An efficient invariant-region-preserving central scheme for hyperbolic conservation laws

Ruifang Yan [a,b], Wei Tong [a,b], Guoxian Chen [a,b,*]

[a] *School of Mathematics and Statistics, Wuhan University, Wuhan, 430072, PR China*
[b] *Computational Science Hubei Key Laboratory, Wuhan University, Wuhan, 430072, PR China*

ARTICLE INFO

ABSTRACT

Due to the Riemann solver free and avoiding characteristic decomposition, the central scheme is a simple and efficient tool for numerical solution of hyperbolic conservation laws (Nessyahu and Tadmor, J. Comput. Phys., 87(2):314-329,1990). But the theoretical Courant number CFL in order to preserve the invariant region of the numerical solution is very small, and there is lack of the stability proof for nonlinear systems. By adding a limiter on the reconstructed slope without requiring clipping condition, we enlarge the value of the CFL to admit larger time step. Then a widely applicable stability proof, which is suitable for general hyperbolic conservation laws, is given by writing the evolved solution as convex combinations in terms of the Lax-Friedrichs scheme. Some numerical experiments are carried out to verify the robustness.

© 2022 Elsevier Inc. All rights reserved.

1. Introduction

The hyperbolic conservation laws belong to an important class of partial differential equations (PDE) which take the form

$$\mathbf{u}_t + \mathbf{f}(\mathbf{u})_x = 0, \quad \mathbf{u}(x, t = 0) = \mathbf{u}_0(x) \tag{1}$$

where $\mathbf{u} = (u_1, u_2, \ldots, u_p)^\top \in \mathbb{R}^p$ is the conservative vector with p components, $\mathbf{f} = (f_1, f_2, \ldots, f_p)^\top$ is the flux function which is assumed nonlinear and takes values in $\mathbb{R}^p$.

The admissible states Ω of (1) is convex and forms an invariant set in the sense that

$$\mathbf{u}(x, t = 0) \in \Omega, \quad \forall x \in \mathbb{R} \quad \Longrightarrow \quad \mathbf{u}(x, t > 0) \in \Omega \quad \forall x \in \mathbb{R}. \tag{2}$$

The above property is called the *invariant-region-preserving* (IRP) principle. For the scalar problem, such principle becomes the *minimum-maximum-preserving* (MMP) principle, which is related to the entropy inequality that allows one to select the unique physical solution among the weak solutions [1]. For the nonlinear Euler equations describing the inviscid gas dynamics, it is called the *positivity-preserving* (PP) principle, which constrains the positivity of density, pressure, and internal energy.

In the discrete level, we also require IRP principle of the numerical solution. Any numerical method that satisfies these constraints is usually called IRP in the literatures [2–13]. The objective of the present work is to put forward a type of IRP

---

* Corresponding author.
  E-mail address: gxchen.math@whu.edu.cn (G. Chen).

central scheme that is explicit in time, second-order accurate in both space and time. Compared with the upwind Godunov-type scheme [14–20], the central scheme [21–24] requires no Riemann solvers or characteristic decomposition and provides a black box solver which can efficiently and universally simulate a wide variety of problems.

The central scheme is the sequel to the first-order Lax-Friedrichs (LxF) scheme, which is IRP and total-variation-diminishing (TVD) and satisfies a discrete entropy inequality. To enhance the spatial accuracy order, Nessyahu and Tadmor [21] use MUSCL-type interpolations from the known staggered cell-averages at the current time level and give a second-order approximation of the flux according to the midpoint rule at the cell centers. The solution with lower diffusion than that by the LxF scheme is evolved to the new time level at the staggered dual cell. The Nessyahu-Tadmor (NT) scheme gives a higher-order approximation, which compensates for the typical excessive viscosity of the first-order LxF piecewise-constant solution. To transform the NT scheme into an unstaggered-central scheme, the authors reconstruct a piecewise-linear inter-polant through the calculated staggered cell-averages at the new time level and project it on the primal cells [24]. This version of the central scheme not only retains the desirable advantages of the central scheme, but also avoids staggered grids and hence is simpler to implement when complex geometries and boundary conditions are involved. There are also some other variations of the central scheme [25–29]. In recent years, the unstaggered-central scheme is applied to many problems, such as idea magnetohydrodynamics (MHD) equations [30], Euler equations with gravitation [31], shallow water flows [32–36].

The stability of higher-order central schemes are discussed in many papers [21,23,24,28,29]. The scalar non-oscillatory properties of the second-order NT scheme were proved, including TVD, cell entropy inequality, and L_1 loc-error estimation. It does not increase the number of initial extremea(-as does the exact entropy solution operator). The local scalar MMP property is proved in [23] under the assumption that the slope limiter is clipping, which means that the neighboring discrete slopes cannot have opposite signs, which is the key to estimate the difference between two neighboring midvalues. However, the clipping limiter reduces the scheme back to the first-order LxF scheme in an extreme cell because the limiter will set a zero slope in this cell [29]. Additionally, the stable CFL number is too small: it is shown in [23] that CFL ≈ 0.1 for minmod limiter (the generalized minmod with $\theta = 1$) and CFL $= 0$ for the generalized minmod with $\theta = 2$ in two-dimensional case. A small CFL number results in a small time step which is disastrous for central scheme because the numerical viscosity is proportion to $O(1/\Delta t)$ [25,29,37,38]. Although the extension of the systems is carried out by componentwise application of the scalar framework, the IRP principle of general systems and even the PP principle of Euler systems are lack of theoretical analysis. Actually, many of the applications mentioned above are not IRP. For example, the application on the shallow water equations with wet-dry state is not positivity-preserving [34,35] which will be illustrated in [39].

In present paper, we will have an insight into the unstaggered-central scheme. Firstly, the clipping limiter for the initial data reconstruction is replaced by an extended IRP limiter. After the non-oscillatory reconstruction, an IRP limiter [9,10] is applied to the reconstructed slopes such that the interpolation over the cell belongs to the invariant region. This extra limiter does not disturb the accuracy order in the extreme cells. The second contribution is about the proof of the IRP property. Instead of calculating the flux difference between two neighboring midvalues at the half time step using the reconstructed non-clipping slopes of both conservative variables and fluxes, we propose a so called forward-backward splitting method. It is an easy task to prove that the predicted value belongs to the invariant domain, which is achieved by rewriting the predicted value in the form of the LxF scheme. The most important part is the proof that the corrected value also belongs to the invariant domain, which is the core of this paper. It is realized by rewriting the corrected value as a convex decomposition in terms of the LxF schemes solving a forward process of the PDE: $\mathbf{u}_t + \mathbf{f}(\mathbf{u})_x = 0$ and the backward process of the PDE: $\mathbf{u}_t - \mathbf{f}(\mathbf{u})_x = 0$. Finally, we prove that the updated solution belongs to the invariant region under a relaxed CFL condition. The new stability technique and conclusion work for general IRP property including the MMP property for scalar problem and the PP property for the Euler equations.

