Oral presentation | Numerical methods

Numerical methods-I

Tue. Jul 16, 2024 2:00 PM - 4:00 PM Room A

[5-A-01] Positivity-preserving algorithm for implicit finite volume methods simulating compressible flows

*Qian Wang[1], Hanyu Zhou[2] (1. Beijing Computational Science Research Center, 2. Tsinghua University)

Keywords: Positivity-preserving, Implicit time stepping, Finite volume method, Compressible flow

Twelfth International Conference on
Computational Fluid Dynamics (ICCFD12),
Kobe, Japan, July 14-19, 2024

Positivity-preserving algorithm for implicit finite volume methods simulating compressible flows

Qian Wang*, Hanyu Zhou**
Corresponding author: qian.wang@csrc.ac.cn
* Beijing Computational Science Research Center, Beijing, 100193, China.
** Department of Engineering Mechanices, Tsinghua Universiy, Beijing, 100084, China.

Abstract: A commonly encountered issue in the numerical simulation of challenging compressible flow problems is the failure to preserve positivity of density or pressure. A positivity-preserving algorithm for finite volume methods using dual-time stepping is developed in this paper, towards robust numerical simulations of compressible flows. In the positivity-preserving algorithm, admissible cell-averages are obtained by controlling physical and pseudo time step sizes which results in limited solution changes. To overcome the difficulty of unknown solution changes in time step limiting, we propose to use explicit time discretizations to obtain efficient estimations of future states. The allowable time step sizes are determined by limiting the relative solution changes to ensure positivity of updated solutions. Given positive cell-averages, admissible reconstruction polynomials can be obtained by applying a positivity-preserving scaling limiter. The positivity-preserving algorithm is applied to a high-order variational finite volume method with an explicit first stage singly diagonally implicit Runge-Kutta (ESDIRK) time integration. Numerical results for a series of benchmark test cases demonstrate the robustness and high resolution of the positivity-preserving implicit high-order finite volume method.

Keywords: Positivity Preserving, Finite Volume, Implicit Time Stepping, Compressible Flows.

1 Introduction

Computational fluid dynamics (CFD) simulation has been an efficient tool for investigation of flow problems, due to the fast development of computer hardware as well as numerical methods, such as finite difference (FD), finite volume (FV), finite element (FE) and spectral methods. Second-order FV method on unstructured grids [1] is widely used in commercial and open-source CFD codes, thus being the workhorse of engineering flow simulations. Over recent decades, various high-order methods on unstructured grids have been developed, such as the high-order FV [2, 3, 4], discontinuous Galerkin (DG) [5, 6], hybrid FV/DG [7, 8, 9], residual distribution (RD) [10] and flux reconstruction (FR) [11, 12, 13]. Such methods, compared with their second-order counterparts, have higher-order accuracy and lower dissipation/dispersion, while maintaining the capability of handling complex geometries. However, CFD solvers suffers from the deterioration of robustness when applied to challenging flow problems, even on high-quality computational meshes. A commonly encountered robustness issue in compressible flow simulation is the failure to preserve positivity of the density or pressure, which leads to nonphysical solutions or numerical instabilities.

Tremendous efforts have been made on developing positivity-preserving algorithms to improve robustness of numerical methods, especially high-order numerical methods, for compressible flow simulations. For the numerical methods based on explicit time integration, popular positivity-preserving algorithms can be categorized into two types [14]. The first type is the positivity-preserving scaling limiter [15, 16], which compresses the solution polynomials to obtain positive density and pressure distributions, while being accuracy-preserving. It is proved that, by using the scaling limiter, a monotone numerical flux and a certain CFL condition, the FV and DG methods with strong stability preserving (SSP) Runge-Kutta time discretizations are positivity-preserving [16]. This approach has also been extended to FD methods [17, 18]. The second type is the positivity-preserving flux limiter [19, 20, 21], which seeks a convex combination of the first-order monotone flux with the high-order flux that can achieve the positivity-preserving property under certain conditions. Compared with the scaling limiter, the advantage of the flux limiter is the easy application to high-order FD methods solving conservation laws and high-order schemes solving convection-diffusion problems [14]. However, the accuracy-preserving property of the flux limiters are often demonstrated by extensive numerical results [22], as it is difficult to be proved for general cases.

Explicit time stepping methods suffer from the CFL constraint that restricts the time step sizes. For problems with very small allowable time step sizes, such as the high Reynolds number flows with large aspect ratio grids in the boundary layer region, implicit time stepping methods are much more efficient. Although most of the effort has been made for increasing accuracy of the time discretization and for increasing the efficiency of the nonlinear solver, only a few works exist in the literature concerning the positivity-preserving property of implicit methods [23], as it is difficult to verify if an implicit numerical scheme is positivity-preserving, even for a low-order one [24]. This difficulty comes from the fact that, in implicit time stepping, the solutions are updated by solving a system of nonlinear equations iteratively. Parent [25] developed a positivity-preserving dual-time stepping scheme for FV methods on structured grids solving Euler equations based on a Cauchy-Kowalevski procedure. Lian et al. [26] proposed a solution-limited time stepping scheme to enhance the reliability in steady-state compressible flow simulations. Qin et al. [23] developed an implicit high-order positivity-preserving DG method for steady-state compressible Euler equations. Recently, Huang et al. [24] proposed a positivity-preserving algorithm based on an iterative flux correction procedure, for finite volume methods simulating compressible flows on unstructured grids with a second-order backward difference (BDF2) dual-time stepping.

In this paper, we develop a positivity-preserving algorithm for finite volume methods on unstructured grids with an implicit dual-time stepping, towards robust and efficient simulations of unsteady compressible flows. In the positivity-preserving algorithm, admissible cell-averages are obtained by controlling physical and pseudo time step sizes which results in limited solution changes. To overcome the difficulty of unknown solution changes in time step limiting, we propose to use explicit time discretizations to obtain efficient estimations of future states. The allowable time step sizes are determined by limiting the relative solution changes to ensure positivity of updated solutions. Given positive cell-averages, admissible reconstruction polynomials can be obtained by applying a positivity-preserving scaling limiter [16]. The positivity-preserving algorithm is applied to a high-order variational finite volume method [4] with an explicit first stage singly diagonally implicit Runge-Kutta (ESDIRK) time integration [27], and verified using a series of benchmark test cases. Numerical results demonstrate the robustness and high resolution of the positivity-preserving implicit high-order finite volume method.

The remainder of this paper is organized as follows. Section 2 presents the high-order finite volume method on unstructured grids using implicit time stepping. Section 3 presents the positivity-preserving algorithm for implicit finite volume methods. Numerical results are given in Section 4 and concluding remarks are given in Section 5.

2 Implicit Finite Volume Method

2.1 Governing Equations

The Navier-Stokes equations governing unsteady compressible viscous flows can be expressed as

$$\frac{\partial \mathbf{U}}{\partial t} + \nabla \cdot (\mathbf{F} - \mathbf{F}_v) = 0, \tag{1}$$

where $\mathbf{U}$ is the conservative variable vector, $\mathbf{F}$ is the inviscid flux tensor and $\mathbf{F}_v$ is the viscous flux tensor defined by

$$\mathbf{U} = \begin{pmatrix} \rho \\ \rho\mathbf{u} \\ \rho E \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} \rho\mathbf{u} \\ \rho\mathbf{u} \otimes \mathbf{u} + p\mathbf{I} \\ (\rho E + p)\mathbf{u} \end{pmatrix}, \quad \mathbf{F}_v = \begin{pmatrix} 0 \\ \boldsymbol{\tau} \\ \boldsymbol{\tau} \cdot \mathbf{u} + \kappa \nabla T \end{pmatrix}. \tag{2}$$

Here ρ is the density, $\mathbf{u}$ is the velocity, p is the pressure and T is the temperature of the fluid. $\mathbf{I}$ is the identity tensor. E is the total energy defined as

$$E = \frac{1}{\gamma - 1}\frac{p}{\rho} + \frac{1}{2}\mathbf{u} \cdot \mathbf{u}, \tag{3}$$

where γ is the ratio of specific heat. The shear stress tensor $\boldsymbol{\tau}$ is defined as

$$\boldsymbol{\tau} = \mu \left[\nabla\mathbf{u} + (\nabla\mathbf{u})^{\mathrm{T}} - \frac{2}{3}(\nabla \cdot \mathbf{u})\mathbf{I} \right], \tag{4}$$

with μ being the dynamic viscosity. The heat conductivity is computed by $\kappa = C_p\mu/Pr$, where C_p is the specific heat at constant pressure and Pr is the Prandtl number. Equation (1) is closed by an equation

of state

$$p = \rho RT, \tag{5}$$

where $R = (\gamma - 1)\, C_p/\gamma$. Neglecting viscous effects, i.e., $\mu = 0$, the Navier-Stokes equations reduce to the Euler equations governing unsteady compressible inviscid flows.

2.2 High-order Finite Volume Spatial Discretization

This subsection presents the general framework of high-order cell-centered finite volume method on unstructured grids for compressible flows. The computational domain $\mathbf{\Omega}$ is partitioned into N non-overlapping control volumes, i.e., $\mathbf{\Omega} = \cup_{i=1}^{N} \mathbf{\Omega}_i$. By integrating the governing equation (1) over control volume $\mathbf{\Omega}_i$, a semi-discrete finite volume scheme is obtained as

$$\frac{\mathrm{d}\overline{\mathbf{U}}_i}{\mathrm{d}t} = -\frac{1}{\overline{\mathbf{\Omega}}_i} \oint_{\partial \mathbf{\Omega}_i} (\mathbf{F} - \mathbf{F}_v) \cdot \mathbf{n}\, \mathrm{d}A, \tag{6}$$

where $\overline{\mathbf{\Omega}}_i$ is the volume of $\mathbf{\Omega}_i$, $\partial \mathbf{\Omega}_i$ is the boundary of $\mathbf{\Omega}_i$ and $\mathbf{n}$ is the outward unit normal of $\partial \mathbf{\Omega}_i$, as shown in Figure 1. The cell-average defined by

$$\overline{\mathbf{U}}_i\,(t) = \frac{1}{\overline{\mathbf{\Omega}}_i} \int_{\mathbf{\Omega}_i} \mathbf{U}\,(\mathbf{x}, t)\ \mathrm{d}V, \tag{7}$$

is the degree of freedom (DOF) of the finite volume method on cell $\mathbf{\Omega}_i$. The flux integral in (6) is computed by using a Gauss quadrature

$$\oint_{\partial \mathbf{\Omega}_i} (\mathbf{F} - \mathbf{F}_v) \cdot \mathbf{n}\, \mathrm{d}A \approx \sum_{f \in \partial \mathbf{\Omega}_i} \sum_{g=1}^{N_g} w_g \left[\mathbf{F}\left(\mathbf{U}\left(\mathbf{x}_{f,g}, t\right)\right) - \mathbf{F}_v\left(\mathbf{U}\left(\mathbf{x}_{f,g}, t\right), \nabla \mathbf{U}\left(\mathbf{x}_{f,g}, t\right)\right)\right] \cdot \mathbf{n}_f\ A_f, \tag{8}$$

where $\mathbf{n}_f$ and A_f are the outward unit normal and area of element interface f, respectively. N_g is the number of quadrature points. $\mathbf{x}_{f,g}$ and w_g are the position and weight of the g-th quadrature point on f, respectively.

