2} \rho u^2 + \frac{1}{2} \rho v^2 + \rho e, \quad p = (\gamma - 1) \rho e,$$

where  $\rho$  is the density,  $u$  is the velocity in  $x$  direction,  $v$  is the velocity in  $y$  direction,  $m$  and  $n$  are the momenta,  $E$  is the total energy,  $p$  is the pressure,  $e$  is the internal energy. The speed of sound is given by  $c = \sqrt{\gamma p / \rho}$ . The eigenvalues of the Jacobian  $\mathbf{f}'(\mathbf{w})$  are  $u - c$ ,  $u$ ,  $u + c$  and the eigenvalues of the Jacobian  $\mathbf{g}'(\mathbf{w})$  are  $v - c$ ,  $v$ ,  $v + c$ . The pressure function  $p$  is concave with respect to  $\mathbf{w}$  if  $\rho \geq 0$ , thus the set of admissible states

$$G = \left\{ \mathbf{w} = \begin{pmatrix} \rho \\ m \\ n \\ E \end{pmatrix} \left| \rho > 0 \text{ and } p = (\gamma - 1) \left( E - \frac{1}{2} \frac{m^2}{\rho} - \frac{1}{2} \frac{n^2}{\rho} \right) > 0 \right. \right\}$$

is convex.

We are interested in high order schemes that can be proven to keep the numerical solutions in the set  $G$ . First order positivity preserving schemes can be written as

$$\mathbf{w}_K^{n+1} = \mathbf{w}_K^n - \frac{\Delta t}{|K|} \sum_{i=1}^3 \mathbf{h}(\mathbf{w}_K^n, \mathbf{w}_{K(i)}^n, v^i) l_K^i, \quad (5.3)$$

where  $\mathbf{w}_K^n$  is the numerical approximation to the cell average in each element  $K$ , and  $\mathbf{h}(\cdot, \cdot, \cdot)$  is taken as the positivity preserving flux. In the rest of this section, we will use Lax-Friedrichs as an example:

$$\mathbf{h}(\mathbf{u}, \mathbf{v}, v) = \frac{1}{2} [\mathbf{F}(\mathbf{u}) \cdot v + \mathbf{F}(\mathbf{v}) \cdot v - a(\mathbf{v} - \mathbf{u})], \quad (5.4)$$

where  $\mathbf{F} = \langle \mathbf{f}, \mathbf{g} \rangle$ ,  $a = \max\{\|(|u| + c)\|_\infty, \|(|v| + c)\|_\infty\}$ . If  $\overline{\mathbf{w}}_K^n \in G$  for all  $K$ , then the scheme (5.3) and (5.4) satisfies  $\overline{\mathbf{w}}_K^{n+1} \in G$  under the CFL condition

$$a \frac{\Delta t}{|K|} \sum_{i=1}^3 l_K^i \leq 1. \quad (5.5)$$

This fact was proved in [14] for the more restrictive CFL number  $\frac{1}{2}$ , however the more relaxed CFL condition (5.5) can be proved along the same lines as in Remark 2.4 of [23].

A second order extension for positivity preserving schemes on triangular meshes was made in [14]. In the next subsection, we will use the techniques in Sect. 2 to construct arbitrarily high order schemes preserving positivity on triangular meshes.

## 5.2 High Order Schemes

Consider the RKDG scheme [3] solving (5.1). The scheme satisfied by the cell averages in the RKDG method with Euler forward time discretization is

$$\bar{\mathbf{w}}_K^{n+1} = \bar{\mathbf{w}}_K^n - \frac{\Delta t}{|K|} \sum_{i=1}^3 \int_{e_K^i} \mathbf{h}(\mathbf{w}_i^{int(K)}, \mathbf{w}_i^{ext(K)}, v^i) ds,$$

(this is also the form of a finite volume scheme). Here  $\bar{\mathbf{w}}_K^n$  is the cell average over  $K$  of the numerical solutions, and  $\mathbf{w}_i^{int(K)}, \mathbf{w}_i^{ext(K)}$  are the approximations to the values on the edge  $e_K^i$  obtained from the interior and the exterior of  $K$ . The flux  $\mathbf{h}$  is defined in (5.4). Assume the DG polynomial of degree  $k$  on the triangle  $K$  is  $\mathbf{q}_K(x, y) = (\rho_K(x, y), m_K(x, y), n_K(x, y), E_K(x, y))^T$  with the cell average  $\bar{\mathbf{w}}_K^n = (\bar{\rho}_K^n, \bar{m}_K^n, \bar{n}_K^n, \bar{E}_K^n)^T$ , then the edge integral should be approximated by the  $(k + 1)$ -point Gauss quadrature. The scheme becomes

$$\bar{\mathbf{w}}_K^{n+1} = \bar{\mathbf{w}}_K^n - \frac{\Delta t}{|K|} \sum_{i=1}^3 \sum_{\beta=1}^{k+1} \mathbf{h}(\mathbf{w}_{i,\beta}^{int(K)}, \mathbf{w}_{i,\beta}^{ext(K)}, v^i) w_{\beta} l_K^i. \quad (5.6)$$

**Theorem 5.1** *For the scheme (5.6) with the polynomial  $\mathbf{q}_K(x, y)$  (either reconstruction or DG polynomial) of degree  $k$  to satisfy the positivity property  $\bar{\mathbf{w}}_K^{n+1} \in G$ , a sufficient condition is that each  $\mathbf{q}_K(x, y)$  satisfies  $\mathbf{q}_K(x, y) \in G, \forall (x, y) \in S_k^K$  where  $S_k^K$  is defined in (3.3), under the CFL condition*

$$a \frac{\Delta t}{|K|} \sum_{i=1}^3 l_K^i \leq \frac{2}{3} \hat{w}_1. \quad (5.7)$$

Here  $\hat{w}_1$  is the quadrature weight of the  $N$ -point Gauss-Lobatto rule on  $[-\frac{1}{2}, \frac{1}{2}]$  for the first quadrature point. For  $k = 2, 3$ ,  $\hat{w}_1 = \frac{1}{6}$  and for  $k = 4, 5$ ,  $\hat{w}_1 = \frac{1}{12}$ .

*Proof* Rewrite the scheme (5.6) as

$$\begin{aligned} \bar{\mathbf{w}}_K^{n+1} &= \bar{\mathbf{w}}_K^n - \frac{\Delta t}{|K|} \sum_{i=1}^3 \sum_{\beta=1}^{k+1} \mathbf{h}(\mathbf{w}_{i,\beta}^{int(K)}, \mathbf{w}_{i,\beta}^{ext(K)}, v^i) w_{\beta} l_K^i \\ &= \bar{\mathbf{w}}_K^n - \frac{\Delta t}{|K|} \sum_{\beta=1}^{k+1} w_{\beta} \left( \sum_{i=1}^3 \mathbf{h}(\mathbf{w}_{i,\beta}^{int(K)}, \mathbf{w}_{i,\beta}^{ext(K)}, v^i) l_K^i \right). \end{aligned} \quad (5.8)$$