The structure of this paper is as follows. In Section 2, we review the general procedure of the unstaggered-central scheme. A new invariant-region-preserving limiter is defined and stability proof of the scheme is proved based on a forward-backward splitting method in Section 3. Section 4 and Section 5 are about applications of the present scheme to one-dimensional scalar equation and Euler equations, respectively. In Section 6, numerical tests are shown to demonstrate the robustness of the proposed scheme. The conclusion is given in Section 7.

2. The second-order MUSCL-type unstaggered-central scheme

In this section, we review the general procedure of the second-order MUSCL-type unstaggered-central scheme.

Given a uniform grid $x_\alpha := \alpha \Delta x$ of the domain, we define the primal cells $\mathcal{I}_i := \left[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}} \right]$ and the staggered dual cells $\mathcal{I}_{i+\frac{1}{2}} := [x_i, x_{i+1}]$. We also denote by $\chi_\alpha(x)$, the characteristic function of the cell $\mathcal{I}_\alpha$, i.e. $\chi_\alpha(x) := \mathbf{1}_{\mathcal{I}_\alpha}$. The finite volume schemes approximate the averages of solution in the primal cells

$$\mathbf{u}_i^n \approx \frac{1}{\Delta x} \int_{\mathcal{I}_i} \mathbf{u}(x, t^n) dx. \tag{3}$$

The scheme evolves the numerical solution on a single grid, which avoids the staggered spatial grids and the solution of the Riemann problems arising at the cell interfaces. The second-order unstaggered-central scheme includes three steps: forward projection, evolution and backward projection.

Forward projection. In this step, the solution on the primal cells is projected onto the "ghost" staggered dual cells. To enhance the spatial accuracy order, a piecewise-linear approximation (MUSCL-type) of the solution is constructed from the primal cell-average

$$\mathcal{L}_i(x; \mathbf{u}^n) = \mathbf{u}_i^n + \Delta \mathbf{u}_i^n \frac{x - x_i}{\Delta x}, \quad \forall x \in \mathcal{I}_i, \tag{4}$$

where $\Delta \mathbf{u}_i^n \approx \mathbf{u}(x_{i+\frac{1}{2}}, t^n) - \mathbf{u}(x_{i-\frac{1}{2}}, t^n)$ denotes the limited spatial increments of the solution in primal cells $\mathcal{I}_i$. It is done by a slope limiter method to ensure a non-oscillatory nature and second-order accuracy. The initial data reconstruction will be detailed in the next section. For the convenience of notation, we denote the sampled point values by

$$\mathbf{u}_i^{n, \vartheta} = \mathcal{L}_i(x_i + \vartheta \Delta x; \mathbf{u}^n) = \mathbf{u}_i^n + \vartheta \Delta \mathbf{u}_i^n \quad \text{for all } -\frac{1}{2} \leq \vartheta \leq \frac{1}{2}. \tag{5}$$

We aim to calculate the staggered cell-average of the grid function (4) in this step:

$$\mathbf{u}_{i+\frac{1}{2}}^n = \frac{1}{\Delta x} \int_{\mathcal{I}_{i+\frac{1}{2}}} \sum_j \mathcal{L}_j(x; \mathbf{u}^n) \chi_j(x) dx = \frac{1}{\Delta x} \sum_j \int_{\mathcal{I}_{i+\frac{1}{2}} \cap \mathcal{I}_j} \mathcal{L}_j(x; \mathbf{u}^n) dx. \tag{6}$$

Only two neighbouring primal cells $\mathcal{I}_i$ and $\mathcal{I}_{i+1}$ have nonempty intersections with staggered cell $\mathcal{I}_{i+\frac{1}{2}}$,

$$\mathcal{I}_{i+\frac{1}{2}} \bigcap \mathcal{I}_j = \begin{cases} [x_i, x_{i+\frac{1}{2}}], & \text{if } j = i, \\ [x_{i+\frac{1}{2}}, x_{i+1}], & \text{if } j = i+1, \\ \phi, & \text{otherwise.} \end{cases} \tag{7}$$

Thus the integral in (6) can be exactly calculated

$$\sum_j \int_{\mathcal{I}_{i+\frac{1}{2}} \cap \mathcal{I}_j} \mathcal{L}_j(x; \mathbf{u}^n) dx = \int_{x_i}^{x_{i+\frac{1}{2}}} \mathcal{L}_i(x; \mathbf{u}^n) dx + \int_{x_{i+\frac{1}{2}}}^{x_{i+1}} \mathcal{L}_{i+1}(x; \mathbf{u}^n) dx$$
$$= \frac{\Delta x}{2} \mathcal{L}_i \left(x_{i+\frac{1}{4}}; \mathbf{u}^n \right) + \frac{\Delta x}{2} \mathcal{L}_{i+1} \left(x_{i+\frac{3}{4}}; \mathbf{u}^n \right) = \frac{\Delta x}{2} \left(\mathbf{u}_i^{n, \frac{1}{4}} + \mathbf{u}_{i+1}^{n, -\frac{1}{4}} \right),$$

which, divided by the cell length Δx, gives the staggered cell-average

$$\mathbf{u}_{i+\frac{1}{2}}^n = \frac{\mathbf{u}_i^{n, \frac{1}{4}} + \mathbf{u}_{i+1}^{n, -\frac{1}{4}}}{2}. \tag{8}$$

These conclude the forward projection.

Evolution. The second step is the key step of general central scheme whether the grids are staggered or not. The solution on the primal cells at current time level t^n is evolved onto the staggered dual cells at the new time level t^{n+1}. The procedure can be written in a two-step predictor-corrector method. Integrating (1) on the rectangle $\mathcal{I}_{i+\frac{1}{2}} \times [t^n, t^{n+1}]$ and substituting the forward projected value $\mathbf{u}_{i+\frac{1}{2}}^n$ by (8), the numerical solution $\mathbf{u}_{i+\frac{1}{2}}^{n+1}$ at time t^{n+1} in the staggered dual cells is obtained

$$\mathbf{u}_{i+\frac{1}{2}}^{n+1} = \mathbf{u}_{i+\frac{1}{2}}^n - \frac{1}{\Delta x} \int_{t^n}^{t^{n+1}} \left(\mathbf{f}(\mathbf{u}(x_{i+1}, \tau)) - \mathbf{f}(\mathbf{u}(x_i, \tau)) \right) d\tau. \tag{9}$$