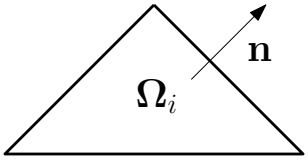


Figure 1: A triangular element.

Given known cell-averages, a reconstruction is performed to obtain an approximate solution distribution on the computational domain $\mathbf{\Omega}$, to compute the states at quadrature points on element interfaces. Specifically, the solution on each control volume is approximated by a polynomial, i.e.,

$$\mathbf{U}_i(\mathbf{x}, t) = \overline{\mathbf{U}}_i\,(t) + \sum_{l=1}^{\mathrm{N_b}(k)} \mathbf{U}_i^l\,(t)\, \varphi_{i,l}(\mathbf{x}), \quad \forall\ \mathbf{x} \in \mathbf{\Omega}_i, \quad i = 1, \cdots, N, \tag{9}$$

where k is the degree of the polynomial, $\{\varphi_{i,l}(\mathbf{x})\}$ are the polynomial basis functions and $\mathrm{N_b}(k)$ is the number of basis functions. The basis coefficients $\{\mathbf{U}_i^l\}$ are determined by using a reconstruction scheme, such as the variational reconstruction [4] that will be presented in Section 2.3.

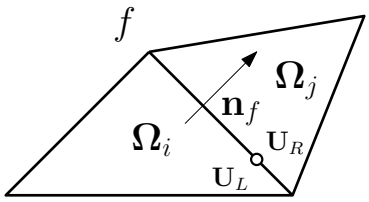


Figure 2: Riemann problem on cell interface.

Flux schemes based on solutions of Riemann problems are used to compute the numerical flux at the quadrature points in (8), since the piece-wise polynomial distribution is discontinuous across cell interfaces, as shown in Figure 2. The numerical flux is computed as

$$\left[\mathbf{F}\left(\mathbf{U}\left(\mathbf{x}_{f,g},t\right)\right)-\mathbf{F}_v\left(\mathbf{U}\left(\mathbf{x}_{f,g},t\right),\nabla\mathbf{U}\left(\mathbf{x}_{f,g},t\right)\right)\right]\cdot\mathbf{n}_f=\tilde{\mathbf{F}}\left(\mathbf{U}_L,\mathbf{U}_R,\mathbf{n}_f\right)-\tilde{\mathbf{F}}_v\left(\mathbf{U}_L,\nabla\mathbf{U}_L,\mathbf{U}_R,\nabla\mathbf{U}_R,\mathbf{n}_f\right), \tag{10}$$

where $\tilde{\mathbf{F}}$ and $\tilde{\mathbf{F}}_v$ are the schemes used to compute the inviscid and viscous numerical fluxes, respectively, with the left and right states defined by

$$\begin{cases} \mathbf{U}_L=\mathbf{U}_i\left(\mathbf{x}_{f,g},t\right), & \nabla\mathbf{U}_L=\nabla\mathbf{U}_i\left(\mathbf{x}_{f,g},t\right), \\ \mathbf{U}_R=\mathbf{U}_j\left(\mathbf{x}_{f,g},t\right), & \nabla\mathbf{U}_R=\nabla\mathbf{U}_j\left(\mathbf{x}_{f,g},t\right). \end{cases} \tag{11}$$

In this work, $\tilde{\mathbf{F}}$ is the local Lax-Friedrichs scheme and $\tilde{\mathbf{F}}_v$ is the viscous flux scheme by Wang et al. [4].

The semi-discrete finite volume scheme (6) can be rewritten in an ordinary differential equation (ODE) form

$$\frac{\mathrm{d}\overline{\mathbf{U}}_i}{\mathrm{d}t}=\mathbf{R}_i\left(t,\left\{\mathbf{U}_j\right\}_{j\in S_i}\right), \tag{12}$$

where can be integrated in time to update the cell-average in a step-by-step manner. The time integration method used in this work will be presented in Section 2.4.

2.3 Variational Reconstruction

This subsection presents a variational reconstruction scheme [4] that can achieve arbitrary high-order accuracy on a compact stencil involving only the current cell and its face-neighboring cells. The compact reconstruction stencil $S_i=\{i,j_1,j_2,j_3\}$ of a triangular element i is shown in Figure 3. In general, the conservative variables are reconstructed separately, allowing an implementation of solution reconstruction in a variable-by-variable manner. For the sake of presentation, the two-dimensional case is considered to illustrate the variational reconstruction procedure.

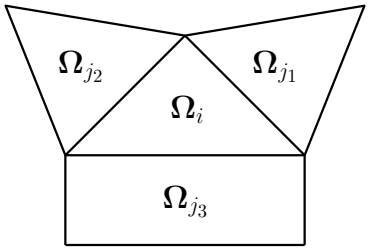


Figure 3: Compact reconstruction stencil.

On control volume $\boldsymbol{\Omega}_i$, a conservative variable $u\in\mathbf{U}$ is approximated as

$$u_i\left(\mathbf{x}\right)=\overline{u}_i+\sum_{l=1}^{\mathrm{N_b}(k)}u_i^l\varphi_{i,l}\left(\mathbf{x}\right), \tag{13}$$

where $\{\varphi_{i,l}\}$ are the basis functions, $\{u_i^l\}$ are the unknown basis coefficients and $\mathrm{N_b}(k)$ is the number of basis functions. For a two-dimensional case, $\mathrm{N_b}(k)=(k+2)(k+1)/2-1$. In (13), the time t is omitted as it remains a constant during the reconstruction procedure. In this work, the basis functions with the zero-mean property

$$\frac{1}{\overline{\boldsymbol{\Omega}}_i}\int_{\boldsymbol{\Omega}_i}\varphi_{i,l}\left(\mathbf{x}\right)\ \mathrm{d}V=0, \tag{14}$$

are used to make the reconstruction polynomial (13) automatically satisfy the conservation condition

$$\frac{1}{\overline{\boldsymbol{\Omega}}_i}\int_{\boldsymbol{\Omega}_i}u_i\left(\mathbf{x}\right)\ \mathrm{d}V=\overline{u}_i. \tag{15}$$

The zero-mean basis functions are defined by

$$\varphi_{i,l} = \left(\frac{x-x_i}{\Delta x_i}\right)^{p_l}\left(\frac{y-y_i}{\Delta y_i}\right)^{q_l} - \overline{\left(\frac{x-x_i}{\Delta x_i}\right)^{p_l}\left(\frac{y-y_i}{\Delta y_i}\right)^{q_l}}, \tag{16}$$

with

$$\overline{\left(\frac{x-x_i}{\Delta x_i}\right)^{p_l}\left(\frac{y-y_i}{\Delta y_i}\right)^{q_l}} = \frac{1}{\overline{\mathbf{\Omega}}_i}\int_{\mathbf{\Omega}_i}\left(\frac{x-x_i}{\Delta x_i}\right)^{p_l}\left(\frac{y-y_i}{\Delta y_i}\right)^{q_l}\mathrm{d}V, \tag{17}$$

where (x_i, y_i) and $(\Delta x_i, \Delta y_i)$ are the barycenter and characteristic length scales of cell i, respectively, and (p_l, q_l) are the powers of the basis functions organized in ascending order of $p_l + q_l$. For instance, the powers of the cubic ($k = 3$) reconstruction polynomial are

$$\{(p_l, q_l)\}_{l=1}^{9} = \{(1,0),(0,1),(2,0),(1,1),(0,2),(3,0),(2,1),(1,2),(0,3)\}. \tag{18}$$

The characteristic length scales are used to non-dimensionalize the basis functions, to avoid a growth of the condition number of the reconstruction matrix during grid refinement [28, 29]. In this work, the length scales are set as $\Delta x_i = \Delta y_i = \max_{\mathbf{x}\in\mathbf{\Omega}_i}\{\|\mathbf{x}-\mathbf{x}_i\|_2\}$.

The objective of a reconstruction is to determine the unknown basis coefficients u_i^l, $l = 1, \cdots, \mathrm{N_b}(k)$, $i = 1, \cdots, N$, given the cell-averages $\overline{u}_j$, $j = 1, \cdots, N$. In the variational reconstruction [4], the linear equation system to determine the unknown basis coefficients is derived by minimizing a cost function using the variational method. Different cost function results in different reconstruction schemes. The cost function is defined as

$$I = \sum_{f=1}^{N_f} I_f, \tag{19}$$

where I_f is an interfacial jump integration (IJI) on cell interface f and N_f is the total number of cell interfaces on the computational domain. In this work, the IJI is defined as

$$I_f = \omega_f^G \sum_{p+q=0}^{k}\int_f\left[\omega_f^D(p,q)\left(\frac{\partial^{p+q}u_L}{\partial x^p\partial y^q} - \frac{\partial^{p+q}u_R}{\partial x^p\partial y^q}\right)\right]^2\mathrm{d}A \tag{20}$$

where L and R are the two cells sharing interface f, ω_f^G is the geometric weight and ω_f^D is the derivative weight. The IJI measures the jumps of the reconstruction polynomial and its spatial derivatives on cell interface. The derivative weights of the cubic variational reconstruction are

$$\begin{aligned}
&\omega_f^D(0,0) = \omega_D(0),\\
&\omega_f^D(1,0) = \omega_f^D(0,1) = \omega_D(1),\\
&\omega_f^D(2,0) = \omega_f^D(0,2) = \omega_D(2),\ \omega_f^D(1,1) = \sqrt{2}\,\omega_D(2),\\
&\omega_f^D(3,0) = \omega_f^D(0,3) = \omega_D(3),\ \omega_f^D(1,2) = \omega_f^D(2,1) = \sqrt{3}\,\omega_D(3),
\end{aligned} \tag{21}$$

with

$$\omega_D(0) = 1,\ \omega_D(1) = d_{LR},\ \omega_D(2) = \frac{(d_{LR})^2}{2},\ \omega_D(3) = \frac{(d_{LR})^3}{6}, \tag{22}$$

where $d_{LR} = \|\mathbf{x}_L - \mathbf{x}_R\|_2$ is the distance between the barycenters of cell L and R. As only nearly uniform and isotropic meshes are used in the numerical simulations in this paper, the geometric weights ω_f^G are all set as 1. The unknown basis coefficients are determined by minimizing the cost function (19), resulting in a "smoothest" piece-wise polynomial distribution which has the smallest jumps measured by the IJIs on cell interfaces. Linear reconstruction equations can be derived by using the variational method, i.e.,

$$\frac{\partial I}{\partial u_i^l} = 0,\ l = 1, \cdots, \mathrm{N_b}(k),\ i = 1, \cdots N. \tag{23}$$