Then decompose the flux term inside the bracket. Let

$$\begin{aligned} &\sum_{i=1}^3 \mathbf{h}(\mathbf{w}_{i,\beta}^{int(K)}, \mathbf{w}_{i,\beta}^{ext(K)}, v^i) l_K^i \\ &= \mathbf{h}(\mathbf{w}_{1,\beta}^{int(K)}, \mathbf{w}_{1,\beta}^{ext(K)}, v^1) l_K^1 + \mathbf{h}(\mathbf{w}_{2,\beta}^{int(K)}, \mathbf{w}_{2,\beta}^{ext(K)}, v^2) l_K^2 + \mathbf{h}(\mathbf{w}_{3,\beta}^{int(K)}, \mathbf{w}_{3,\beta}^{ext(K)}, v^3) l_K^3 \end{aligned}$$

$$\begin{aligned}
&= \mathbf{h}(\mathbf{w}_{1,\beta}^{int(K)}, \mathbf{w}_{1,\beta}^{ext(K)}, v^1)l_K^1 + \mathbf{h}(\mathbf{w}_{1,\beta}^{int(K)}, \mathbf{w}_{2,\beta}^{int(K)}, -v^1)l_K^1 \\
&\quad + \mathbf{h}(\mathbf{w}_{2,\beta}^{int(K)}, \mathbf{w}_{1,\beta}^{int(K)}, v^1)l_K^1 + \mathbf{h}(\mathbf{w}_{2,\beta}^{int(K)}, \mathbf{w}_{2,\beta}^{ext(K)}, v^2)l_K^2 + \mathbf{h}(\mathbf{w}_{2,\beta}^{int(K)}, \mathbf{w}_{3,\beta}^{int(K)}, v^3)l_K^3 \\
&\quad + \mathbf{h}(\mathbf{w}_{3,\beta}^{int(K)}, \mathbf{w}_{2,\beta}^{int(K)}, -v^3)l_K^3 + \mathbf{h}(\mathbf{w}_{3,\beta}^{int(K)}, \mathbf{w}_{3,\beta}^{ext(K)}, v^3)l_K^3.
\end{aligned} \tag{5.9}$$

Plugging (3.5) and (5.9) into (5.8), we get

$$\begin{aligned}
\bar{\mathbf{w}}_K^{n+1} &= \sum_{i=1}^3 \sum_{\beta=1}^{k+1} \frac{2}{3} w_\beta \hat{w}_1 \mathbf{w}_{i,\beta}^{int(K)} + \sum_{\gamma=1}^L \tilde{w}_\gamma \mathbf{w}_\gamma^{int} - \frac{\Delta t}{|K|} \sum_{\beta=1}^{k+1} w_\beta \left( \sum_{i=1}^3 \mathbf{h}(\mathbf{w}_{i,\beta}^{int(K)}, \mathbf{w}_{i,\beta}^{ext(K)}, v^i)l_K^i \right) \\
&= \sum_{\gamma=1}^L \tilde{w}_\gamma \mathbf{w}_\gamma^{int} + \sum_{\beta=1}^{k+1} \frac{2}{3} w_\beta \hat{w}_1 [H_{1,\beta} + H_{2,\beta} + H_{3,\beta}]
\end{aligned}$$

where  $H_{1,\beta}$ ,  $H_{2,\beta}$ , and  $H_{3,\beta}$  are

$$\begin{aligned}
H_{1,\beta} &= \mathbf{w}_{1,\beta}^{int(K)} - \frac{3\Delta t}{2\hat{w}_1|K|} \left[ \mathbf{h}(\mathbf{w}_{1,\beta}^{int(K)}, \mathbf{w}_{1,\beta}^{ext(K)}, v^1)l_K^1 + \mathbf{h}(\mathbf{w}_{1,\beta}^{int(K)}, \mathbf{w}_{2,\beta}^{int(K)}, -v^1)l_K^1 \right], \\
H_{2,\beta} &= \mathbf{w}_{2,\beta}^{int(K)} - \frac{3\Delta t}{2\hat{w}_1|K|} \left[ \mathbf{h}(\mathbf{w}_{2,\beta}^{int(K)}, \mathbf{w}_{1,\beta}^{int(K)}, v^1)l_K^1 + \mathbf{h}(\mathbf{w}_{2,\beta}^{int(K)}, \mathbf{w}_{2,\beta}^{ext(K)}, v^2)l_K^2 \right. \\
&\quad \left. + \mathbf{h}(\mathbf{w}_{2,\beta}^{int(K)}, \mathbf{w}_{3,\beta}^{int(K)}, v^3)l_K^3 \right], \\
H_{3,\beta} &= \mathbf{w}_{3,\beta}^{int(K)} - \frac{3\Delta t}{2\hat{w}_1|K|} \left[ \mathbf{h}(\mathbf{w}_{3,\beta}^{int(K)}, \mathbf{w}_{2,\beta}^{int(K)}, -v^3)l_K^3 + \mathbf{h}(\mathbf{w}_{3,\beta}^{int(K)}, \mathbf{w}_{3,\beta}^{ext(K)}, v^3)l_K^3 \right].
\end{aligned}$$

Under the CFL condition (5.7),  $H_{2,\beta}$  is a formal first order positivity preserving scheme, namely, the same type as (5.3), therefore  $H_{2,\beta} \in G$ .  $H_{1,\beta}$  and  $H_{3,\beta}$  are formal one-dimensional first order positivity preserving schemes (see the appendix in [14] and Remark 2.4 of [23]), thus  $H_{1,\beta}, H_{3,\beta} \in G$ .

Notice that  $\bar{\mathbf{w}}_K^{n+1}$  is a convex combination of  $\mathbf{w}_\gamma^{int}$  and  $H_{i,\beta}$ . Since  $G$  is a convex set, we have  $\bar{\mathbf{w}}_K^{n+1} \in G$ .  $\square$

We can use SSP high order time discretization and it will keep the positivity preserving property because of the convexity.

### 5.3 Limiter and Its Implementation for the DG Method

Given the DG polynomials  $\mathbf{q}_K(x, y) = (\rho_K(x, y), m_K(x, y), n_K(x, y), E_K(x, y))^T$  with its cell average  $\bar{\mathbf{w}}_K = (\bar{\rho}_K^n, \bar{m}_K^n, \bar{n}_K^n, \bar{E}_K^n)^T \in G$ , we would like to modify  $\mathbf{q}_K(x, y)$  into  $\tilde{\mathbf{q}}_K(x, y)$  such that it satisfies:

- Accuracy: For smooth solutions, the limiter does not destroy accuracy

$$\|\tilde{\mathbf{q}}_K(x, y) - \mathbf{q}_K(x, y)\| = O(\Delta x^{k+1}), \quad \forall (x, y) \in K,$$

where  $\|\cdot\|$  denotes the Euclidean norm.

- Positivity:  $\tilde{\mathbf{q}}_K(x, y) \in G$  for any  $(x, y) \in S_K^K$ .

- Conservativity:

$$\frac{1}{\Delta x \Delta y} \iint_K \tilde{\mathbf{q}}_K(x, y) dx dy = \bar{\mathbf{w}}_K^n.$$

Define  $\bar{\rho}_K^n = (\gamma - 1)(\bar{E}_K^n - \frac{1}{2}(\bar{m}_K^n)^2/\bar{\rho}_K^n - \frac{1}{2}(\bar{n}_K^n)^2/\bar{\rho}_K^n)$ . Then  $\bar{\rho}_K^n > 0$  and  $\bar{p}_K^n > 0$  for all  $K$ . Assume there exists a small number  $0 < \varepsilon \ll 1$  such that  $\bar{\rho}_K^n \geq \varepsilon$  and  $\bar{p}_K^n \geq \varepsilon$  for all  $K$ .