The second-order accurate mid-point quadrature rule is applied on the above temporal integrals for the fluxes at the cell centers:

$$\int_{t^n}^{t^{n+1}} \mathbf{f}(\mathbf{u}(x_i, \tau)) d\tau \approx \Delta t \mathbf{f}(\mathbf{u}_i^{n+\frac{1}{2}}), \tag{10}$$

where $\mathbf{u}_i^{n+\frac{1}{2}}$ is the predicted solution at the primal cell centers $x = x_i$ and at time $t^{n+\frac{1}{2}} = t^n + \frac{1}{2} \Delta t$. The required midvalue is given by the Taylor expansion

$$\mathbf{u}_i^{n+\frac{1}{2}} \approx \mathbf{u}_i^n - \frac{\Delta t}{2} (\partial_x \mathbf{f})_i^n, \tag{11}$$

where $(\partial_x \mathbf{f})_i^n$ is the discrete derivative of the flux in the primal cell $\mathcal{I}_i$. To avoid the calculation of the Jacobian $\mathbf{f}_u'$, we proceed with a straightforward computation by difference quotient

$$(\partial_x \mathbf{f})_i^n \approx \frac{\mathbf{f}(\mathbf{u}_i^{n, \frac{1}{2}}) - \mathbf{f}(\mathbf{u}_i^{n, -\frac{1}{2}})}{\Delta x}, \tag{12}$$

in which the interpolated values $\mathbf{u}_i^{n, \pm\frac{1}{2}}$ at the two boundaries of primal cell $\mathcal{I}_i$ are obtained by (5). Substituting the above discrete derivative into (11) yields a simple and time-saving prediction step

$$\mathbf{u}_i^{n+\frac{1}{2}} = \mathbf{u}_i^n - \frac{\Delta t}{2\Delta x} \left(\mathbf{f}(\mathbf{u}_i^{n, \frac{1}{2}}) - \mathbf{f}(\mathbf{u}_i^{n, -\frac{1}{2}}) \right). \tag{13}$$
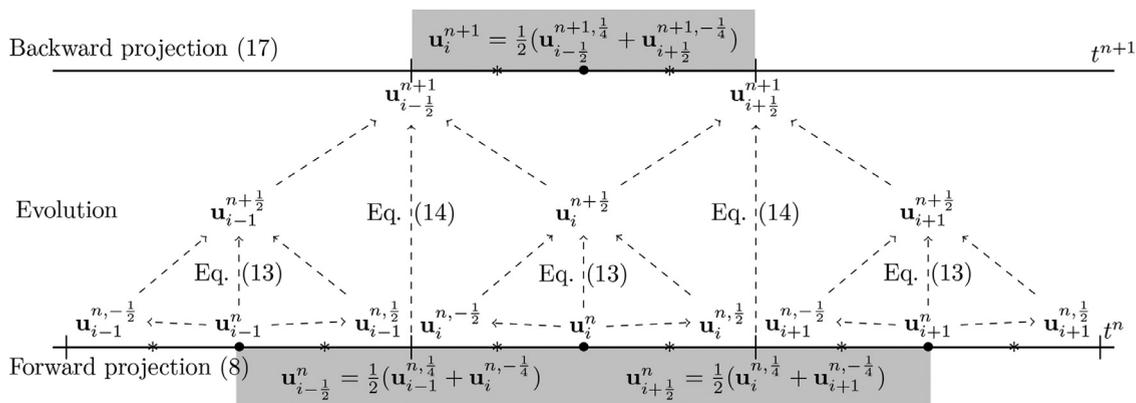
Fig. 1. The procedure of unstaggered-central scheme.

The integral of fluxes (10) is computed from the predicted value (13) and then substituted into (9), which completes the correction step

$$\mathbf{u}_{i+\frac{1}{2}}^{n+1} = \mathbf{u}_{i+\frac{1}{2}}^{n} - \frac{\Delta t}{\Delta x}\left(\mathbf{f}\left(\mathbf{u}_{i+1}^{n+\frac{1}{2}}\right) - \mathbf{f}\left(\mathbf{u}_i^{n+\frac{1}{2}}\right)\right). \tag{14}$$

Backward projection. Finally, the updated solution in the "ghost" staggered cells is projected back into the primal cells resulting an unstaggered formulation. Similar to the forward projection, we reconstruct a piecewise-linear function from the staggered cell-average at time t^{n+1}:

$$\mathcal{L}_{i+\frac{1}{2}}(x; \mathbf{u}^{n+1}) = \mathbf{u}_{i+\frac{1}{2}}^{n+1} + \Delta\mathbf{u}_{i+\frac{1}{2}}^{n+1}\frac{x - x_{i+\frac{1}{2}}}{\Delta x}, \quad \forall x \in \mathcal{I}_{i+\frac{1}{2}}, \tag{15}$$

where $\Delta\mathbf{u}_{i+\frac{1}{2}}^{n+1} \approx \mathbf{u}(x_{i+1}, t^{n+1}) - \mathbf{u}(x_i, t^{n+1})$ denotes the limited value which is the approximate spatial increment of the exact solution in the staggered dual cells. The sampled point values are denoted by

$$\mathbf{u}_{i+\frac{1}{2}}^{n+1,\vartheta} := \mathcal{L}_{i+\frac{1}{2}}(x_{i+\frac{1}{2}} + \vartheta\,\Delta x; \mathbf{u}^{n+1}) = \mathbf{u}_{i+\frac{1}{2}}^{n+1} + \vartheta\,\Delta\mathbf{u}_{i+\frac{1}{2}}^{n+1} \quad \text{for all} \quad -\frac{1}{2} \le \vartheta \le \frac{1}{2}. \tag{16}$$

The projected values, i.e. the updated values in the primal cells, are obtained after integrating the reconstructed solution (15) in the primal cells. Therefore, the primal cell-average is shown below

$$\mathbf{u}_i^{n+1} = \frac{\mathbf{u}_{i-\frac{1}{2}}^{n+1,\frac{1}{4}} + \mathbf{u}_{i+\frac{1}{2}}^{n+1,-\frac{1}{4}}}{2}. \tag{17}$$

We summarize the description of the algorithm as shown in Fig. 1. The forward and backward projection steps are fulfilled by (8) and (17), respectively. The two-step predictor-corrector formulas for the evolution are done by (13) and (14), respectively. The spatial accuracy is enhanced by the piecewise-linear initial data reconstruction. The temporal accuracy is achieved by the midpoint integral rule (10) of the flux. As a result, the scheme has second-order accuracy. In the next section, we detail the initial data reconstruction and give the stability analysis.

3. Stability analysis

This paper only focuses on the stability of the second-order unstaggered-central scheme. To make the analysis valid for general IRP property, we introduce an extended IRP property of the slope limiter for the initial data reconstruction. The second novelty is the introduction of a convex splitting of the scheme, in which the convex splitting is a combination of some LxF schemes. We call it the forward-backward decomposition. Finally, the stability result is proved under a suitable CFL condition which is larger than that used in [21,23,24]. The stability result works for general hyperbolic conservation laws.

3.1. IRP *initial data reconstruction*

In [23], a clipping limiter feature is the key to retain the MMP property at the reconstruction step for the scalar problem. The limiter will reduce the scheme back to a first-order scheme in an extreme cell because the limiter will set zero slope in this cell [29]. On the other hand, the clipping condition can not be extended to nonlinear systems. We will relax the clipping requisition of the slope limiter.