By substituting (13) into (20), (19) and (23), a linear equation system is obtained as following

$$\mathbf{A}_i\mathbf{u}_i = \sum_{j\in S_i, j\neq i}\mathbf{B}_i^j\mathbf{u}_j + \mathbf{b}_i, \quad i = 1, \cdots, N, \tag{24}$$

where the elements of matrices $\mathbf{A}_i$, $\mathbf{B}_i^j \in \mathbb{R}^{\mathrm{N_b}(k)\times \mathrm{N_b}(k)}$ and vectors $\mathbf{u}_i$, $\mathbf{u}_j$, $\mathbf{b}_i \in \mathbb{R}^{\mathrm{N_b}(k)}$ are

$$\begin{aligned}
\mathbf{A}_i[m,n] &= \sum_{j\in S_i, j\neq i} \omega_f^G \int_{\mathbf{\Omega}_i\cap\mathbf{\Omega}_j} \sum_{p+q=0}^{k} \left(\omega_f^D(p,q)\right)^2 \frac{\partial^{p+q}\varphi_{i,m}}{\partial x^p \partial y^q} \frac{\partial^{p+q}\varphi_{i,n}}{\partial x^p \partial y^q}\ \mathrm{d}A, \\
\mathbf{B}_i^j[m,n] &= \omega_f^G \int_{\mathbf{\Omega}_i\cap\mathbf{\Omega}_j} \sum_{p+q=0}^{k} \left(\omega_f^D(p,q)\right)^2 \frac{\partial^{p+q}\varphi_{i,m}}{\partial x^p \partial y^q} \frac{\partial^{p+q}\varphi_{j,n}}{\partial x^p \partial y^q}\ \mathrm{d}A, \\
\mathbf{b}_i[m] &= \sum_{j\in S_i, j\neq i} \omega_f^G \left(\omega_f^D(0,0)\right)^2 \int_{\mathbf{\Omega}_i\cap\mathbf{\Omega}_j} \varphi_{i,m}\left(\overline{u}_j - \overline{u}_i\right)\ \mathrm{d}A, \\
\mathbf{u}_i[m] &= u_i^m, \\
\mathbf{u}_j[m] &= u_j^m.
\end{aligned} \tag{25}$$

It is observed from (24) that the variational reconstruction is implicit, as the unknown coefficients of the face-neighboring cells are required to determine the unknown coefficients of the current cell. By assembling the linear equations (24) of all cells, we obtain a global linear equation system

$$\mathbf{A}\mathbf{u} = \mathbf{b}, \tag{26}$$

where

$$\begin{aligned}
&\mathbf{A} = \mathbf{D} - \mathbf{L} - \mathbf{U},\ \mathbf{D} = \{\mathbf{A}_i\},\ \mathbf{L} = \left\{\mathbf{B}_i^j, j < i\right\},\ \mathbf{U} = \left\{\mathbf{B}_i^j, j > i\right\}, \\
&\mathbf{u} = \{\mathbf{u}_i\},\ \mathbf{b} = \{\mathbf{b}_i\}.
\end{aligned} \tag{27}$$

It is proved in [4] that the large and sparse reconstruction matrix $\mathbf{A}$ is symmetric and positive definite, which guarantees the existence and uniqueness of the solution of the linear equation system (26). This is a significant advantage of the variational reconstruction over other existing high-order reconstructions on unstructured grids. The linear equation system (26) is solved iteratively using the block Gauss-Seidel method, of which the convergence is proved in [4]. Each block Gauss-Seidel iteration is compact as it only relies on the information of the current and face-neighboring cells.

In the reconstructions of solutions involving discontinuities, a WBAP limiter [30, 31] is used to suppress spurious oscillations. The limiting process on the element i in Figure 3 is presented for illustration. The limited reconstruction polynomial

$$\tilde{u}_i(\mathbf{x}) = \overline{u}_i + \sum_{l=1}^{\mathrm{N_b}(k)} \tilde{u}_i^l \varphi_{i,l}(\mathbf{x}) \tag{28}$$

is a nonlinear weighted average of the reconstruction polynomial $u_i(\mathbf{x})$ and three additional candidate polynomials $u_{j\to i}(\mathbf{x}), j = j_1, j_2, j_3$ obtained through a secondary reconstruction [32]. The limited coefficients $\tilde{u}_i^l$ is computed by

$$\tilde{u}_i^l = L\left(u_i^l, u_{j_1\to i}^l, u_{j_2\to i}^l, u_{j_3\to i}^l\right), \tag{29}$$

with a limiting function

$$L(a_0, a_1, ..., a_J) = a_0 \cdot W\left(1, \frac{a_1}{a_0}, ..., \frac{a_J}{a_0}\right), \quad W = \frac{n + \sum_{m=1}^{J} 1/\theta_m^{p-1}}{n + \sum_{m=1}^{J} 1/\theta_m^{p}}, \tag{30}$$

where $n = 10$ and $p = 4$. The limiting is performed in a successive manner in characteristic space [31].

2.4 Implicit Time Marching

As mentioned in the previous subsection, the linear equation system of the variational reconstruction is solved iteratively. The reconstruction will be very expensive if the iteration needs to reach convergence at each single time step. A reconstruction and implicit dual-time stepping coupled iteration procedure has been proposed in [33] to address this computational efficiency issue. In the coupled iteration procedure, the reconstruction iteration is performed only once at each pseudo time step. The coupling between the reconstruction and the time integration makes these two procedures achieve convergence synchronously. By using the coupled iteration, the implicit nature of the variational reconstruction does not result in additional cost, thus ensuring the high computational efficiency of the variational finite volume method.

In this work, the time integration scheme for unsteady flow simulation is the six-stage, fourth-order explicit first stage singly diagonal implicit Runge-Kutta (ESDIRK4) [27]. By using this method, (12) is integrated in time as

$$\overline{\mathbf{U}}_i^{n,1} = \overline{\mathbf{U}}_i^n, \tag{31a}$$

$$\overline{\mathbf{U}}_i^{n,s} = \overline{\mathbf{U}}_i^n + \Delta t^n \sum_{q=1}^{s} a_{sq} \mathbf{R}_i^{n,q}, \quad s = 2, \cdots, 6, \tag{31b}$$

$$\overline{\mathbf{U}}_i^{n+1} = \overline{\mathbf{U}}_i^{n,6}, \tag{31c}$$

where

$$\mathbf{R}_i^{n,q} = \mathbf{R}_i \left(t^n + c_q \Delta t^n, \left\{ \mathbf{U}_j^{n,q} \right\}_{j \in S_i} \right), \tag{32}$$

with $\Delta t^n = t^{n+1} - t^n$ being the time step size, which is in general a user-specified constant. The Runge-Kutta coefficients a_{sq} and c_q of ESDIRK4 can be found in [27].

The implicit and nonlinear equation (31b) is solved iteratively using a dual-time stepping technique, in which a pseudo-time variable τ is introduced. The pseudo-time integration scheme for (31b) is

$$\frac{\overline{\mathbf{U}}_i^{n,s,m+1} - \overline{\mathbf{U}}_i^{n,s,m}}{\Delta \tau_i} + \frac{\overline{\mathbf{U}}_i^{n,s,m+1} - \overline{\mathbf{U}}_i^{n}}{\Delta t^n} = \sum_{q=1}^{s-1} a_{sq} \mathbf{R}_i^{n,q} + a_{ss} \mathbf{R}_i^{n,s,m+1}, \tag{33}$$

where m is the index for pseudo-time step, and

$$\mathbf{R}_i^{n,s,m+1} = \mathbf{R}_i \left(t^n + c_s \Delta t^n, \left\{ \mathbf{U}_j^{n,s,m+1} \right\}_{j \in S_i} \right). \tag{34}$$

By using the linear approximation

$$\mathbf{R}_i^{n,s,m+1} \approx \mathbf{R}_i^{n,s,m} + \sum_{j \in S_i} \frac{\partial \mathbf{R}_i}{\partial \overline{\mathbf{U}}_j} \left(\overline{\mathbf{U}}_j^{n,s,m+1} - \overline{\mathbf{U}}_j^{n,s,m} \right), \tag{35}$$

we obtain a linear equation system to update the cell-average in pseudo-time direction as following

$$\left(\frac{\mathbf{I}}{\Delta \tau_i} + \frac{\mathbf{I}}{\Delta t^n} - a_{ss} \frac{\partial \mathbf{R}_i}{\partial \overline{\mathbf{U}}_i} \right) \Delta \overline{\mathbf{U}}_i^{n,s,m} - a_{ss} \sum_{j \in S_i, j \neq i} \frac{\partial \mathbf{R}_i}{\partial \overline{\mathbf{U}}_j} \Delta \overline{\mathbf{U}}_j^{n,s,m} = \tilde{\mathbf{R}}_i^{n,s,m}, \tag{36}$$

where

$$\Delta \overline{\mathbf{U}}_i^{n,s,m} = \overline{\mathbf{U}}_i^{n,s,m+1} - \overline{\mathbf{U}}_i^{n,s,m}, \quad \tilde{\mathbf{R}}_i^{n,s,m} = \sum_{q=1}^{s-1} a_{sq} \mathbf{R}_i^{n,q} + a_{ss} \mathbf{R}_i^{n,s,m} - \frac{\overline{\mathbf{U}}_i^{n,s,m} - \overline{\mathbf{U}}_i^{n}}{\Delta t^n}. \tag{37}$$

The linear equation system (36) is solved by using the matrix-free LU-SGS approach [34]. The solution is advanced in pseudo-time until $\overline{\mathbf{U}}_i^{n,s,m+1}$ converges to $\overline{\mathbf{U}}_i^{n,s}$, forming an inner iteration at the s-th stage of ESDIRK4. The convergence criterion of the inner iteration is that the L_1 norm of the pseudo-time derivative decreases by a certain number of orders of magnitude.

The local pseudo-time step size is computed by

$$\Delta \tau_i = \frac{\mathrm{CFL}_\tau \overline{\mathbf{\Omega}}_i}{\sum_{f \in \partial \mathbf{\Omega}_i} A_f \lambda_f} \tag{38}$$

where λ_f is the spectral radius estimated on cell interface f and CFL_τ is the CFL number used to control the convergence speed of inner iteration.

3 Positivity-Preserving Algorithm

In compressible flow simulations, density and pressure need to be preserved positive, to avoid nonphysical solutions or numerical instabilities. Therefore, a positivity-preserving property is desired for numerical methods, especially high-order numerical methods, to perform robust simulations of compressible flows.