The first step is to limit the density. Replace  $\rho_K(x, y)$  by

$$\hat{\rho}_K(x, y) = \theta_1(\rho_K(x, y) - \bar{\rho}_K^n) + \bar{\rho}_K^n, \quad (5.10)$$

$$\theta_1 = \min \left\{ \frac{\bar{\rho}_K^n - \varepsilon}{\bar{\rho}_K^n - \rho_{\min}}, 1 \right\}, \quad \rho_{\min} = \min_{(x,y) \in S_K^K} \rho_K(x, y). \quad (5.11)$$

The second step is to enforce the positivity of the pressure. Let

$$\hat{\mathbf{q}}_K(x, y) = (\hat{\rho}_K(x, y), m_K(x, y), n_K(x, y), E_K(x, y))^T.$$

For any fixed  $\mathbf{x} = (x, y) \in G$ , denote  $\hat{\mathbf{q}}_{\mathbf{x}} = \hat{\mathbf{q}}_K(x, y)$ . Define

$$\mathbf{s}_{\mathbf{x}}(t) = (1 - t)\bar{\mathbf{w}}_K^n + t\hat{\mathbf{q}}_{\mathbf{x}}.$$

Then calculate

$$t_{\mathbf{x}} = \begin{cases} 1, & \text{if } p(\hat{\mathbf{q}}_{\mathbf{x}}) \geq \varepsilon, \\ \text{the solution of } p(\mathbf{s}_{\mathbf{x}}(t)) = \varepsilon, & \text{if } p(\hat{\mathbf{q}}_{\mathbf{x}}) < \varepsilon. \end{cases} \quad (5.12)$$

We have the following new vector of polynomials

$$\tilde{\mathbf{q}}_K(x, y) = \theta_2(\hat{\mathbf{q}}_K(x, y) - \bar{\mathbf{w}}_K^n) + \bar{\mathbf{w}}_K^n, \quad (5.13)$$

$$\theta_2 = \min_{\mathbf{x} \in S_K^K} t_{\mathbf{x}}. \quad (5.14)$$

Following the same lines as in [23], we can show the limiting process (5.10), (5.11), (5.13) and (5.14) returns  $\tilde{\mathbf{q}}_K(x, y)$  satisfying the accuracy, positivity and conservativity.

Assume  $\bar{\rho}_K^n > 0$  and  $\bar{p}_K^n > 0$  for all  $K$ . With a fixed small number, for example, set  $\varepsilon = 10^{-13}$ , the algorithm flowchart of our algorithm for the Euler forward is

- In each cell, modify the density first:
  - If  $\bar{\rho}_K^n \geq \varepsilon$ , then evaluate  $\min_{(x,y) \in S_K^K} \rho_K(x, y)$  and get  $\hat{\rho}_K(x, y)$  by (5.10) and (5.11), set  $\hat{\mathbf{q}}_K(x, y) = (\hat{\rho}_K(x, y), m_K(x, y), n_K(x, y), E_K(x, y))^T$ .
  - If  $\bar{\rho}_K^n < \varepsilon$ , set  $\hat{\mathbf{q}}_K(x, y) = (\bar{\rho}_K^n, m_K(x, y), n_K(x, y), E_K(x, y))^T$ .
- Then modify the pressure:
  - If  $\bar{p}_K^n \geq \varepsilon$ , then solve  $t_{\mathbf{x}}$  in (5.12) and get  $\tilde{\mathbf{q}}_K(x, y)$  by (5.13) and (5.14).
  - If  $\bar{p}_K^n < \varepsilon$ , set  $\tilde{\mathbf{q}}_K(x, y) = \bar{\mathbf{w}}_K^n$ .
- Replace  $\mathbf{q}_K(x, y)$  by  $\tilde{\mathbf{q}}_K(x, y)$  for each  $K$  in the DG scheme with Euler forward in time under the CFL condition (5.7).

For SSP high order time discretizations, we need to use the limiter in each stage for a Runge-Kutta method or in each step for a multistep method.

**Table 1** Accuracy test for the linear equation  $u_t + u_x + u_y = 0$  at time  $t = 2$ , using the  $P^2$  polynomial DG scheme and the third order SSP Runge-Kutta method

Mesh size	Without limiter				With limiter			
	$L^\infty$ error	Order	$L^2$ error	Order	$L^\infty$ error	Order	$L^2$ error	Order
1/8	3.48E-02	–	3.39E-03	–	4.98E-02	–	8.07E-03	–
1/16	5.42E-03	2.68	4.14E-04	3.04	5.72E-03	3.12	5.84E-04	3.79
1/32	6.99E-04	2.95	5.17E-05	3.00	7.00E-04	3.03	5.57E-05	3.39
1/64	8.96E-05	2.96	6.47E-06	3.00	8.96E-05	2.96	6.56E-06	3.09
1/128	1.14E-05	2.97	8.08E-07	3.00	1.14E-05	2.97	8.12E-07	3.06

## 6 Numerical Tests

In this section, we implement the third order  $P^2$  DG method on triangular meshes with the limiter described in the previous sections. The numerical flux is global Lax-Friedrichs. The time discretization is third order SSP Runge-Kutta (2.14). All the meshes are unstructured, generated by EasyMesh [13].

### 6.1 Scalar Conservation Laws

We test the accuracy for the limiter on scalar conservation laws.

*Example 6.1* Linear equation  $u_t + u_x + u_y = 0$  with periodic boundary conditions.

The initial condition is  $u_0(x, y) = \sin(2\pi(x + y))$  in the domain  $[0, 1] \times [0, 1]$ . The analytic solution is  $u(x, y, t) = \sin(2\pi(x + y - 2t))$ . The time step is taken as that indicated by the CFL condition (3.6). The errors of the DG method with and without the limiter are listed in Table 1. We observe the designed order of accuracy for our method, justifying that the linear scaling limiter (3.9) and (3.10) does not destroy the accuracy for smooth solutions.

*Example 6.2* Burgers' equation  $u_t + (u^2)_x + (u^2)_y = 0$  with periodic boundary conditions.

The initial condition is  $u_0(x, y) = 0.5 \sin(2\pi(x + y))$  in the domain  $[0, 1] \times [0, 1]$ . The time step is taken as that indicated by the CFL condition (3.6). The errors of the DG method with and without the limiter are listed in Table 2, at time  $t = 0.2$ , when the solution is still smooth. For this nonlinear problem we can also see close to third order of accuracy. In particular, the errors with or without the limiter are comparable.

### 6.2 Two-Dimensional Incompressible Euler Equation

In this subsection, we test the scheme in Sect. 4 solving the two-dimensional incompressible Euler equation in vorticity stream-function formulation (4.1) and (4.2).

We use  $P^2$  polynomial continuous Galerkin finite element method to solve the Poisson's equation, coupled with the  $P^2$  DG scheme and the third order SSP Runge-Kutta method with the linear scaling limiter for the time-dependent equation.

*Example 6.3* Accuracy test, initial data is  $\omega_0(x, y) = -2 \sin(x) \sin(y)$  in the domain  $[0, 2\pi] \times [0, 2\pi]$  with periodic boundary conditions.

**Table 2** Accuracy test for the Burgers' equation  $u_t + (u^2)_x + (u^2)_y = 0$  at time  $t = 0.2$ , using the  $P^2$  polynomial DG scheme and the third order TVD Runge-Kutta method

Mesh size	Without limiter				With limiter			
	$L^\infty$ error	Order	$L^2$ error	Order	$L^\infty$ error	Order	$L^2$ error	Order
1/8	1.19E-02	–	1.25E-03	–	1.19E-02	–	1.32E-03	–
1/16	1.37E-03	2.87	1.64E-04	2.93	1.37E-03	3.12	1.66E-04	2.99
1/32	1.89E-04	2.85	2.50E-05	2.71	1.89E-04	2.85	2.51E-05	2.73
1/64	3.76E-05	2.33	4.09E-06	2.61	3.76E-05	2.33	4.09E-06	2.62
1/128	6.65E-06	2.50	5.97E-07	2.78	6.65E-06	2.50	5.97E-07	2.78
1/256	1.06E-06	2.65	8.67E-08	2.78	1.06E-06	2.65	8.67E-08	2.78

**Table 3** Accuracy test for the two-dimensional incompressible Euler equation, at time  $t = 2\pi$ , using the  $P^2$  DG scheme and the third order SSP Runge-Kutta method, with the limiter

Mesh size	$L^\infty$ error	Order	$L^2$ error	Order
$\pi/4$	3.11E-02	–	6.14E-03	–
$\pi/8$	4.07E-03	2.93	6.51E-04	3.24
$\pi/16$	6.36E-04	2.68	8.28E-05	2.98
$\pi/32$	7.23E-05	3.14	9.93E-06	3.06
$\pi/64$	8.96E-06	3.01	1.23E-06	3.01

The time step is taken as that indicated by the CFL condition (4.8). The exact solution is  $\omega(x, y, t) = -2 \sin(x) \sin(y)$ . We can see the designed third order of the accuracy for this problem at time  $t = 2\pi$  in Table 3.