In the upwind-type scheme including WENO and DG, et. al., the slope limiters were proposed to preserve the MMP property for the scalar problem [9] and the PP property for the Euler equations [10]. Generally, these demands that the reconstruction must be IRP inside every cell.

Definition 3.1. A reconstruction is said IRP, if

$$\mathbf{u}_\alpha^n \in \Omega \text{ for all } \alpha \Longrightarrow \mathbf{u}_\alpha^{n,\pm\frac{1}{2}} \in \Omega \text{ for all } \alpha \tag{18}$$

with $\alpha = i$(resp. $i + \frac{1}{2}$) for primal cells (resp. staggered dual cells).

Due to the convexity of the domain Ω, an equivalent definition of IRP reconstruction is that any sampled state inside the whole cell belongs to the invariant set:

$$\mathbf{u}_\alpha^{n,\pm\vartheta} = \mathcal{L}_\alpha(x_\alpha \pm \vartheta \Delta x; \mathbf{u}^n) \in \Omega, \quad \forall \vartheta \in [0, \frac{1}{2}]. \tag{19}$$

Specially, the quantities $\mathbf{u}_\alpha^{n,\pm\frac{1}{4}} \in \Omega$ used in forward and backward projection steps (8) and (17), respectively. Actually, they can be viewed as the average of the cell average value and cell boundary value

$$\mathbf{u}_\alpha^{n,\pm\frac{1}{4}} = \frac{1}{2}\mathbf{u}_\alpha^n + \frac{1}{2}\mathbf{u}_\alpha^{n,\pm\frac{1}{2}}. \tag{20}$$

In order to improve the stability of the scheme, instead of the clipping condition we ask for a little bit more restriction on the initial data reconstruction by extending the interpolation points outside to the neighbouring cells in the forward projection step.

Definition 3.2. A reconstruction is said extended IRP, if there exists a parameter $\vartheta \geq \frac{1}{2}$

$$\mathbf{u}_i^n \in \Omega \text{ for all } i \Longrightarrow \mathbf{u}_i^{n,\pm\vartheta} \in \Omega \text{ for all } i. \tag{21}$$

Due to the convexity of the Ω and the linearity of the reconstructed conservative variables, we immediately get that (21) implies (18) because $\mathbf{u}_i^{n,\pm\frac{1}{2}}$ can be viewed as a convex combination of $\mathbf{u}_i^n$ and $\mathbf{u}_i^{n,\pm\vartheta}$ where $\vartheta \geq \frac{1}{2}$. It means that an extended IRP reconstruction must be IRP.

By replacing the cell boundary state in (20) with the extrapolated quantity (21), the quantities used in the forward projection step (8) can be written as a new convex combination:

$$\mathbf{u}_i^{n,\pm\frac{1}{4}} = \frac{1}{4\vartheta}\mathbf{u}_i^{n,\pm\vartheta} + \left(1 - \frac{1}{4\vartheta}\right)\mathbf{u}_i^n. \tag{22}$$

The implementations of the IRP limiter will be realized in Section 4 for scalar problem and in Section 5 for the Euler systems, respectively.

3.2. Solution's convex decomposition

The clipping condition for the slope limiter has been relaxed to the extended IRP, and then the technique of proof in [23] cannot be used here. There is no difficulty in proving the stability of the predicted solution at the half time step in the predictor step (13). The main difficulty is proving the IRP property of the updated solution in the corrector step (14). In present part, we will introduce a convex decomposition of the updated solution.

The Lax-Friedrichs scheme [40] is the prototype of the central scheme and many of its well-known stability properties are proved, including IRP principle, the TVD property, and the discrete entropy inequality. We will split the updated solution to be a convex summation of the reconstructed states at the initial time and some other states by the LxF scheme which takes the form

$$\mathcal{H}(\mathbf{u}, \mathbf{v}, a) := \frac{\mathbf{u} + \mathbf{v}}{2} - a\Big(\mathbf{f}(\mathbf{v}) - \mathbf{f}(\mathbf{u})\Big), \quad \forall \mathbf{u}, \mathbf{v} \in \Omega, a \in R. \tag{23}$$

The following lemma gives the decomposition.

Lemma 3.3. *(Forward-backward splitting) The updated solution by (14) in the staggered dual cell at the time level t^{n+1} is equivalent to the following combination*

$$\begin{aligned}
\boldsymbol{u}_{i+\frac{1}{2}}^{n+1} = &\frac{1}{8\vartheta}\Big(\boldsymbol{u}_i^{n,\vartheta} + \boldsymbol{u}_{i+1}^{n,-\vartheta}\Big) \\
&+\kappa\boldsymbol{u}_{i+\frac{1}{2}}^{n+1,(\wedge)} + \frac{1}{2}\Big(1 - \frac{1}{4\vartheta} - \kappa\Big)\Big(\boldsymbol{u}_{i+\frac{1}{2}}^{n+1,\left(-\frac{1}{4},\vee\right)} + \boldsymbol{u}_{i+\frac{1}{2}}^{n+1,\left(\frac{1}{4},\vee\right)}\Big),
\end{aligned} \tag{24}$$

where $\mathbf{u}_i^{n,\pm\vartheta}$ are sampled quantities in (21) and

$$\boldsymbol{u}_{i+\frac{1}{2}}^{n+1,(\wedge)} = \mathcal{H}\left(\mathbf{u}_i^{n+\frac{1}{2}}, \mathbf{u}_{i+1}^{n+\frac{1}{2}}, \frac{\Delta t}{\kappa \Delta x}\right), \tag{25}$$

$$\mathbf{u}_{i+\frac{1}{2}}^{n+1,(\pm\frac{1}{4},\vee)} = \mathcal{H}\left(\mathbf{u}_{i+\frac{1}{2}\pm\frac{1}{2}}^{n,\frac{1}{2}}, \mathbf{u}_{i+\frac{1}{2}\pm\frac{1}{2}}^{n,-\frac{1}{2}}, \frac{\frac{\kappa}{2}}{1-\frac{1}{4\vartheta}-\kappa}\frac{\Delta t}{\Delta x}\right) \tag{26}$$

for any given parameter κ.