The set of admissible states is defined by

$$G = \left\{ \mathbf{U} = \begin{pmatrix} \rho \\ \rho\mathbf{u} \\ \rho E \end{pmatrix} \middle| \rho > 0 \text{ and } p = (\gamma - 1)\left(\rho E - \frac{1}{2}\rho\mathbf{u}\cdot\mathbf{u}\right) > 0 \right\}. \tag{39}$$

The density ρ is a linear function of $\mathbf{U}$, and satisfies

$$\rho\left(w\mathbf{U}_1 + (1-w)\mathbf{U}_2\right) = w\rho\left(\mathbf{U}_1\right) + (1-w)\rho\left(\mathbf{U}_2\right), \quad 0 \le w \le 1. \tag{40}$$

The pressure p is a concave function of $\mathbf{U}$ if $\rho\left(\mathbf{U}\right) \ge 0$, and satisfies Jensen's inequality

$$p\left(w\mathbf{U}_1 + (1-w)\mathbf{U}_2\right) \ge wp\left(\mathbf{U}_1\right) + (1-w)p\left(\mathbf{U}_2\right), \quad 0 \le w \le 1, \quad \text{if } \rho\left(\mathbf{U}_1\right) \ge 0,\ \rho\left(\mathbf{U}_2\right) \ge 0. \tag{41}$$

A positivity-preserving numerical scheme satisfies that $\mathbf{U}^{n+1} \in G$, if $\mathbf{U}^n \in G$. In a finite volume method, the numerical solution at t^{n+1} is

$$\mathbf{U}_i^{n+1}(\mathbf{x}) = \overline{\mathbf{U}}_i^{n+1} + \sum_{l=1}^{\mathrm{N_b}(k)} \mathbf{U}_i^{n+1,l}\varphi_{i,l}(\mathbf{x}), \quad i = 1, \cdots, N. \tag{42}$$

An implicit finite volume solution $\mathbf{U}^{n+1} \in G$ can be obtained through the following two steps:

(1) Given $\mathbf{U}^n \in G$, obtain $\overline{\mathbf{U}}^{n+1} \in G$ by using a proper implicit time-stepping scheme;

(2) Given $\overline{\mathbf{U}}^{n+1} \in G$, obtain $\mathbf{U}^{n+1} \in G$ by using a proper limiting algorithm.

The second step can be accomplished by using the scaling limiter of Zhang et al. [16]. Therefore, the key of developing a positivity-preserving implicit finite volume method is to design an implicit time-stepping scheme that leads to admissible updated cell-averages.

In this paper, we propose a positivity-preserving algorithm for finite volume methods with implicit dual-time discretizations. In this algorithm, the physical and pseudo time step sizes are controlled by limiting the solution changes, to obtain admissible cell-averages at the next physical and pseudo time steps, respectively. The details of the proposed positivity-preserving algorithm will be presented in the following subsections.

3.1 Physical Time Step Limiting

In an unsteady flow simulation using implicit time marching, the physical time step size is usually set as a user-specified constant, i.e., $\Delta t^n = \Delta t_{max}$, based on a prior estimation of the time scale of the problem. For challenging problems, there are transient states with very low density or pressure, making it difficult for the numerical method to preserve positivity using the original time step size Δt_{max}. The positivity-preserving difficulty may be overcome by reducing the time step size temporarily. Therefore, we propose to control the physical time step size dynamically based on a positivity-preserving principle

$$\rho\left(\overline{\mathbf{U}}_i^{n+1}\right) > 0, \quad p\left(\overline{\mathbf{U}}_i^{n+1}\right) > 0, \tag{43}$$

for all cells $i = 1, \cdots, N$. However, the solution $\overline{\mathbf{U}}_i^{n+1}$ is unknown at the beginning of the time step. To overcome this difficulty, we propose an solution estimation based on a forward Euler time stepping, which is given by

$$\overline{\mathbf{U}}_i^{n+1} \approx \overline{\mathbf{U}}_i^n + \Delta t^n \mathbf{R}_i^n. \tag{44}$$

A time step size Δt_{pp}^n satisfying the positivity-preserving principle (43) can be determined by limiting the solution changes as follows

$$\begin{aligned} \rho\left(\overline{\mathbf{U}}_i^n + \Delta t_{pp}^n \mathbf{R}_i^n\right) &\ge (1-\eta_t)\rho\left(\overline{\mathbf{U}}_i^n\right) > 0, \\ p\left(\overline{\mathbf{U}}_i^n + \Delta t_{pp}^n \mathbf{R}_i^n\right) &\ge (1-\eta_t)p\left(\overline{\mathbf{U}}_i^n\right) > 0, \end{aligned} \tag{45}$$

for all cells $i = 1, \cdots, N$, where $\eta_t \in (0, 1)$ is the parameter used to control the relative solution changes. It is noted that the forward Euler scheme is first-order accurate, and thus the error of the estimation in

(44) is $\mathcal{O}(\Delta t^2)$. As the allowable solution changes in (45) are $\eta_t \rho\left(\overline{\mathbf{U}}_i^n\right) \sim \mathcal{O}(1)$ and $\eta_t p\left(\overline{\mathbf{U}}_i^n\right) \sim \mathcal{O}(1)$, the positivity of $\overline{\mathbf{U}}_i^{n+1}$ can be preserved, provided that η_t is not overly large.

The allowable time step size for cell i can be determined in two steps according to (45). First, we assume an allowable time step size $\alpha_{t,i,\rho}\Delta t_{max}$, with $\alpha_{t,i,\rho} \in (0,1]$, to preserve positivity of density. According to (40) and (45), we have

$$\rho\left(\overline{\mathbf{U}}_i^n + \alpha_{t,i,\rho}\Delta t_{max}\mathbf{R}_i^n\right) = (1-\alpha_{t,i,\rho})\,\rho\left(\overline{\mathbf{U}}_i^n\right) + \alpha_{t,i,\rho}\rho\left(\overline{\mathbf{U}}_i^n + \Delta t_{max}\mathbf{R}_i^n\right) \geq (1-\eta_t)\rho\left(\overline{\mathbf{U}}_i^n\right), \quad (46)$$

and thus obtain

$$\alpha_{t,i,\rho} = \begin{cases} 1, & \text{if } \rho\left(\overline{\mathbf{U}}_i^n + \Delta t_{max}\mathbf{R}_i^n\right) \geq \rho\left(\overline{\mathbf{U}}_i^n\right), \\ \min\left(1, \dfrac{-\eta_t\rho\left(\overline{\mathbf{U}}_i^n\right)}{\rho\left(\overline{\mathbf{U}}_i^n + \Delta t_{max}\mathbf{R}_i^n\right) - \rho\left(\overline{\mathbf{U}}_i^n\right)}\right), & \text{else.} \end{cases} \quad (47)$$

Second, we assume an allowable time step size $\alpha_{t,i,p}\alpha_{t,i,\rho}\Delta t_{max}$, with $\alpha_{t,i,p} \in (0,1]$, to preserve positivity of pressure. According to (41) and (45), we have

$$p\left(\overline{\mathbf{U}}_i^n + \alpha_{t,i,p}\alpha_{t,i,\rho}\Delta t_{max}\mathbf{R}_i^n\right) \geq (1-\alpha_{t,i,p})\,p\left(\overline{\mathbf{U}}_i^n\right) + \alpha_{t,i,p}p\left(\overline{\mathbf{U}}_i^n + \alpha_{t,i,\rho}\Delta t_{max}\mathbf{R}_i^n\right) \geq (1-\eta_t)p\left(\overline{\mathbf{U}}_i^n\right), \quad (48)$$

and thus obtain a sufficient condition

$$\alpha_{t,i,p} = \begin{cases} 1, & \text{if } p\left(\overline{\mathbf{U}}_i^n + \alpha_{t,i,\rho}\Delta t_{max}\mathbf{R}_i^n\right) \geq p\left(\overline{\mathbf{U}}_i^n\right), \\ \min\left(1, \dfrac{-\eta_t p\left(\overline{\mathbf{U}}_i^n\right)}{p\left(\overline{\mathbf{U}}_i^n + \alpha_{t,i,\rho}\Delta t_{max}\mathbf{R}_i^n\right) - p\left(\overline{\mathbf{U}}_i^n\right)}\right), & \text{else.} \end{cases} \quad (49)$$

The global time step size is computed by

$$\Delta t^n = \alpha_t \Delta t_{max} \quad (50)$$

where

$$\alpha_t = \min_i(\alpha_{t,i,\rho}\alpha_{t,i,p}). \quad (51)$$

It can be observed from the above equations that $\Delta t^n \leq \Delta t_{max}$, as $0 < \alpha_t \leq 1$.

3.2 Pseudo Time Step Limiting

The physical time step limiting strategy described in the previous subsection is used to obtain a solution $\overline{\mathbf{U}}_i^{n+1} \in G$, which is the converged solution of the last stage of ESDIRK4 (31). However, intermediate solutions with negative density or pressure might be encountered during the inner iteration at each stage. Therefore, we propose to limit the solution changes by controlling the pseudo-time step sizes during the inner iteration process, similar to the limiting strategy in physical time direction in Section 3.1.

For the m-th inner iteration at the s-stage of ESDIRK4, we have the following positivity-preserving requirement

$$\rho\left(\overline{\mathbf{U}}_i^{n,s,m+1}\right) > 0, \quad p\left(\overline{\mathbf{U}}_i^{n,s,m+1}\right) > 0, \quad (52)$$

for all cells $i = 1, \cdots, N$. However, $\overline{\mathbf{U}}_i^{n,sm+1}$ is unknown at the beginning of the pseudo-time step. Based on an explicit dual-time discretization of (31b) given by

$$\frac{\overline{\mathbf{U}}_i^{n,s,m+1} - \overline{\mathbf{U}}_i^{n,s,m}}{\Delta\tau_i} + \frac{\overline{\mathbf{U}}_i^{n,s,m} - \overline{\mathbf{U}}_i^{n}}{\Delta t^n} \approx \sum_{q=1}^{s-1} a_{sq}\mathbf{R}_i^{n,q} + a_{ss}\mathbf{R}_i^{n,s,m}, \quad (53)$$

we obtain an solution estimation

$$\overline{\mathbf{U}}_i^{n,s,m+1} \approx \overline{\mathbf{U}}_i^{n,s,m} + \Delta\tau_i\tilde{\mathbf{R}}_i^{n,s,m}, \quad (54)$$

with $\tilde{\mathbf{R}}_i^{n,s,m}$ defined in (37). A pseudo-time step size $\Delta\tau_{i,pp}$ satisfying the positivity-preserving principle (52) can be determined by limiting the solution changes as follows

$$\begin{aligned} \rho\left(\overline{\mathbf{U}}_i^{n,s,m}+\Delta\tau_{i,pp}\tilde{\mathbf{R}}_i^{n,s,m}\right) &\geq (1-\eta_\tau)\rho\left(\overline{\mathbf{U}}_i^{n,s.m}\right)>0,\\ p\left(\overline{\mathbf{U}}_i^{n,s,m}+\Delta\tau_{i,pp}\tilde{\mathbf{R}}_i^{n,s,m}\right) &\geq (1-\eta_\tau)p\left(\overline{\mathbf{U}}_i^{n,s,m}\right)>0, \end{aligned} \tag{55}$$

where $\eta_\tau \in (0,1)$ is the parameter used to control the relative solution changes.