**Example 6.4** The double shear layer problem in the domain  $[0, 2\pi] \times [0, 2\pi]$  with periodic boundary condition.

The initial condition is

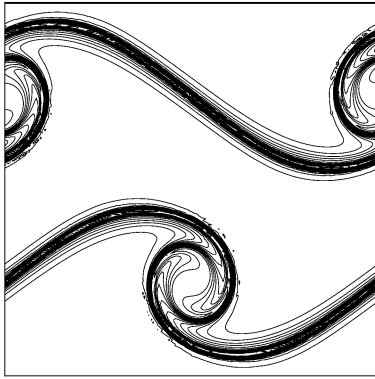
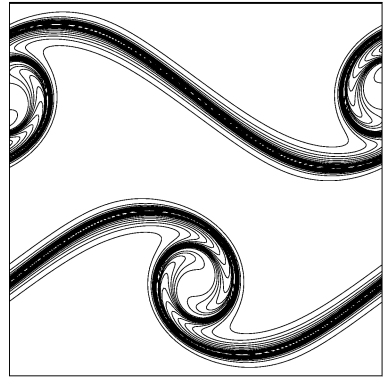
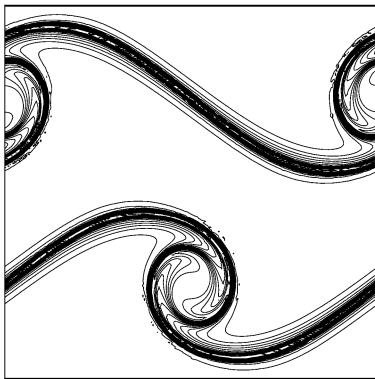
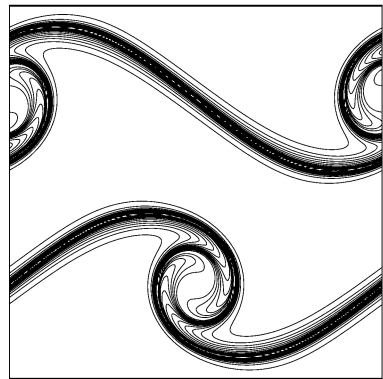
$$\omega(x, y, 0) = \begin{cases} \delta \cos(x) - \frac{1}{\rho} \operatorname{sech}^2((y - \frac{\pi}{2})/\rho), & y \leq \pi, \\ \delta \cos(x) - \frac{1}{\rho} \operatorname{sech}^2((y - \frac{\pi}{2})/\rho), & y > \pi, \end{cases}$$

with  $\delta = 0.05$  and  $\rho = \pi/15$ . In Figs. 3 and 4, we plot the contours of the vorticity  $\omega$  at time  $t = 6$  and  $t = 8$ , computed by the scheme with and without the limiter. Although one can barely see any difference between the results with the limiter and without the limiter from the contour, we point out that the numerical solutions of the scheme with the limiter are in the range  $[-\delta - \frac{1}{\rho}, \delta + \frac{1}{\rho}]$ .

**Example 6.5** The vortex patch problem in the domain  $[0, 2\pi] \times [0, 2\pi]$  with periodic boundary condition.

The initial condition is

$$\omega(x, y, 0) = \begin{cases} -1, & \frac{\pi}{2} \leq x \leq \frac{3\pi}{2}, \frac{\pi}{4} \leq y \leq \frac{3\pi}{4}; \\ 1, & \frac{\pi}{2} \leq x \leq \frac{3\pi}{2}, \frac{5\pi}{4} \leq y \leq \frac{7\pi}{4}; \\ 0, & \text{otherwise.} \end{cases}$$

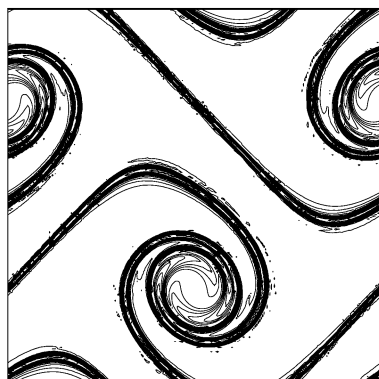
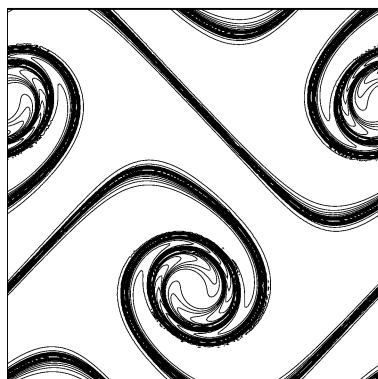
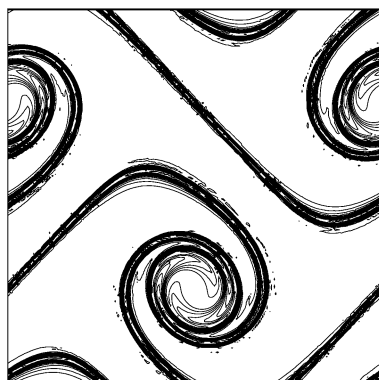
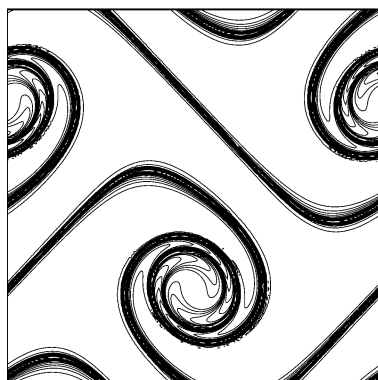
(a) Mesh size is  $\pi/32$ , without limiter(b) Mesh size is  $\pi/64$ , without limiter(c) Mesh size is  $\pi/32$ , with limiter(d) Mesh size is  $\pi/64$ , with limiter**Fig. 3** Vorticity at time  $t = 6$ ,  $P^2$ , 30 equally spaced contours from  $-4.9$  to  $4.9$ 

The contour plots of the vorticity  $\omega$  at time  $t = 5$  and  $t = 10$  are given in Figs. 5 and 6, computed by the schemes with and without the limiter. We also plot the cuts of the numerical results along the diagonal in these figures. We can not see any significant difference between the two results in the contour plots. But the cut plots show the advantage of the limiter, that the numerical results with the limiter is strictly in the range  $[-1, 1]$ .

### 6.3 Compressible Euler Equations

In this section, the numerical scheme solving (5.1) is the third order DG method with TVB limiter and the positivity preserving limiter. If discontinuities emerge in the solution, then we should use the characteristic TVB limiter [3] in the DG scheme. Although the positivity limiter can successfully preserve the positivity of density and pressure, the TVB limiter is still necessary for shocks. The DG scheme without the TVB limiter will produce blow-ups for the blast waves even if we use the positivity limiter.