Proof. Using (22), the summation of two sampled quantities in (24) takes the form

$$\mathbf{u}_i^{n,\vartheta} + \mathbf{u}_{i+1}^{n,-\vartheta} = 4\vartheta(\mathbf{u}_i^{n,\frac{1}{4}} + \mathbf{u}_{i+1}^{n,-\frac{1}{4}}) - (4\vartheta-1)(\mathbf{u}_i^n + \mathbf{u}_{i+1}^n). \tag{27}$$

Applying (23) to (25), in which $\frac{\Delta t}{\kappa\Delta x}$ is in place of a, yields

$$\mathbf{u}_{i+\frac{1}{2}}^{n+1,(\wedge)} = \frac{\mathbf{u}_i^{n+\frac{1}{2}} + \mathbf{u}_{i+1}^{n+\frac{1}{2}}}{2} - \frac{\Delta t}{\kappa\Delta x}\left(\mathbf{f}(\mathbf{u}_{i+1}^{n+\frac{1}{2}}) - \mathbf{f}(\mathbf{u}_i^{n+\frac{1}{2}})\right). \tag{28}$$

Substituting the predicted values $\mathbf{u}_i^{n+\frac{1}{2}}$ and $\mathbf{u}_{i+1}^{n+\frac{1}{2}}$ by (13) into the first two terms in the right hand side of above formula, we obtain

$$\mathbf{u}_{i+\frac{1}{2}}^{n+1,(\wedge)} = \frac{\mathbf{u}_i^n + \mathbf{u}_{i+1}^n}{2} - \frac{\Delta t}{4\Delta x}\sum_{l=i}^{i+1}\left(\mathbf{f}(\mathbf{u}_l^{n,\frac{1}{2}}) - \mathbf{f}(\mathbf{u}_l^{n,-\frac{1}{2}})\right) - \frac{\Delta t}{\kappa\Delta x}\left(\mathbf{f}(\mathbf{u}_{i+1}^{n+\frac{1}{2}}) - \mathbf{f}(\mathbf{u}_i^{n+\frac{1}{2}})\right). \tag{29}$$

Adding the two equations in (26) and applying (23) with $\frac{\frac{\kappa}{2}}{1-\frac{1}{4\vartheta}-\kappa}\frac{\Delta t}{\Delta x}$ in place of a, it is derived that

$$\mathbf{u}_{i+\frac{1}{2}}^{n+1,(-\frac{1}{4},\vee)} + \mathbf{u}_{i+\frac{1}{2}}^{n+1,(\frac{1}{4},\vee)} = \mathbf{u}_i^n + \mathbf{u}_{i+1}^n - \frac{\frac{\kappa}{2}}{1-\frac{1}{4\vartheta}-\kappa}\frac{\Delta t}{\Delta x}\sum_{l=i}^{i+1}\left(\mathbf{f}(\mathbf{u}_l^{n,-\frac{1}{2}}) - \mathbf{f}(\mathbf{u}_l^{n,\frac{1}{2}})\right). \tag{30}$$

Inserting terms (27), (29) and (30) into the right hand side of (24), we obtain updated solution (14), i.e. (24) holds.
These conclude the proof. $\square$

Remark 3.4. We can check that the summation (24) is convex if the parameter κ to be defined satisfies the following inequality

$$0 \leq \kappa \leq 1 - \frac{1}{4\vartheta}. \tag{31}$$

Remark 3.5. Formula (25) can be viewed as the solution of the forward process of the original PDE (1): $\mathbf{u}_t + \mathbf{f}(\mathbf{u})_x = 0$, while (26) can also be seen as solving the backward process: $\mathbf{u}_t - \mathbf{f}(\mathbf{u})_x = 0$. As a consequence, we call (24) the Forward-Backward Splitting.

3.3. Main results

In this part, we will give main stability results. Using the proposed IRP limiter and the forward-backward splitting, we can simplify the proof of the stability. The stability condition will enlarge the CFL number which admits larger time step. On the other hand, our stability results are suitable for general IRP property of the hyperbolic conservation laws including nonlinear systems.

To verify the stability of the convex combination (24) of updated solution, we first need to know the stability property of the LxF scheme (23).

Lemma 3.6. Given the flux function $\mathbf{f}(\mathbf{u})$ and two states $\mathbf{u}, \mathbf{v}$ in Ω. If there is a positive parameter $a > 0$ satisfying

$$a \cdot \lambda(\mathbf{u}, \mathbf{v}) \leq \frac{1}{2} \quad \text{with} \quad \lambda(\mathbf{u}, \mathbf{v}) = \max_{0 \leq \theta \leq 1}\rho(\mathbf{f}'(\mathbf{u} + \theta(\mathbf{v} - \mathbf{u}))), \tag{32}$$

where $\rho(\mathbf{f}'(\mathbf{w}))$ is the spectral radius of the Jacobian $\mathbf{f}'(\mathbf{w})$ for any $\mathbf{w} \in \Omega$. Then $\mathcal{H}(\mathbf{u}, \mathbf{v}, a) \in \Omega$.

In the case of the genuinely nonlinear or linearly degeneration, these local maximal speeds can be easily evaluated as

$$\lambda(\mathbf{u}, \mathbf{v}) = \max\left(\rho(\mathbf{f}'(\mathbf{u})), \rho(\mathbf{f}'(\mathbf{v}))\right). \tag{33}$$

The proof of the Lemma 3.6 is omitted here and is referred to [40].
The following theorem is the main stability result which gives a sufficient condition for IRP property.

Theorem 3.7. Given an extended IRP initial data reconstruction with the parameter $\vartheta \geq \frac{1}{2}$. Under the assumption that

$$\frac{\Delta t}{\Delta x}\max_i\left(\lambda_i^n, \lambda_{i+\frac{1}{2}}^{n+\frac{1}{2}}\right) \leq \text{CFL} = \frac{\sqrt{3-\frac{1}{2\vartheta}}-1}{2}, \tag{34}$$

with $\lambda_i^n := \lambda(\mathbf{u}_i^{n,-\frac{1}{2}}, \mathbf{u}_i^{n,\frac{1}{2}})$ and $\lambda_{i+\frac{1}{2}}^{n+\frac{1}{2}} := \lambda(\mathbf{u}_i^{n+\frac{1}{2}}, \mathbf{u}_{i+1}^{n+\frac{1}{2}})$, then the scheme is IRP: $u_i^{n+1} \in \Omega$.

Proof. The IRP reconstruction enforces $\mathbf{u}_i^{n,\pm\frac{1}{4}} \in \Omega$ and then $\mathbf{u}_{i+\frac{1}{2}}^n \in \Omega$, which is on account of the convex combination of (8). We assert that

$$\mathbf{u}_{i+\frac{1}{2}}^{n+1} \in \Omega. \tag{35}$$

Similarly, we have $\mathbf{u}_{i+\frac{1}{2}}^{n+1,\pm\frac{1}{4}} \in \Omega$ and $\mathbf{u}_{i+\frac{1}{2}}^{n+1} \in \Omega$ for the backward projection step. Thus the final conclusion is valid.

We next focus on the discussion of (35) in the evolution step, which is done by a two-step predictor-corrector method.

We first consider the predicted value by (13) at the half time step. By the linearity of the reconstructed solution (4), we can rewrite the predicted value by (13) in a form of (23) with $\frac{\Delta t}{2\Delta x}$ in place of a

$$\mathbf{u}_i^{n+\frac{1}{2}} = \mathcal{H}\left(\mathbf{u}_i^{n,-\frac{1}{2}}, \mathbf{u}_i^{n,\frac{1}{2}}, \frac{\Delta t}{2\Delta x}\right). \tag{36}$$

By the Lemma 3.6, we have $\mathbf{u}_i^{n+\frac{1}{2}} \in \Omega$ under the condition

$$\frac{\Delta t}{2\Delta x} \max_i \lambda_i^n < \frac{1}{2}, \tag{37}$$

which is ensured by (34).