The allowable pseudo-time step size for cell i can be determined according to (55) using the same two-step procedure in Section 3.1. The allowable time step size is $\Delta\tau_{i,pp}=\alpha_{\tau,i,p}\alpha_{\tau,i,\rho}\Delta\tau_i$, where

$$\alpha_{\tau,i,\rho}=\begin{cases} 1, & \text{if } \rho\left(\overline{\mathbf{U}}_i^{n,s,m}+\Delta\tau_i\tilde{\mathbf{R}}_i^{n,s,m}\right)\geq\rho\left(\overline{\mathbf{U}}_i^{n,s,m}\right),\\ \min\left(1,\dfrac{-\eta_\tau\rho\left(\overline{\mathbf{U}}_i^{n,s,m}\right)}{\rho\left(\overline{\mathbf{U}}_i^{n,s,m}+\Delta\tau_i\tilde{\mathbf{R}}_i^{n,s,m}\right)-\rho\left(\overline{\mathbf{U}}_i^{n,s,m}\right)}\right), & \text{else,}\end{cases} \tag{56}$$

and

$$\alpha_{\tau,i,p}=\begin{cases} 1, & \text{if } p\left(\overline{\mathbf{U}}_i^{n,s,m}+\alpha_{\tau,i,\rho}\Delta\tau_i\tilde{\mathbf{R}}_i^{n,s,m}\right)\geq p\left(\overline{\mathbf{U}}_i^{n,s,m}\right),\\ \min\left(1,\dfrac{-\eta_\tau p\left(\overline{\mathbf{U}}_i^{n,s,m}\right)}{p\left(\overline{\mathbf{U}}_i^{n,s,m}+\alpha_{\tau,i,\rho}\Delta\tau_i\tilde{\mathbf{R}}_i^{n,s,m}\right)-p\left(\overline{\mathbf{U}}_i^{n,s,m}\right)}\right), & \text{else.}\end{cases} \tag{57}$$

It is noted that $\overline{\mathbf{U}}_i^{n,s,m+1}$ is obtained by performing one LU-SGS iteration for (36), thus being not the converged solution of the linear equation system. Therefore, there is not theoretical result for the error of estimation in (54), and thus not guarantee for the positivity of the solution $\overline{\mathbf{U}}_i^{n,s,m+1}$. To address this issue, an increment correction will be introduced in the next subsection.

3.3 Increment Correction

As pointed out in the previous subsection, although the physical and pseudo time steps are limited, the intermediate results may have negative density or pressure before the convergence of the LU-SGS iteration. Therefore, a pseudo-time increment correction is designed to ensure the positivity of the updated solutions. The increment correction is based on the following conditions

$$\begin{aligned} \rho\left(\overline{\mathbf{U}}_i^{n,s,m}+\alpha_{\Delta,i}\Delta\overline{\mathbf{U}}_i^{n,s,m}\right) &\geq (1-\eta_\Delta)\rho\left(\overline{\mathbf{U}}_i^{n,s,m}\right),\\ p\left(\overline{\mathbf{U}}_i^{n,s,m}+\alpha_{\Delta,i}\Delta\overline{\mathbf{U}}_i^{n,s,m}\right) &\geq (1-\eta_\Delta)p\left(\overline{\mathbf{U}}_i^{n,s,m}\right), \end{aligned} \tag{58}$$

where $\eta_\Delta \in (0,1)$ is a relaxation parameter to control the relative solution changes. By using the same two-step approach used in physical/pseudo time step limiting, the relaxation parameter is computed by

$$\alpha_{\Delta,i}=\alpha_{\Delta,i,\rho}\alpha_{\Delta,i,p}, \tag{59}$$

where

$$\alpha_{\Delta,i,\rho}=\begin{cases} 1, & \text{if } \rho\left(\overline{\mathbf{U}}_i^{n,s,m}+\Delta\overline{\mathbf{U}}_i^{n,s,m}\right)\geq\rho\left(\overline{\mathbf{U}}_i^{n,s,m}\right),\\ \min\left(1,\dfrac{-\eta_\Delta\rho\left(\overline{\mathbf{U}}_i^{n,s,m}\right)}{\rho\left(\overline{\mathbf{U}}_i^{n,s,m}+\Delta\overline{\mathbf{U}}_i^{n,s,m}\right)-\rho\left(\overline{\mathbf{U}}_i^{n,s,m}\right)}\right), & \text{else,}\end{cases} \tag{60}$$

and

$$\alpha_{\Delta,i,p} = \begin{cases} 1, & \text{if } p\left(\overline{\mathbf{U}}_i^{n,s,m} + \alpha_{\Delta,i,\rho}\Delta\overline{\mathbf{U}}_i^{n,s,m}\right) \geq p\left(\overline{\mathbf{U}}_i^{n,s,m}\right), \\ \min\left(1, \dfrac{-\eta_\Delta p\left(\overline{\mathbf{U}}_i^{n,s,m}\right)}{p\left(\overline{\mathbf{U}}_i^{n,s,m} + \alpha_{\Delta,i,\rho}\Delta\overline{\mathbf{U}}_i^{n,s,m}\right) - p\left(\overline{\mathbf{U}}_i^{n,s,m}\right)}\right), & \text{else.} \end{cases} \tag{61}$$

A solution $\overline{\mathbf{U}}_i^{n,s,m+1} \in G$ is then computed by

$$\overline{\mathbf{U}}_i^{n,s,m+1} = \overline{\mathbf{U}}_i^{n,s,m} + \alpha_{\Delta,i}\Delta\overline{\mathbf{U}}_i^{n,s,m}. \tag{62}$$

The positivity-preserving algorithm developed in this paper is based on the scaling limiter [16], physical time step limiting, pseudo time step limiting and pseudo time increment correction. The scaling limiter is accuracy-preserving [16]. It is known that reducing physical time step sizes does not affect the order of accuracy of the time integration. Furthermore, reducing pseudo time step sizes and increments at intermediate steps do not affect the accuracy of the converged solutions of the inner iterations. Therefore, the proposed positivity-preserving approach is accuracy-preserving. Furthermore, it is noted that the proposed positivity-preserving algorithm is applicable to general finite volume methods with implicit dual-time stepping, not limited to finite volume methods with high-order variational reconstruction or high-order ESDIRK time discretization.

4 Numerical Results

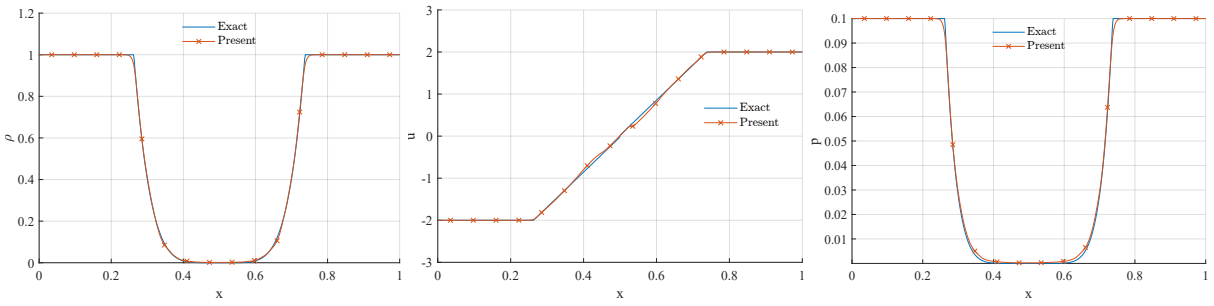
The proposed positivity-preserving algorithm is verified using a series of benchmark test cases. The numerical scheme used is the fourth-order accurate implicit finite volume method based on a cubic variational reconstruction and an ESDIRK4 time integration. In the simulations, the relaxation parameters are set as $\eta_t = 0.8$, $\eta_\tau = 0.5$ and $\eta_\Delta = 0.9$, unless otherwise specified. The convergence criterion for inner iteration is that the L_1 norm of the pseudo time derivative decreases by three orders of magnitude.

4.1 Double Rarefaction

The one-dimensional double rarefaction problem [35] is a Riemann problem with the following initial conditions

$$(\rho, u, p) = \begin{cases} (1, -2, 0.1), & x < 0.5, \\ (1, 2, 0.1), & \text{else.} \end{cases} \tag{63}$$

The flow is inviscid and the ratio of specific heat is $\gamma = 1.4$. In this problem, a vacuum lies in in the middle, adjacent to the ends of two rarefaction waves. The computational domain is $[0, 1]$. We perform a numerical simulation on a uniform mesh with grid size $\Delta x = 1/400$, up to $t = 0.1$ using a user-specified time step size $\Delta t_{max} = 5 \times 10^{-3}$. In each inner iteration, the CFL number for local pseudo time step CFL_τ is initiated as 0.1 and increased gradually up to its maximum 10 at the tenth iteration. The numerical results are shown in Figure 4. It is observed from Figure 4 that, the numerical solution is essentially oscillation-free and agrees well with the reference solution.



(a) Density (b) Velocity (c) Pressure

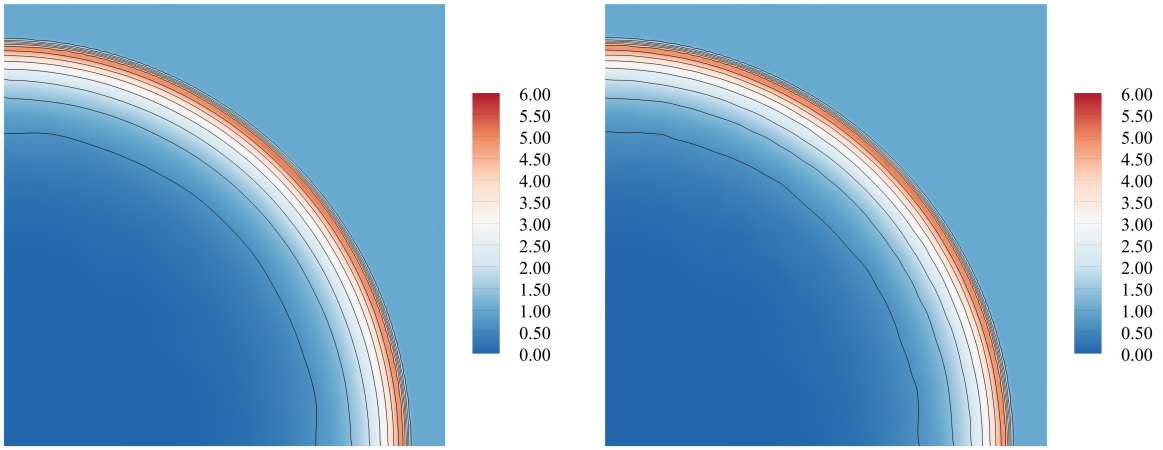
Figure 4: Results of the double rarefaction problem at $t = 0.1$.