(a) Mesh size is  $\pi/32$ , without limiter(b) Mesh size is  $\pi/64$ , without limiter(c) Mesh size is  $\pi/32$ , with limiter(d) Mesh size is  $\pi/64$ , with limiter**Fig. 4** Vorticity at time  $t = 8$ ,  $P^2$ , 30 equally spaced contours from  $-4.9$  to  $4.9$ 

In the TVB limiter, there is a TVB corrected minmod function defined by

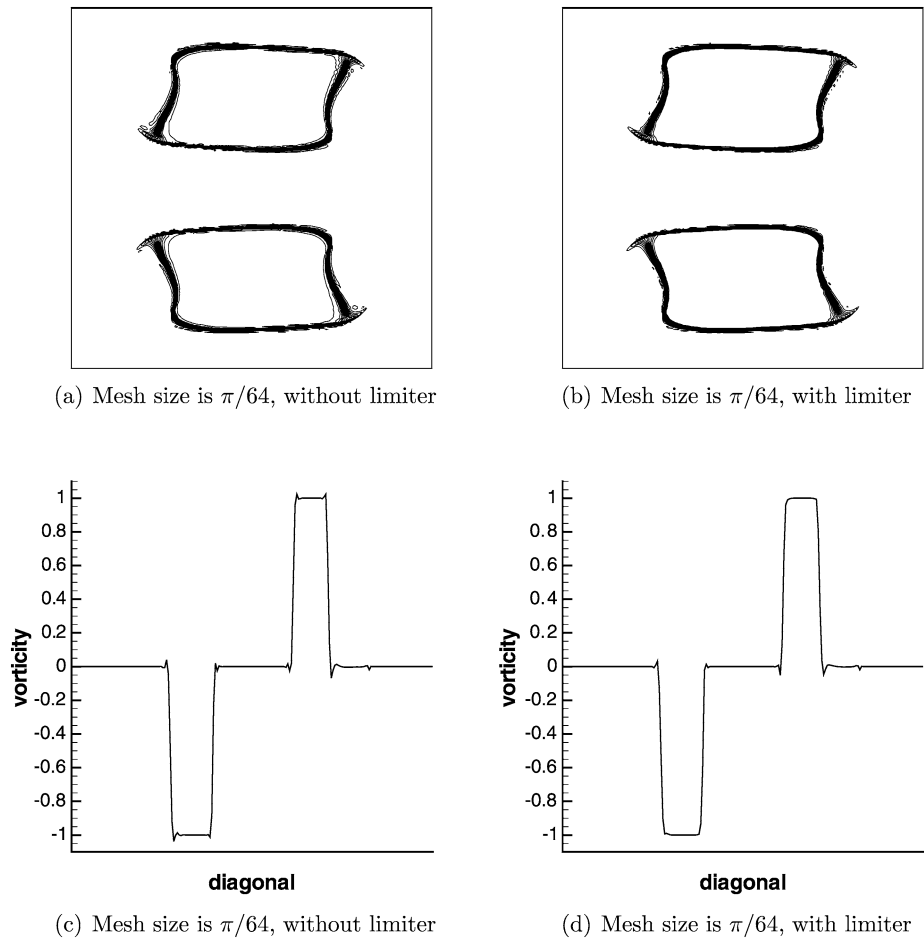
$$\overline{m}(a_1, \dots, a_m) = \begin{cases} a_1, & \text{if } |a_1| \leq M \Delta x^2, \\ m(a_1, \dots, a_m), & \text{otherwise,} \end{cases} \quad (6.1)$$

with the minmod function  $m$  defined by

$$m(a_1, \dots, a_m) = \begin{cases} s \min_i |a_i|, & \text{if } s = \text{sign}(a_1) = \dots = \text{sign}(a_m), \\ 0, & \text{otherwise.} \end{cases}$$

$M$  in (6.1) is a parameter. If  $M = 0$ , then the TVB limiter will reduce to the TVD limiter which is only first order accurate at smooth extrema. The TVB limiter will not destroy accuracy for large enough  $M$ , see [3] for more details of the TVB limiter.  $M$  will be taken as 50 in the following examples, unless otherwise specified.

**Example 6.6** Accuracy test, the domain is  $[0, 1] \times [0, 1]$  and the boundary condition is periodic.



**Fig. 5** Vorticity of the vortex patch problem, at time  $t = 5$ .  $P^2$ . *Top*: 30 equally spaced contours from  $-1.1$  to  $1.1$ . *Bottom*: cut along the diagonal

The initial condition is

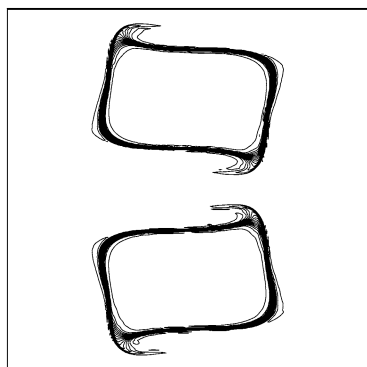
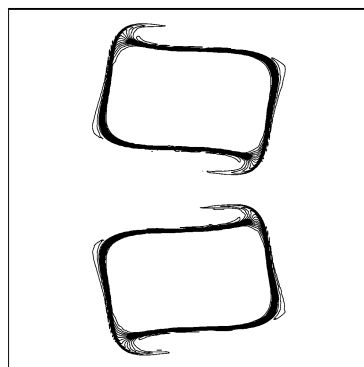
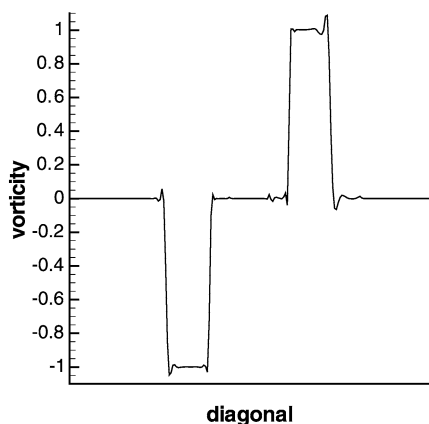
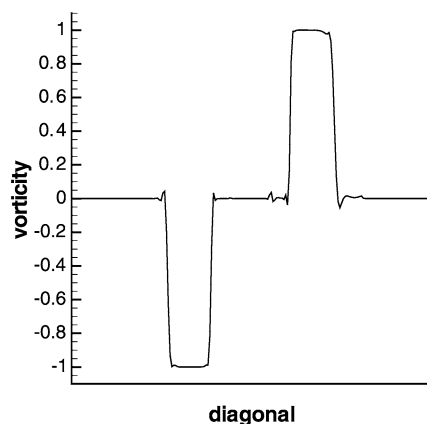
$$\rho_0(x, y) = 1 + 0.99 \sin(2\pi(x + y)), \quad u_0(x, y) = 1, \quad v_0(x, y) = 1, \quad p_0(x, y) = 1.$$

The exact solution is

$$\begin{aligned} \rho(x, y, t) &= 1 + 0.99 \sin(2\pi(x + y - 2t)), & u(x, y, t) &= 1, & v(x, y, t) &= 1, \\ p(x, y, t) &= 1. \end{aligned}$$

The minimum density of the exact solution is 0.01. We clearly observe the designed order of accuracy for this low density problem in Table 4.

*Example 6.7* The Sedov blast wave problem.