Secondly, we consider the corrector step by (14) at full time level t^{n+1}. By Lemma 3.3, the corrected value has the forward-backward splitting form (24). The IRP reconstruction enforces $\mathbf{u}_i^{n,\pm\vartheta} \in \Omega$. It is easy to check that $\kappa = 2 \cdot \text{CFL}$, where CFL defined in (34) is a root of the equation

$$\frac{\kappa}{2} = \frac{1 - \frac{1}{4\vartheta} - \kappa}{\kappa}. \tag{38}$$

Thus the condition (34) implies that

$$\begin{aligned} \frac{\Delta t}{\kappa \Delta x} \max_i \lambda_{i+\frac{1}{2}}^{n+\frac{1}{2}} &\le \frac{\text{CFL}}{\kappa} = \frac{1}{2}, \\ \frac{\frac{\kappa}{2}}{1 - \frac{1}{4\vartheta} - \kappa} \frac{\Delta t}{\Delta x} \max_i \lambda_i^n &\le \frac{\frac{\kappa}{2}}{1 - \frac{1}{4\vartheta} - \kappa} \text{CFL} = \frac{1}{2}. \end{aligned} \tag{39}$$

Applying Lemma 3.6 on (25) (resp. (26)) can ensure $\mathbf{u}_{i+\frac{1}{2}}^{n+1,(\wedge)} \in \Omega$ (resp. $\mathbf{u}_{i+\frac{1}{2}}^{n+1,(\pm\frac{1}{4},\vee)} \in \Omega$ where we have used the fact that $\lambda(\mathbf{u}, \mathbf{v}) = \lambda(\mathbf{v}, \mathbf{u})$ for any two vectors $\mathbf{u}$ and $\mathbf{v}$ in Ω). On the other hand, all coefficients in (24) are positive by the choice of κ and their summation equals one. Consequently, $\mathbf{u}_{i+\frac{1}{2}}^{n+1} \in \Omega$ since it is a convex combination of some quantities belonging to Ω by (24).

These conclude the proof. □

Remark 3.8. It can be seen that the time step calculated by (34) is an implicit value, because it depends not only on the initial data $\mathbf{u}_i^{n,\pm\frac{1}{2}}$, but also on the predicted value $\mathbf{u}_{i+\frac{1}{2}\pm\frac{1}{2}}^{n+\frac{1}{2}}$. In the real application, we can choose varying time steps. In the predictor step, we compute the solution $\mathbf{u}_i^{n+\frac{1}{2}} = \mathbf{u}_i^n - \frac{\widehat{\Delta t}}{2}(\partial_x \mathbf{f})_i^n$ by (13) with time step $\widehat{\Delta t}$ in place of Δt. Here $\widehat{\Delta t}$ is obtained by

$$\widehat{\Delta t} = \text{CFL} \frac{\Delta x}{\max_i \lambda_i^n}, \tag{40}$$

using the data $\mathbf{u}_i^{n,\pm\frac{1}{2}}$ at the initial time $t = t^n$. While the corrected solution $\mathbf{u}_{i+\frac{1}{2}}^{n+1}$ is computed by (14) with time step Δt, which is obtained by

$$\Delta t = \text{CFL} \frac{\Delta x}{\max_i \left(\lambda_i^n, \lambda_{i+\frac{1}{2}}^{n+\frac{1}{2}}\right)} \tag{41}$$

using the data $\mathbf{u}_i^{n,\pm\frac{1}{2}}$ at the initial time $t = t^n$ and $\mathbf{u}_{i+\frac{1}{2}\pm\frac{1}{2}}^{n+\frac{1}{2}}$ at the half time step $t = t^n + \frac{1}{2}\widehat{\Delta t}$.

It is clear to see that $\Delta t \le \widehat{\Delta t}$. Thus the IRP principle stated in Theorem 3.7 still holds. On the other hand, $\widehat{\Delta t} = \Delta t + \mathcal{O}(\Delta t^2)$ is established because of $\mathbf{u}_i^{n+\frac{1}{2}} = \mathbf{u}_i^n + \mathcal{O}(\Delta t)$. Thus the error of the numerical integral in $[t^n, t^n + \Delta t]$ by (10) is enforced to be $\mathcal{O}(\Delta t^3)$, even though $\mathbf{u}_i^{n+\frac{1}{2}}$ is sampled at $t^n + \frac{1}{2}\widehat{\Delta t}$ instead of at $t^n + \frac{1}{2}\Delta t$. Accordingly, the numerical accuracy order of the scheme is maintained. Another way to realize the stable time discretization with different time steps can be found in [41].
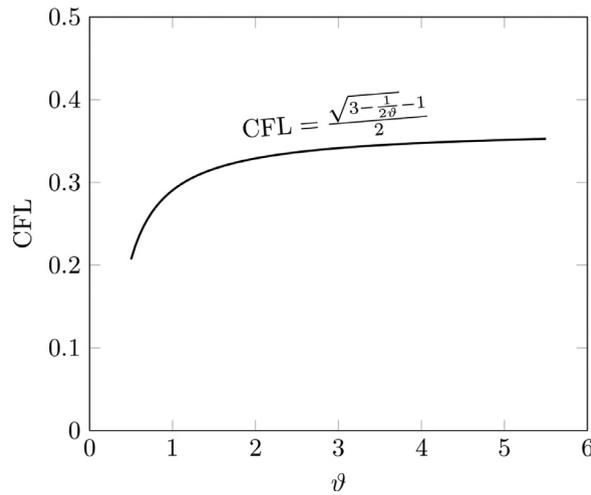
Fig. 2. CFL numbers as a function of ϑ.

Remark 3.9. The CFL conditions are sufficient conditions for IRP principle. We plot the CFL number as a function of ϑ in Fig. 2. Actually, the CFL number is increasing as a function of ϑ. As discussed before, IRP reconstruction is also extended IRP if we set $\vartheta = \frac{1}{2}$ which activates $\text{CFL}_{\vartheta=\frac{1}{2}} = \frac{\sqrt{2}-1}{2} \approx 0.207$. If we set larger ϑ, then we can get larger CFL number and larger time step. The possible largest value is $\text{CFL}_{\vartheta\to\infty} = \frac{\sqrt{3}-1}{2} \approx 0.366$. But if $\vartheta > \frac{1}{2}$, we need more serious restriction on the reconstructed slopes obtained in the forward projection step. In the real application, we set $\vartheta = \frac{3}{2}$ resulting $\text{CFL}_{\vartheta=\frac{3}{2}} \approx 0.316$. The extended IRP limiters should be given for different systems.