4.2 Sedov Blast Wave

The Sedov blast problem [17, 36] is a popular case to test positivity-preserving properties of numerical schemes. The governing equations are the Euler equations and the ratio of specific heat is $\gamma = 1.4$. The computational domain is $[0, 1.1] \times [0, 1.1]$, partitioned into rectangular cells with grid size $\Delta x = \Delta y = 1.1/160$. The initial conditions are

$$(\rho, u, v, p) = \begin{cases} (1, 0, 0, 4 \times 10^{-9}), & \text{if } x > \Delta x \text{ or } y > \Delta y, \\ (1, 0, 0, \dfrac{(\gamma - 1)\,\varepsilon^0}{\Delta x \Delta y}), & \text{else,} \end{cases} \tag{64}$$

where ε^0 is the total amount of release energy. By choosing $\varepsilon^0 = 2.44816 \times 10^5$, the solution consists of a diverging infinite strength shock wave whose front is located at radius $r = 1$ at $t = 10^{-3}$, with a peak density reaching 6. All boundaries are slip walls. Two simulations are performed up to $t = 10^{-3}$, using 1026 time steps with $\Delta t_{max} = 10^{-6}$ and 297 time steps with $\Delta t_{max} = 10^{-5}$, respectively.



(a) $\Delta t_{max} = 10^{-6}$ (b) $\Delta t_{max} = 10^{-5}$

Figure 5: Density contours of the Sedov blast wave problem.

The computed density contours at $t = 10^{-3}$ are shown in Figure 5. It is observed that the numerical solutions are essentially oscillation-free. To check the solutions in a more intuitive way, we plot both the analytical and numerical density distributions along the $y = x$ line in Figure 6. It is shown in Figure 6 that, both numerical solutions agree well with the analytic solution, and the one with the smaller time step size has higher resolution.

4.3 Mach 2000 Jet

The Mach 2000 jet problem [16] is a challenging case as it has extremely strong discontinuities. Following the practice of [24], the current paper studies two cases with $Re = \infty$ and $Re = 100$ to demonstrate the capability of the proposed positivity-preserving algorithm to deal with inviscid and viscous compressible flows, respectively. The governing equations are the Navier-Stokes equations with $\gamma = 5/3$. The computational domain is $[0, 1] \times [-0.25, 0.25]$. Far-field boundary conditions are imposed on the upper, lower and right boundaries. The following inflow conditions are imposed on the left boundary:

$$(\rho, u, v, p) = \begin{cases} (5, 800, 0, 0, 4127), & \text{if } -0.05 \le y \le 0.05, \\ (0.5, 0, 0, 0.4127), & \text{else.} \end{cases} \tag{65}$$

The initial conditions are $(\rho, u, v, p) = (0.5, 0, 0, 0.4127)$. The numerical simulations are performed on a uniform rectangular mesh with 800×800 cells, up to $t = 10^{-3}$ using a user-specified time step size $\Delta t_{max} = 1 \times 10^{-6}$. The relaxation parameters are $\eta_t = 0.6$, $\eta_\tau = 0.5$, $\eta_\Delta = 0.9$. In each inner iteration, the CFL number for local pseudo time step CFL_τ is initiated as 0.1 and increased gradually up to its maximum 2 at the tenth iteration.

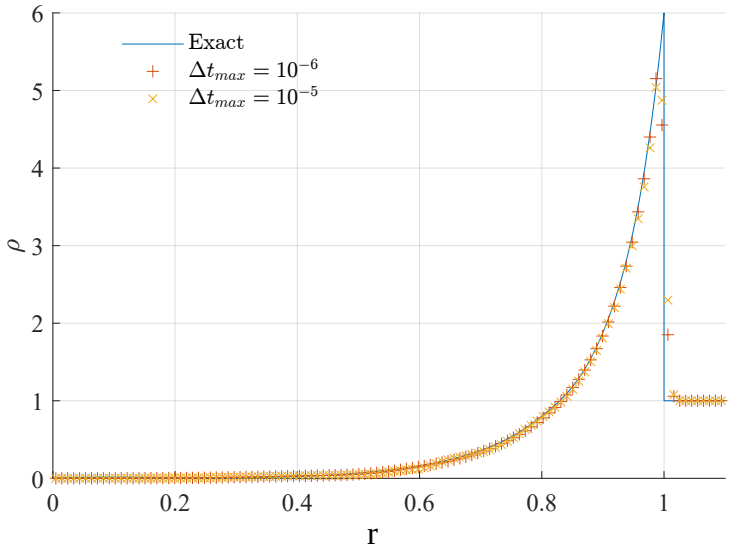


Figure 6: Density distributions of the Sedov blast wave problem along the diagonal line.

The computed density and pressure contours at $t = 10^{-3}$ in logarithmic scales are shown in Figure 7, 8, 9 and 10. It is observed from these figures that the numerical results are essentially non-oscillatory, and small-scale structures are smeared out in the $Re = 100$ case due to the viscous effect.

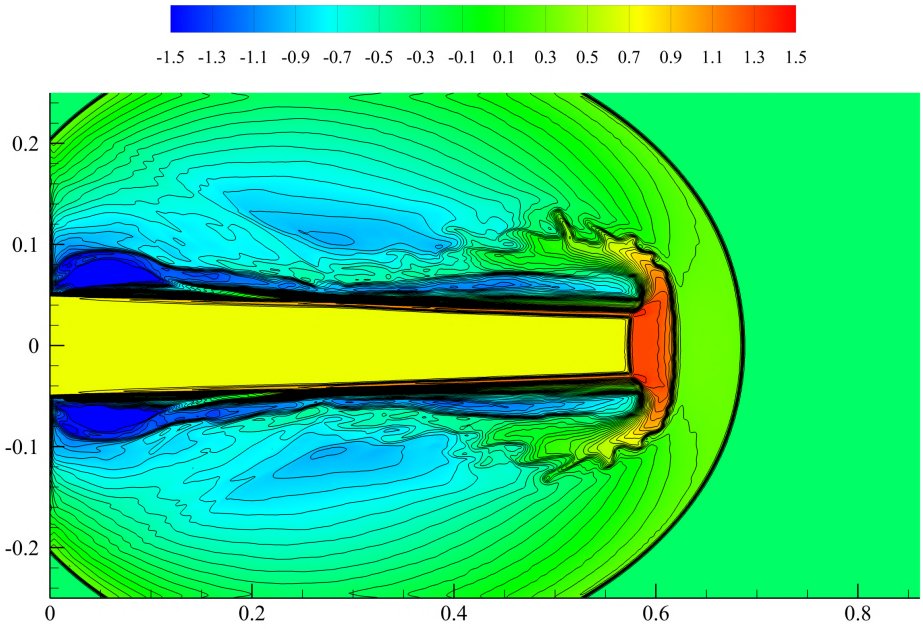


Figure 7: $\log_{10}(\rho)$ of $Re = \infty$ Mach 2000 jet, 40 contour lines from -1.5 to 1.5.

4.4 Shock Diffraction

The shock diffraction problem [16], in which a shock passes a backward facing corner, is used to test the positivity-preserving capability of the proposed algorithm. The governing equations are the Euler equations with $\gamma = 1.4$. The computational domain is the union of $[0, 1] \times [6, 11]$ and $[1, 13] \times [0, 11]$. The initial condition is a Mach number 5.09 shock located at $x = 0.5$, moving into undisturbed air ahead of the shock at a state $(\rho, u, v, p) = (1.4, 0, 0, 1)$. Inflow and slip solid wall boundary conditions are imposed on the left side $x = 0$, $6 \leq y \leq 11$ and other boundaries, respectively. The simulation is performed on

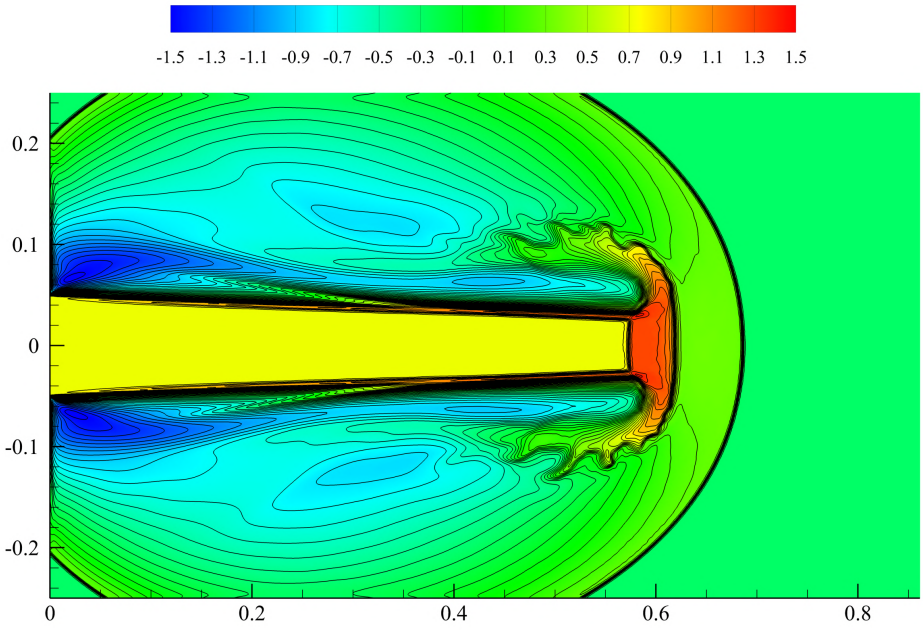


Figure 8: $\log_{10}(\rho)$ of $Re = 100$ Mach 2000 jet, 40 contour lines from -1.5 to 1.5.