(a) Mesh size is  $\pi/64$ , without limiter(b) Mesh size is  $\pi/64$ , with limiter(c) Mesh size is  $\pi/64$ , without limiter(d) Mesh size is  $\pi/64$ , with limiter

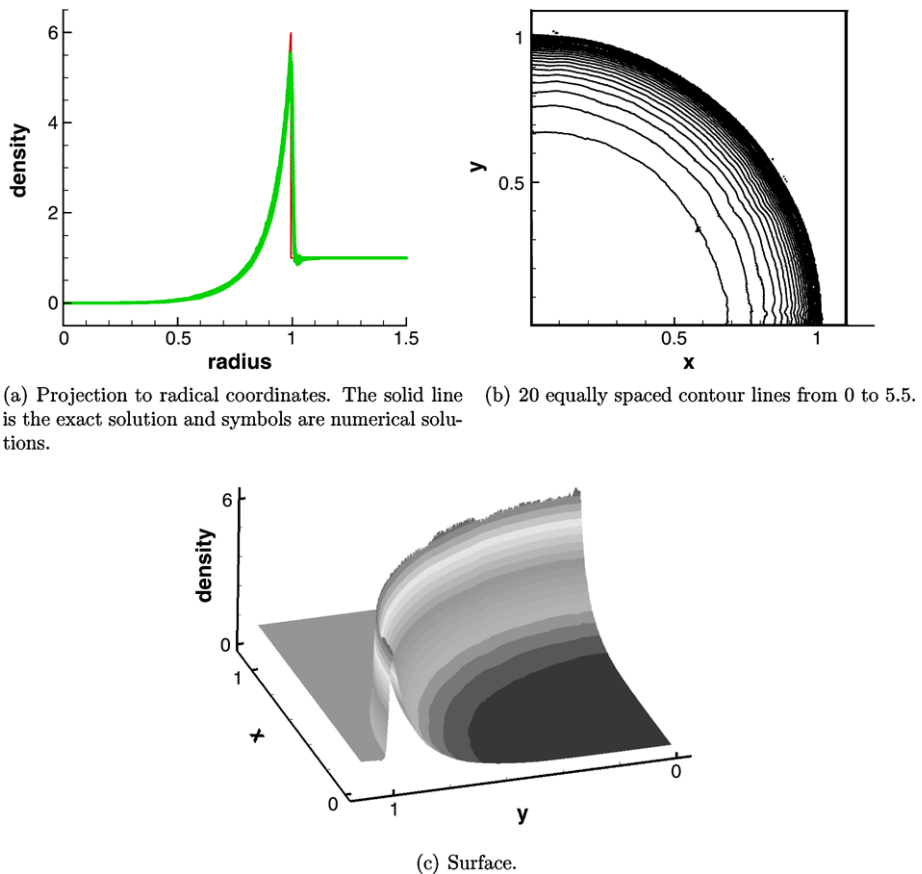
**Fig. 6** Vorticity of the vortex patch problem, at time  $t = 10$ .  $P^2$ . *Top*: 30 equally spaced contours from  $-1.1$  to  $1.1$ . *Bottom*: cut along the diagonal

**Table 4** The third order DG scheme for compressible Euler equation with initial condition  $1 + 0.99 \sin(2\pi(x + y))$  at  $t = 0.1$

Mesh size	$L^1$ error	Order	$L^\infty$ error	Order
1/10	3.66E-3	—	2.60E-2	—
1/20	5.42E-4	2.75	5.28E-3	2.30
1/40	8.48E-5	2.68	6.45E-4	3.03
1/80	1.12E-5	2.91	1.01E-4	2.68
1/160	1.37E-6	3.03	1.36E-5	2.89

The Sedov point-blast wave is a typical low density problem involving shocks. The exact solution formula can be found in [9, 17].

The computational domain is a square. For the initial condition, the density is 1, velocity is zero, total energy is  $10^{-12}$  everywhere except that the energy in the lower left corner cell is the constant  $\frac{0.244816}{|K|}$ ,  $\gamma = 1.4$ . The numerical boundary treatment is, extending  $\rho, v, E$



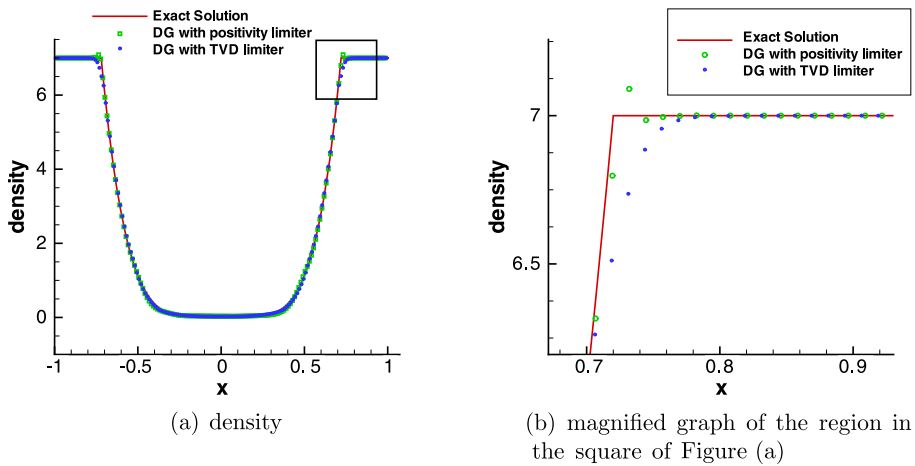
**Fig. 7** 2D Sedov blast, plot of density.  $T = 1$ .  $P^2$ . Mesh size is  $1/160$ . TVB limiter parameter  $M = 5000$

of the DG solutions as even functions and  $u$  as an odd function with respect to the left edge; extending  $\rho$ ,  $u$ ,  $E$  of the DG solutions as even functions and  $v$  as an odd function with respect to the bottom edge (symmetry). See Fig. 7. The computational result is comparable to those in the literature, e.g. [12] (which uses a Lagrangian method to compute this problem).

**Example 6.8** An extreme Riemann problem with vacuum emerging.

Computational domain is a square. The initial condition is  $\rho_L = \rho_R = 7$ ,  $u_L = -1$ ,  $u_R = 1$ ,  $v_L = v_R = 0$ ,  $p_L = p_R = 0.2$ . The exact solution contains vacuum. Since there is no shock, we do not need the TVB limiter for this problem. See Fig. 8 for the result of the DG scheme with only the positivity limiter. We can see that the low pressure and the low density are both captured very well. Without the positivity limiter, the DG scheme with the TVB limiter will blow up for this example. Even though the TVD limiter can make it stable, the result is more smeared when compared to the one with only the positivity limiter, see the comparison in Fig. 8.

**Example 6.9** A shock wave diffracts at a convex corner.



**Fig. 8** Double rarefaction problem.  $T = 0.6$ . Mesh size is  $1/80$ .  $P^2$ . Cut at  $y = 0$  for the 2D problem. The solid line is the exact solution. Symbols are numerical solutions. The left figure is the comparison between the result of the DG scheme with positivity limiter and the result of the DG scheme with TVD limiter. The right figure is the magnified graph of the region in the square of the left figure

A shock wave diffracting at a sharp convex corner is a benchmark problem in computational fluid dynamics. Many experimental and numerical results have been published, for example, [20, 21]. When the Mach number of the shock wave becomes larger, low density or pressure may appear. A Mach 5 shock diffracting at a ninety-degree convex edge will cause third order DG scheme blow up (see [3]). Our positivity preserving method on rectangular meshes in [23] gives satisfying results for this problem.