Remark 3.10. The proof by Nessyahu and Tadmor in [21] is only for scalar problem. The basis is that the reconstructed conservative variables and fluxes are clipping, i.e. the neighboring discrete slopes can not have opposite signs. The proof technique is the mean-value theorem. Both their clipping reconstruction and proof technique are not suitable for nonlinear systems. In this paper, the extended IRP reconstruction is more practical, and the forward-backward splitting method is more adaptable. We will apply our method to scalar problem and Euler equations respectively in the upcoming sections.

4. Application to scalar problems

In this section, we apply our method to the scalar problem, i.e. $p = 1$ in (1) which becomes

$$u_t + f(u)_x = 0, \quad u(x, t = 0) = u_0(x). \tag{42}$$

In this case, the invariant domain is an interval

$$\Omega = [m, M], \tag{43}$$

where $m = \min_x u_0(x)$ and $M = \max_x u_0(x)$ are the minimum and maximum values of the solution at the initial time, respectively. The IRP principle becomes *minimum-maximum-preserving*(MMP) principle.

To ensure non-oscillatory nature of the reconstruction and so as to avoid spurious oscillations in the numerical solution, one has to use a nonlinear limiter on the reconstructed piecewise-linear function in the forward and backward projection steps.

In [23], a clipping condition of the initial data reconstruction is the key to prove the MMP property for the scalar problem. This condition demands that the reconstructed slopes of the solution in two neighbouring cells have the same sign. It is done by the generalized minmod method. Given cell-averages u_α with $\alpha = i$(resp. $\alpha = i + \frac{1}{2}$) on primal cells(resp. staggered dual cells), the cell increments are estimated by

$$\Delta u_\alpha = \text{minmod}\left(\theta(u_\alpha - u_{\alpha-1}), \frac{u_{\alpha+1} - u_{\alpha-1}}{2}, \theta(u_{\alpha+1} - u_\alpha)\right), \tag{44}$$

which is a three components function with a parameter θ. The minmod function is defined as:

$$\text{minmod}(z_1, z_2, \dots) := \begin{cases} \min_j\{z_j\}, & \text{if} \quad z_j > 0 \quad \forall j, \\ \max_j\{z_j\}, & \text{if} \quad z_j < 0 \quad \forall j, \\ 0, & \text{otherwise}. \end{cases} \tag{45}$$

The parameter $\theta \in [1, 2]$ affects the numerical viscosity of the scheme. If $\theta = 1$, it degenerates to the original minmod limiter which results in large viscosity. More robust choice is $\theta = 1.3$ for many applications.

It can be checked that the generalized minmod method satisfies the extended IRP property by setting $\vartheta = \frac{1}{\theta}$. Thus the CFL in (34) will become

$$\text{CFL}_\theta = \frac{\sqrt{3 - \frac{1}{2}\theta} - 1}{2}. \tag{46}$$

This generalized minmod limiter (44) will reduces the scheme back to a first-order scheme in an extreme cell because the limiter will set zero slope in this cell [29]. On the other hand, the clipping condition can not be extended to nonlinear systems. The more serious problem is applying the method to solve the non-convex problem resulting nonphysical solution near the joint point of a rarefaction wave and a shock wave. Thus an adaptive trick is added in the vicinity of the problem points to switch back to the parameter $\theta = 1$ [27].

Pioneered by the IRP property (21), we realize the reconstruction in following two steps. The preliminary step is estimating the first-order accurate nonoscillatory slope in every cell. The second step is adjusting the preliminary reconstructed slope to enforce the MMP property. Without the clipping restriction we can realize the full second-order interpolation near the extreme point.

- *Preliminary step*: We estimate the cell increments using the uniformly high-order nonoscillatory (UNO) limiter [24,42]

$$\Delta^\# u_\alpha := \quad \text{minmod}\Big(\Delta_- u_\alpha + \tfrac{1}{2}\text{minmod}\big(\Delta_+ \Delta_- u_{\alpha-1}, \Delta_+ \Delta_- u_\alpha\big), \\ \Delta_+ u_\alpha - \tfrac{1}{2}\text{minmod}\big(\Delta_+ \Delta_- u_\alpha, \Delta_+ \Delta_- u_{\alpha+1}\big)\Big), \tag{47}$$

in which

$$\Delta_- u_\alpha = u_\alpha - u_{\alpha-1}, \quad \Delta_+ u_\alpha = u_{\alpha+1} - u_\alpha.$$

The UNO limiter is the only limiter which is uniformly second-order accurate. The penalty of using the UNO limiter is that it has a wider stencil(five points for computing numerical derivatives). And this limiter can give the correct solution for the nonconvex problems [24].

- *Adjustment step*: We next adjust the preliminary reconstructed slope to satisfy the extended IRP property (21) without violating the nonoscillation property and accuracy order of the interpolation. It can be done by setting

$$\Delta u_\alpha = \text{sgn}(\Delta^\# u_\alpha) \min\left(|\Delta^\# u_\alpha|, \frac{1}{\vartheta}(M - u_\alpha), \frac{1}{\vartheta}(u_\alpha - m)\right), \tag{48}$$

where $\vartheta \geq \frac{1}{2}$.

The following properties of the adjustment can be checked.

  - The MMP property $m \leq u_\alpha \pm \vartheta \Delta u_\alpha \leq M$ is a direct result of the inequality

$$\frac{1}{\vartheta}\max(m - u_\alpha, u_\alpha - M) \leq \Delta u_\alpha \leq \frac{1}{\vartheta}\min(M - u_\alpha, u_\alpha - m), \tag{49}$$

  which is a conclusion of (48).
  - The nonoscillation property is preserved because $|\Delta u_\alpha| \leq |\Delta^\# u_\alpha|$ and they have same sign.
  - The accuracy order of the limiter is preserved. The proof is referred to the paper [9].

Remark 4.1. The reconstruction method in the preliminary step is not limited to the generalized minmod and UNO. We can use any other essential nonoscillatory limiters, e.g. van-Leer's method [43], ENO.

Remark 4.2. Nessyahu and Tadmor [21] use the generalized minmod limiter (44) to give the initial data reconstruction due to its clipping property. They give the stability condition

$$\text{CFL}_\theta^{\text{NT}} = \frac{\sqrt{4 + 4\theta - \theta^2} - 2}{2\theta}, \tag{50}$$

which is related to the parameter θ in generalized minmod limiter (44).

We plot the different CFL numbers (50), (46) and (34) with $\vartheta = \frac{3}{2}$ in Fig. 3 for comparison. It can be seen that (50) is more sharper than (46) for the stable CFL estimation. The reason is that the proof in [21] is taken full consideration of the clipping condition. But clipping condition is not suitable for nonlinear systems. Actually our new estimation (46) is very close to (50). An important thing is that our estimation is suitable for general systems. Another issue is that we can choose any essential nonoscillation reconstruction for the preliminary reconstruction and then we can use the scalar ϑ in the adjustment step to enforce the MMP property. It gives us a new freedom to enlarge the CFL number. It can be calculated that $\text{CFL}_{\theta=1.048}^{\text{NT}} \approx \text{CFL}_{\vartheta=\frac{3}{2}} = 0.316$ which means that $\text{CFL}_{\vartheta=\frac{3}{2}} > \text{CFL}_\theta^{\text{NT}}$ when $\theta \gtrsim 1.048$.