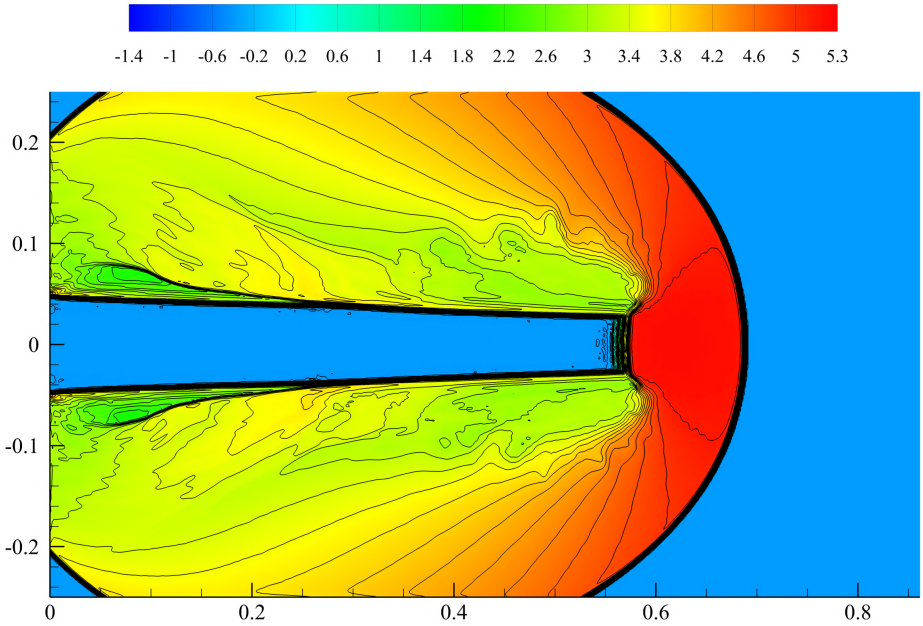


Figure 9: $\log_{10}(p)$ of $Re = \infty$ Mach 2000 jet, 40 contour lines from -1.4 to 5.3.

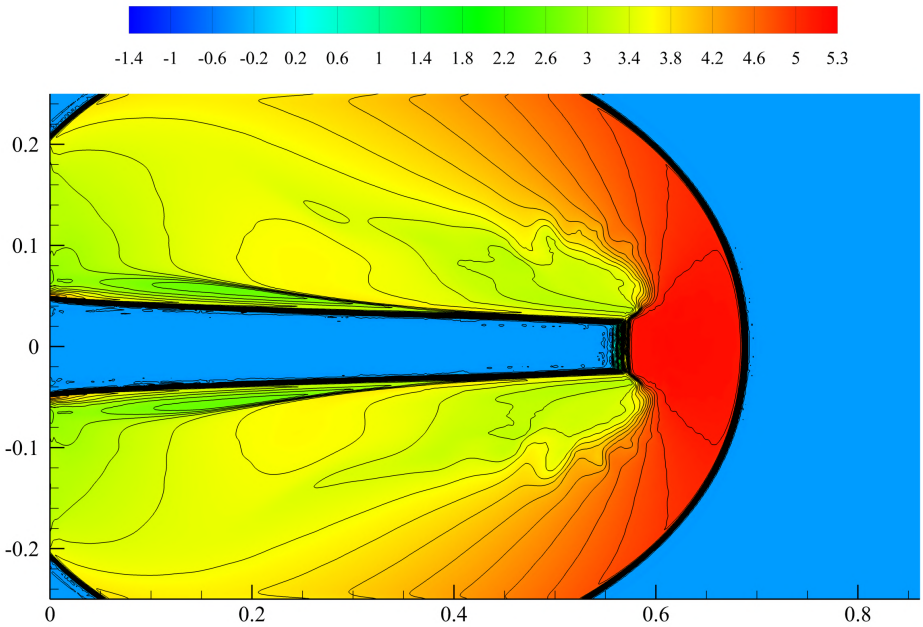


Figure 10: $\log_{10}(p)$ of $Re = 100$ Mach 2000 jet, 40 contour lines from -1.4 to 5.3.

a rectangular mesh with grid size $\Delta x = \Delta y = 1/80$, up to $t = 2.3$ using a user-specified time step size $\Delta t_{max} = 1 \times 10^{-3}$. The CFL number for pseudo-time step is $\text{CFL}_\tau = 1$.

The computed density and pressure contours in logarithmic scales at $t = 2.3$ are shown in Figure 11. It is observed from Figure 11 that, the numerical solutions are essentially oscillation-free and small-scale shock structures are well resolved.

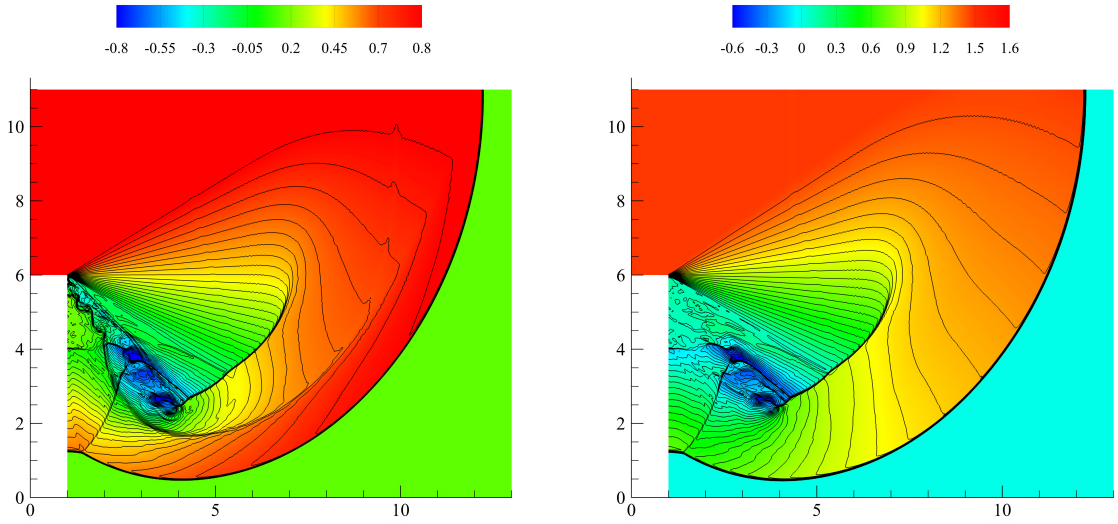
4.5 Shock Reflection and Diffraction Around a Wedge

The shock reflection and diffraction problem [37] is similar to the shock diffraction problem, with the corner replaced by a wedge. The computational domain is $[0, 3] \times [0, 2]$ with a 30° wedge placed at $x \in [0.2, 1.2]$, while the tip of the wedge is at $(x, y) = (1.2, 1/\sqrt{3})$. Initially, there is a Mach 10 shock at $x = 0.2$ moving into undisturbed air at a state $(\rho, u, v, p) = (1.4, 0, 0, 1)$. The $x = 0$ boundary on the left matches the downstream state of the shock, and all other boundaries are inviscid walls. Before the diffraction of the shock, the development of the shock structure is identical with that in the Mach 10 double Mach reflection problem [38]. Similar to the Mach 2000 jet example, $\text{Re} = \infty$ and $\text{Re} = 100$ cases are investigated. A triangular mesh with grid size $h = 1/160$, shown in Figure 12, is used in the simulations. The simulations are performed up to $t = 0.245$ using a user-specified time step size $\Delta t_{max} = 2 \times 10^{-4}$. In each inner iteration, the CFL number for local pseudo time step CFL_τ is initiated as 0.01 and increased gradually up to its maximum 1 at the tenth iteration.

The computed density and pressure contours are shown in Figure 13 and 14. It is observed that the numerical solutions are essentially non-oscillatory. The Kelvin-Helmholtz instability induced in the shear layer is well resolved in Figure 13 for the $\text{Re} = \infty$ case. While in Figure 14 for the $\text{Re} = 100$ case, the small-scale flow structures around the shear layer and the wedge are smeared due to the existence of physical viscosity.

5 Conclusions

This paper presents a positivity-preserving algorithm for finite volume methods using dual-time stepping for compressible flow simulations. In the positivity-preserving algorithm, admissible cell-averages are obtained by limiting solution changes, which is accomplished by controlling physical and pseudo time step sizes. To overcome the difficulty of unknown solution changes, we propose to use explicit time discretizations to obtain efficient estimations of future states. The allowable time step sizes are determined by



(a) $\log_{10}(\rho)$, 40 contour lines from -0.8 to 0.8

(b) $\log_{10}(p)$, 40 contour lines from -0.6 to 1.6

Figure 11: Density and pressure contours of the shock diffraction problem.

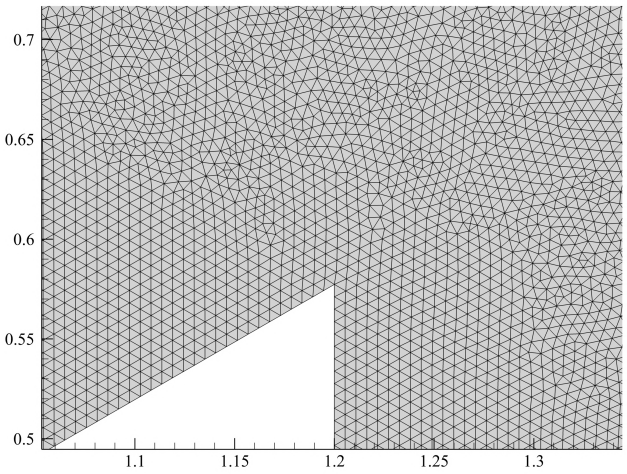
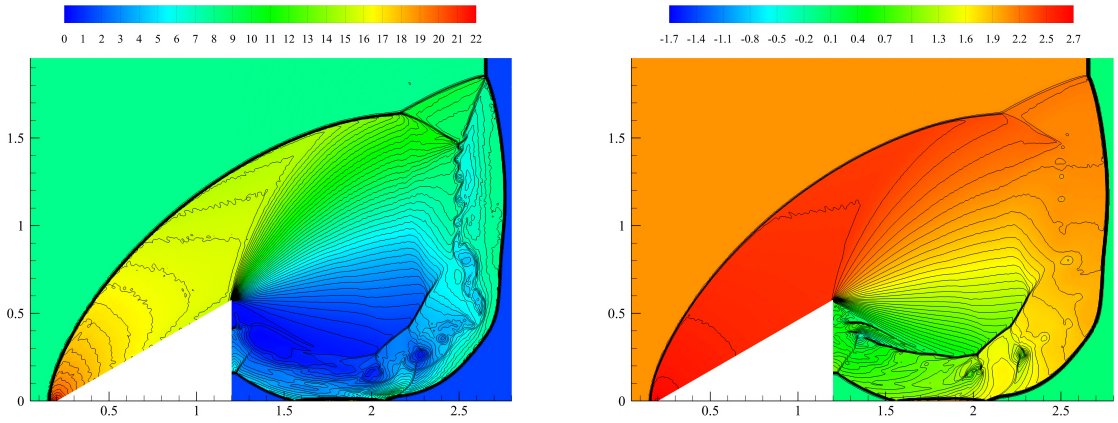