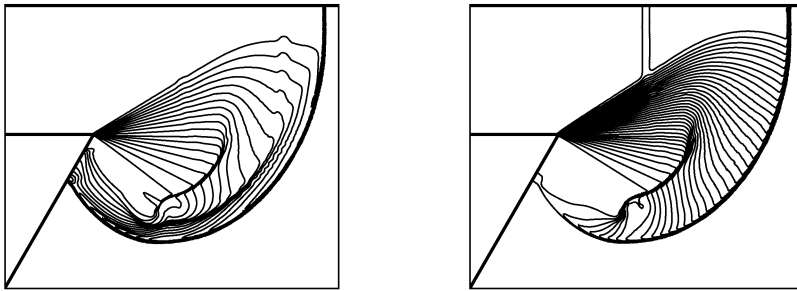
Here we would like to study a Mach 10 shock diffracting at a  $120^\circ$  convex corner, whose domain can not be discretized by a rectangular mesh any more. The contour plots of density and pressure at  $T = 0.9$  are given in Fig. 9, which look nice. See Fig. 10(a) for the illustration of the computational domain. The initial condition is a pure right-moving shock of  $Mach = 10$ , initially located at  $x = 3.4$  and  $6 \leq y \leq 11$ , moving into undisturbed air ahead of the shock with a density of 1.4 and pressure of 1. The high order DG method without the positivity limiter may blow up since the lowest density is very close to zero.

*Example 6.10* Schardin's Problem: shock waves pass a finite wedge.

The problem is named after the famous experiment by Schardin in 1957 [16], a planar shock passes a finite wedge and is reflected and diffracted. There are also many computational results, for instance, see [1, 15].

To validate our scheme, we first compute a Mach 1.34 shock passing an equilateral triangle, for which the experimental and computational results are available in [1]. See Fig. 10(b) for the computational domain. The initial condition is a pure right-moving shock of  $Mach = 1.34$ , initially located at  $x = 1$ , moving into undisturbed air ahead of the shock with a density of 1.4 and pressure of 1. The plot of density is shown in Fig. 11, which is comparable to the results in the literature.

Next, we compute a Mach 10 shock passing the same triangle. See Fig. 10(c) for the computational domain. The initial condition is a pure right-moving shock of  $Mach = 10$ , initially located at  $x = 0.2$ , moving into undisturbed air ahead of the shock with a density



(a) Density: 20 equally spaced contour lines from  $\rho = 0.07$  to  $\rho = 8.1$ . (b) Pressure: 40 equally spaced contour lines from  $p = 0.8$  to  $p = 115$ .

**Fig. 9** Mach 10 Shock diffracting at a  $120^\circ$  corner.  $T = 0.9$ .  $P^2$ . Mesh size is  $1/20$

of 1.4 and pressure of 1. For this problem, the density and pressure in the region behind the triangle will drop close to zero after the diffracting of the shock, which will cause blow-ups for many high order schemes. See Fig. 12 for our results, which look very good.

*Remark* All the examples in this subsection are computed in a parallel manner, implemented by MPI. The original RKDG method with the TVB limiter is fully local thus high parallel efficiency can be easily achieved. Our positivity limiter is also a purely local operation, so we still have ideal parallel efficiency for the RKDG with TVB and positivity limiters.

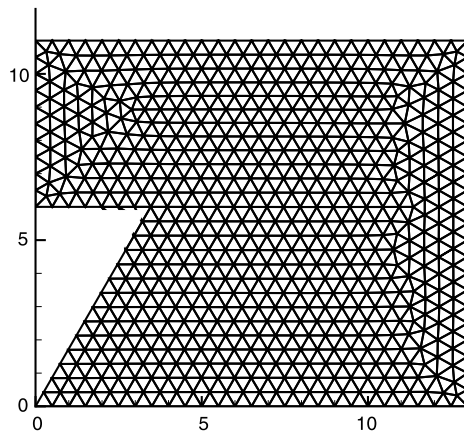
## 7 Concluding Remarks

In [22], it is the first time that genuinely high order schemes on rectangular meshes are obtained which satisfy strict maximum principle especially for multidimensional nonlinear problems. In this paper, we have extended the results in [22] to a general framework of constructing arbitrarily high order accurate maximum-principle-satisfying schemes for two-dimensional scalar conservation laws on triangular meshes. The main difficulty is how to decompose the cell average into a convex combination of the point values of the approximation polynomial, including all the Gauss quadrature points for each edge. We propose a special quadrature which gives us a very natural extension of the method in [22] to triangular meshes. The implementation is quite easy. With the addition of the limiter in this paper, which involves a small additional computational cost, to the DG scheme or the finite volume scheme (e.g. ENO and WENO), the numerical solutions will satisfy the maximum principle under suitable CFL condition.

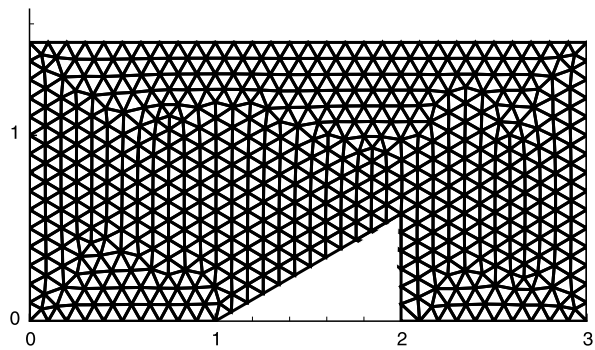
In addition, we show that the same result holds for the DG or finite volume schemes solving the two dimensional incompressible Euler equations in the vorticity stream-function formulation [10]. The result is also valid for any passive convection equation with divergence-free velocity coefficients.

The same techniques also allow us to extend the positivity preserving method for compressible Euler equation in [14, 23] to arbitrarily high order schemes on the triangular meshes. The positivity preserving limiter is easy to implement and purely local. Following the same idea on triangular meshes, it is also straightforward to extend all the results to three-dimensional schemes on tetrahedral meshes without difficulty. The RKDG method

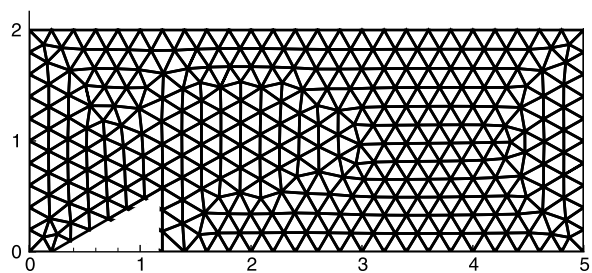
**Fig. 10** Illustration of the computational domain and the unstructured meshes



(a) Mach 10 Shock diffracts at a  $120^\circ$  corner.

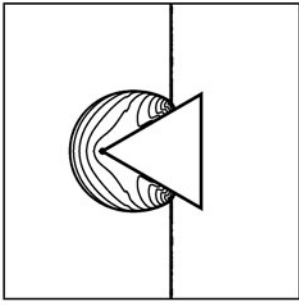


(b) Mach 1.34 shock passes an equilateral triangle.

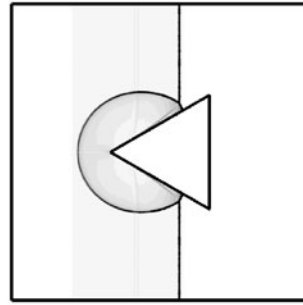


(c) Mach 10 shock passes an equilateral triangle.

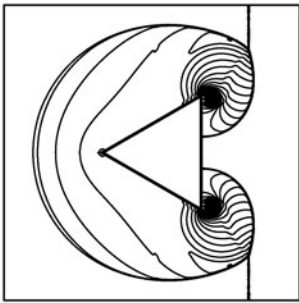
with TVB limiter and our positivity limiter is high order accurate, highly parallelizable and very robust for compressible Euler equations.



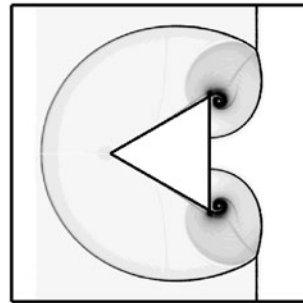
(a)  $T = 0.521$ , 30 equally spaced contour lines from  $\rho = 1.39$  to  $\rho = 2.88$



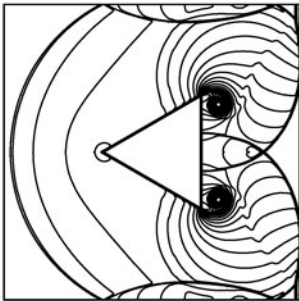
(b)  $T = 0.521$



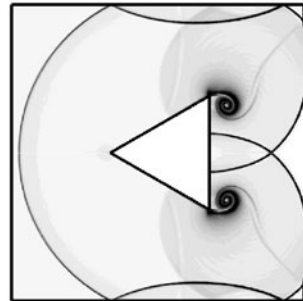
(c)  $T = 1.104$ , 30 equally spaced contour lines from  $\rho = 0.6$  to  $\rho = 2.9$



(d)  $T = 1.104$



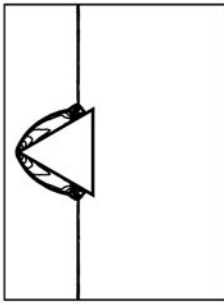
(e)  $T = 1.46$ , 30 equally spaced contour lines from  $\rho = 0.5$  to  $\rho = 2.9$



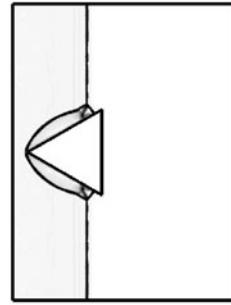
(f)  $T = 1.46$

**Fig. 11** Mach 1.34 shock passes an equilateral triangle.  $P^2$ . Mesh size is  $1/160$ . *Left*: the contour plot of the density. *Right*: the numerical Schlieren image of density

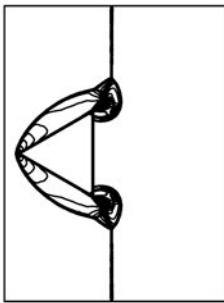




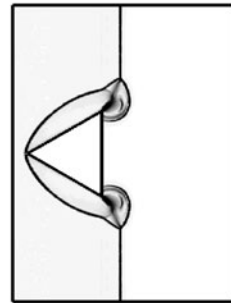
(a)  $T = 0.08$ , 30 equally spaced contour lines from  $\rho = 1.39$  to  $\rho = 21$



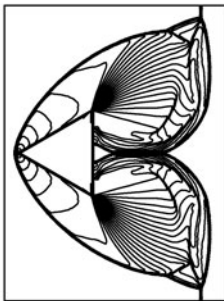
(b)  $T = 0.08$



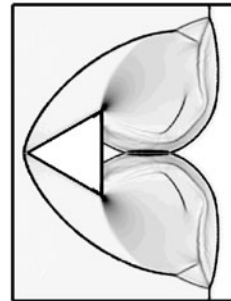
(c)  $T = 0.125$ , 30 equally spaced contour lines from  $\rho = 0.24$  to  $\rho = 22$



(d)  $T = 0.125$



(e)  $T = 0.245$ , 30 equally spaced contour lines from  $\rho = 0.078$  to  $\rho = 22$



(f)  $T = 0.245$

**Fig. 12** Mach 10 shock passes an equilateral triangle.  $P^2$ . Mesh size is  $1/80$ . *Left*: the contour plot of the density. *Right*: the numerical Schlieren image of density

## References

1. Chang, S., Chang, K.: On the shock-vortex interaction in Schardin's problem. *Shock Waves* **10**, 333–343 (2000)

2. Cockburn, B., Shu, C.-W.: TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws II: general framework. *Math. Comput.* **52**, 411–435 (1989)
3. Cockburn, B., Shu, C.-W.: The Runge-Kutta discontinuous Galerkin method for conservation laws V: multidimensional systems. *J. Comput. Phys.* **141**, 199–224 (1998)
4. Dafermos, C.M.: *Hyperbolic Conservation Laws in Continuum Physics*. Springer, Berlin (2000)
5. Gottlieb, S., Ketcheson, D.I., Shu, C.-W.: High order strong stability preserving time discretizations. *J. Sci. Comput.* **38**, 251–289 (2009)
6. Ha, Y., Gardner, C., Gelb, A., Shu, C.-W.: Numerical simulation of high Mach number astrophysical jets with radiative cooling. *J. Sci. Comput.* **24**, 597–612 (2005)
7. Harten, A., Engquist, B., Osher, S., Chakravarthy, S.: Uniformly high order essentially non-oscillatory schemes, III. *J. Comput. Phys.* **71**, 231–303 (1987)
8. Jiang, G.-S., Shu, C.-W.: Efficient implementation of weighted ENO schemes. *J. Comput. Phys.* **126**, 202–228 (1996)
9. Korobeinikov, V.P.: *Problems of Point-Blast Theory*. American Institute of Physics, New York (1991)
10. Liu, J.-G., Shu, C.-W.: A high-order discontinuous Galerkin method for 2D incompressible flows. *J. Comput. Phys.* **160**, 577–596 (2000)
11. Liu, X.-D., Osher, S., Chan, T.: Weighted essentially non-oscillatory schemes. *J. Comput. Phys.* **115**, 200–212 (1994)
12. Liu, W., Cheng, J., Shu, C.-W.: High order conservative Lagrangian schemes with Lax-Wendroff type time discretization for the compressible Euler equations. *J. Comput. Phys.* **228**, 8872–8891 (2009)
13. Niceno, B.: EasyMesh Version 1.4: a two-dimensional quality mesh generator (2001). Available from: <http://www-dinma.univ.trieste.it/nirftc/research/easymesh/>
14. Perthame, B., Shu, C.-W.: On positivity preserving finite volume schemes for Euler equations. *Numer. Math.* **73**, 119–130 (1996)
15. Sambasivan, S.K., UdayKumar, H.S.: Ghost fluid method for strong shock interactions part 2: immersed solid boundaries. *AIAA J.* **47**, 2923–2937 (2009)
16. Schardin, H.: High frequency cinematography in the shock tube. *J. Photosci.* **5**, 19–26 (1957)
17. Sedov, L.I.: *Similarity and Dimensional Methods in Mechanics*. Academic Press, New York (1959)
18. Shu, C.-W.: Total-variation-diminishing time discretizations. *SIAM J. Sci. Stat. Comput.* **9**, 1073–1084 (1988)
19. Shu, C.-W., Osher, S.: Efficient implementation of essentially non-oscillatory shock-capturing schemes. *J. Comput. Phys.* **77**, 439–471 (1988)
20. Skews, B.W.: The perturbed region behind a diffracting shock wave. *J. Fluid Mech.* **29**, 705–719 (1967)
21. Sun, M., Takayama, K.: The formation of a secondary shock wave behind a shock wave diffracting at a convex corner. *Shock Waves* **7**, 287–295 (1997)
22. Zhang, X., Shu, C.-W.: On maximum-principle-satisfying high order schemes for scalar conservation laws. *J. Comput. Phys.* **229**, 3091–3120 (2010)
23. Zhang, X., Shu, C.-W.: On positivity preserving high order discontinuous Galerkin schemes for compressible Euler equations on rectangular meshes. *J. Comput. Phys.* **229**, 8918–8934 (2010)