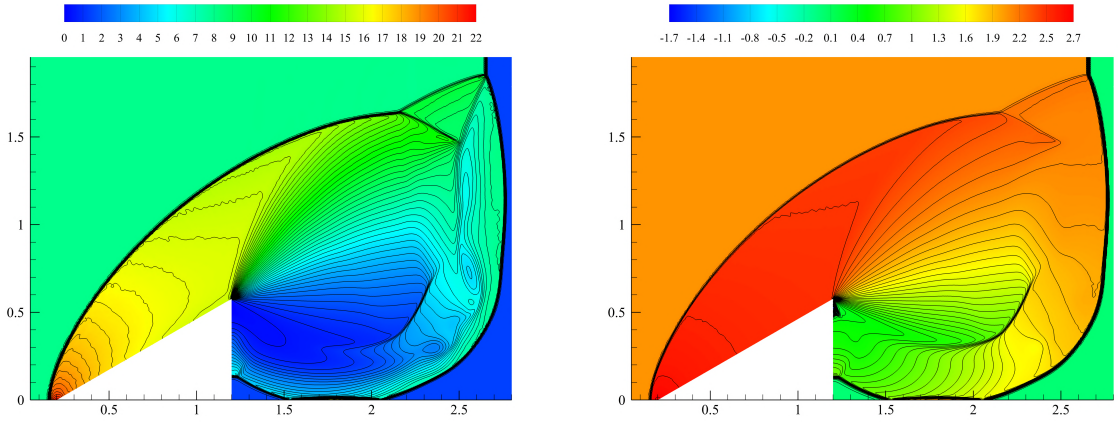
Figure 12: Mesh used in the shock reflection and diffraction problem.



(a) ρ, 60 contour lines from 0 to 22

(b) $\log_{10}(p)$, 60 contour lines from -1.7 to 2.7

Figure 13: Density and pressure contours for the shock reflection and diffraction problem, $\mathrm{Re} = \infty$.



(a) ρ, 60 contour lines from 0 to 22

(b) $\log_{10}(p)$, 60 contour lines from -1.7 to 2.7

Figure 14: Density and pressure contours for the shock reflection and diffraction problem, $\mathrm{Re} = 100$.

limiting the relative solution changes. Given positive cell-averages, admissible reconstruction polynomials can be obtained by applying a positivity-preserving scaling limiter. The proposed positivity-preserving algorithm is accuracy-preserving. Numerical results for a series of benchmark test cases demonstrate the robustness and high resolution of the positivity-preserving implicit high-order finite volume method.

References

[1] Bram Van Leer. Towards the ultimate conservative difference scheme. v. a second-order sequel to godunov's method. *Journal of computational Physics*, 32(1):101–136, 1979.

[2] Carl F Ollivier-Gooch. Quasi-eno schemes for unstructured meshes based on unlimited data-dependent least-squares reconstruction. *Journal of Computational Physics*, 133(1):6–17, 1997.

[3] Michael Dumbser, Martin Käser, Vladimir A Titarev, and Eleuterio F Toro. Quadrature-free non-oscillatory finite volume schemes on unstructured meshes for nonlinear hyperbolic systems. *Journal of Computational Physics*, 226(1):204–243, 2007.

[4] Qian Wang, Yu-Xin Ren, Jianhua Pan, and Wanai Li. Compact high order finite volume method on unstructured grids III: Variational reconstruction. *Journal of Computational Physics*, 337:1–26, 2017.

[5] William H Reed and Thomas R Hill. Triangular mesh methods for the neutron transport equation. Technical report, Los Alamos Scientific Lab., N. Mex.(USA), 1973.

[6] Bernardo Cockburn and Chi-Wang Shu. Runge–kutta discontinuous galerkin methods for convection-dominated problems. *Journal of scientific computing*, 16:173–261, 2001.

[7] Michael Dumbser, Dinshaw S Balsara, Eleuterio F Toro, and Claus-Dieter Munz. A unified framework for the construction of one-step finite volume and discontinuous Galerkin schemes on unstructured meshes. *Journal of Computational Physics*, 227(18):8209–8253, 2008.

[8] Lingquan Li, Xiaodong Liu, and Hong Luo. A reconstructed discontinuous galerkin method based on variational formulation for compressible flows. *Journal of Computational Physics*, 466:111406, 2022.

[9] Laiping Zhang, Wei Liu, Lixin He, Xiaogang Deng, and Hanxin Zhang. A class of hybrid DG/FV methods for conservation laws I: Basic formulation and one-dimensional systems. *Journal of Computational Physics*, 231(4):1081–1103, 2012.

[10] Rémi Abgrall and Mohamed Mezine. Construction of second order accurate monotone and stable residual distribution schemes for unsteady flow problems. *Journal of Computational Physics*, 188(1):16–55, 2003.

[11] HT Huynh. A flux reconstruction approach to high-order schemes including discontinuous Galerkin methods. *AIAA paper*, 4079:1–42, 2007.

[12] Peter E Vincent, Patrice Castonguay, and Antony Jameson. A new class of high-order energy stable flux reconstruction schemes. *Journal of Scientific Computing*, 47:50–72, 2011.

[13] ZJ Wang and Haiyang Gao. A unifying lifting collocation penalty formulation including the discontinuous Galerkin, spectral volume/difference methods for conservation laws on mixed grids. *Journal of Computational Physics*, 228(21):8161–8186, 2009.

[14] Zhengfu Xu and Xiangxiong Zhang. Bound-preserving high-order schemes. In *Handbook of numerical analysis*, volume 18, pages 81–102. Elsevier, 2017.

[15] Xiangxiong Zhang and Chi-Wang Shu. On maximum-principle-satisfying high order schemes for scalar conservation laws. *Journal of Computational Physics*, 229(9):3091–3120, 2010.

[16] Xiangxiong Zhang and Chi-Wang Shu. On positivity-preserving high order discontinuous galerkin schemes for compressible euler equations on rectangular meshes. *Journal of Computational Physics*, 229(23):8918–8934, 2010.

[17] Xiangxiong Zhang and Chi-Wang Shu. Positivity-preserving high order finite difference weno schemes for compressible euler equations. *Journal of Computational Physics*, 231(5):2245–2258, 2012.

[18] Chuan Fan, Xiangxiong Zhang, and Jianxian Qiu. Positivity-preserving high order finite difference weno schemes for compressible navier-stokes equations. *Journal of Computational Physics*, 467:111446, 2022.

[19] Xiangyu Y Hu, Nikolaus A Adams, and Chi-Wang Shu. Positivity-preserving method for high-order conservative schemes solving compressible euler equations. *Journal of Computational Physics*, 242:169–180, 2013.

[20] Andrew J Christlieb, Yuan Liu, Qi Tang, and Zhengfu Xu. High order parametrized maximum-

principle-preserving and positivity-preserving weno schemes on unstructured meshes. *Journal of Computational Physics*, 281:334–351, 2015.

[21] Dmitri Kuzmin, Manuel Quezada de Luna, David I Ketcheson, and Johanna Grüll. Bound-preserving flux limiting for high-order explicit runge–kutta time discretizations of hyperbolic conservation laws. *Journal of Scientific Computing*, 91(1):21, 2022.

[22] Tao Xiong, Jing-Mei Qiu, and Zhengfu Xu. A parametrized maximum principle preserving flux limiter for finite difference rk-weno schemes with applications in incompressible flows. *Journal of Computational Physics*, 252:310–331, 2013.

[23] Tong Qin and Chi-Wang Shu. Implicit positivity-preserving high-order discontinuous galerkin methods for conservation laws. *SIAM Journal on Scientific Computing*, 40(1):A81–A107, 2018.

[24] Qian-Min Huang, Hanyu Zhou, Yu-Xin Ren, and Qian Wang. A general positivity-preserving algorithm for implicit high-order finite volume schemes solving the euler and navier-stokes equations. *Journal of Computational Physics*, 508:112999, 2024.

[25] Bernard Parent. Positivity-preserving dual time stepping schemes for gas dynamics. *Journal of Computational Physics*, 361:391–411, 2018.

[26] Chenzhou Lian, Guoping Xia, and Charles L Merkle. Solution-limited time stepping to enhance reliability in cfd applications. *Journal of Computational Physics*, 228(13):4836–4857, 2009.

[27] Hester Bijl, Mark H Carpenter, Veer N Vatsa, and Christopher A Kennedy. Implicit time integration schemes for the unsteady compressible navier–stokes equations: laminar flow. *Journal of Computational Physics*, 179(1):313–329, 2002.

[28] Rémi Abgrall. On essentially non-oscillatory schemes on unstructured meshes: analysis and implementation. *Journal of Computational Physics*, 114(1):45–58, 1994.

[29] Oliver Friedrich. Weighted essentially non-oscillatory schemes for the interpolation of mean values on unstructured grids. *Journal of computational physics*, 144(1):194–212, 1998.

[30] Wanai Li, Yu-Xin Ren, Guodong Lei, and Hong Luo. The multi-dimensional limiters for solving hyperbolic conservation laws on unstructured grids. *Journal of Computational Physics*, 230(21):7775–7795, 2011.

[31] Wanai Li and Yu-Xin Ren. The multi-dimensional limiters for solving hyperbolic conservation laws on unstructured grids ii: extension to high order finite volume schemes. *Journal of Computational Physics*, 231(11):4053–4077, 2012.

[32] Wanai Li and Yu-Xin Ren. High-order k-exact weno finite volume schemes for solving gas dynamic euler equations on unstructured grids. *International Journal for Numerical Methods in Fluids*, 70(6):742–763, 2012.

[33] Qian Wang, Yu-Xin Ren, and Wanai Li. Compact high order finite volume method on unstructured grids ii: Extension to two-dimensional euler equations. *Journal of Computational Physics*, 314:883–908, 2016.

[34] Hong Luo, Joseph D Baum, and Rainald Löhner. A fast, matrix-free implicit method for compressible flows on unstructured grids. *Journal of Computational Physics*, 146(2):664–690, 1998.

[35] XY Hu and BC Khoo. Kinetic energy fix for low internal energy flows. *Journal of Computational Physics*, 193(1):243–259, 2004.

[36] François Vilar, Chi-Wang Shu, and Pierre-Henri Maire. Positivity-preserving cell-centered lagrangian schemes for multi-material compressible flows: From first-order to high-orders. part i: The one-dimensional case. *Journal of Computational Physics*, 312:385–415, 2016.

[37] Xiangxiong Zhang. On positivity-preserving high order discontinuous galerkin schemes for compressible navier–stokes equations. *Journal of Computational Physics*, 328:301–343, 2017.

[38] Paul Woodward and Phillip Colella. The numerical simulation of two-dimensional fluid flow with strong shocks. *Journal of computational physics*, 54(1):115–173, 1984.