5. Application to nonlinear systems

In this section, we apply our method to solve the nonlinear Euler equations for the inviscid gas dynamics

$$\partial_t \begin{pmatrix} \rho \\ \rho u \\ E \end{pmatrix} + \partial_x \begin{pmatrix} \rho u \\ \rho u^2 + p \\ u(E + p) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \tag{51}$$
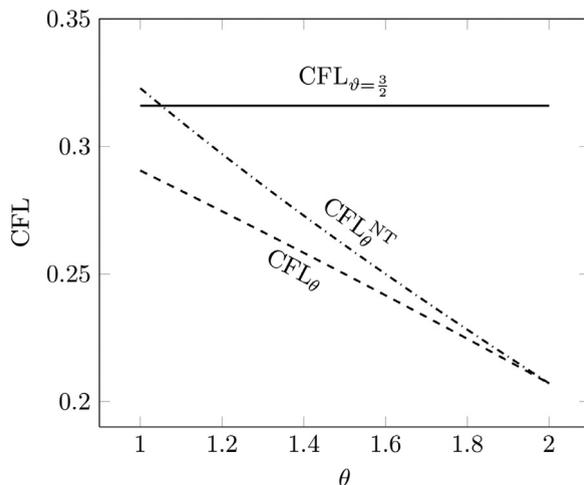
Fig. 3. Comparison of CFL numbers by different methods.

where ρ is the density of gas, u denotes the velocity, p represents the pressure, $E = \frac{1}{2}\rho u^2 + \frac{p}{\gamma-1}$ describes the energy. The eigenvalues are $\lambda_\pm = u \pm c$ with the sound speed $c = \sqrt{\gamma p/\rho}$. So the maximum absolute eigenvalue is $\lambda = |u| + c$. In this case, the invariant domain is

$$\Omega = \left\{ \begin{pmatrix} \rho \\ \rho u \\ E \end{pmatrix} \middle| \rho > 0, p > 0 \right\}, \tag{52}$$

which means that both density and pressure should be positive. The IRP principle becomes the *positivity-preserving*(PP) principle.

Similarly to the scalar case, we do the reconstruction in following two steps. The preliminary step is estimating the first-order accurate nonoscillatory slope in every cell. The second step is adjusting the preliminary reconstructed slope to enforce the PP property.

- *Preliminary step*: Because discontinuities may emerge in the solution even with continuous initial data, we should use a nonlinear characteristic limiter for the initial data reconstruction to avoid spurious oscillations in the numerical solution. We apply the uniformly high-order non-oscillatory (UNO) limiter (47) in each characteristic field. For the sake of notation, we use $\Delta^\# \rho_\alpha$, $\Delta^\#(\rho u)_\alpha$ and $\Delta^\# E_\alpha$ to denote the reconstructed cell increments of all conservative variables. Here $\alpha = i$ stands for the reconstruction in the primal cells $\mathcal{I}_i$ in the forward projection step, while $\alpha = i + \frac{1}{2}$ stands for the reconstruction in the staggered dual cells $\mathcal{I}_{i+\frac{1}{2}}$ in the backward projection step.

- • *Adjustment step*: The reconstructed nonoscillatory cell increments does not ensure the positivity, i.e. the PP principle is destroyed, we need to modify them to obtain a second-order PP and extended PP reconstructions. In present paper, we adopt the method by Berthon [41]. For generality, we find a sufficient condition to realize positivity:

$$\rho_\alpha \pm \vartheta \Delta \rho_\alpha > 0, \tag{53}$$

$$E_\alpha \pm \vartheta \Delta E_\alpha - \frac{\left((\rho u)_\alpha \pm \vartheta \Delta(\rho u)_\alpha\right)^2}{2(\rho_\alpha \pm \vartheta \Delta \rho_\alpha)} > 0, \tag{54}$$

for any given constant $\vartheta \geq \frac{1}{2}$, where (53) is for the positivity of reconstructed density, and (54) is for the positivity of reconstructed pressure. Following [41], the sum of two equations in (54) can be written as

$$2E_\alpha - \frac{1}{2}\left(\frac{\left((\rho u)_\alpha - \vartheta \Delta(\rho u)_\alpha\right)^2}{\rho_\alpha - \vartheta \Delta \rho_\alpha} + \frac{\left((\rho u)_\alpha + \vartheta \Delta(\rho u)_\alpha\right)^2}{\rho_\alpha + \vartheta \Delta \rho_\alpha}\right) > 0. \tag{55}$$

To find $\Delta \rho_\alpha$ and $\Delta(\rho u)_\alpha$ satisfying the above inequality, we propose to consider $\Delta \rho_\alpha$ such that

$$2E_\alpha^n - \frac{1}{2}\left(\frac{(\rho u)_\alpha^2}{\rho_\alpha - \vartheta \Delta \rho_\alpha} + \frac{(\rho u)_\alpha^2}{\rho_\alpha + \vartheta \Delta \rho_\alpha}\right) > 0, \tag{56}$$

which deduces that

$$\vartheta |\Delta \rho_\alpha| < \sqrt{\frac{E_\alpha - \frac{1}{2}\rho_\alpha(u_\alpha)^2}{E_\alpha}} < 1. \tag{57}$$

Table 1
Example 1: linear advection problem (60) with continuous initial condition
(62). Numerical errors in different norms and the corresponding experimental
orders of convergence (EOC) at time $t = 10$. "# cells" stands for the number of
cells.

# cells	L_1 error	EOC	L_2 error	EOC	L_∞ error	EOC
25	1.039E-1	-	8.091E-2	-	7.824E-2	-
50	2.090E-2	2.313	1.565E-2	2.371	1.543E-2	2.342
100	4.933E-3	2.083	3.841E-3	2.026	4.411E-3	1.806
200	1.229E-3	2.006	9.333E-4	2.041	1.146E-3	1.944
400	2.911E-4	2.077	2.219E-4	2.072	2.926E-4	1.969
800	6.923E-5	2.072	5.321E-5	2.060	7.451E-5	1.974

The inequality in (57) is used to modify the cell increment of density which is consistent with the PP condition. Next, with the fixed $\Delta \rho_\alpha$, the inequality (55) gives a relevant choice of $\Delta(\rho u)_\alpha$. Finally, the increment ΔE_α is considered to satisfy (54).

Based on the above discussion, we modify the reconstructed gradients of conservative variables as follows

$$
\begin{cases}
\Delta \rho_\alpha = \text{sgn}(\Delta^\# \rho_\alpha) \min \left(\frac{1}{\vartheta} \rho_\alpha \sqrt{\frac{E_\alpha - \frac{1}{2} \rho_\alpha (u_\alpha)^2}{E_\alpha}}, |\Delta^\# \rho_\alpha| \right), \\
\Delta (\rho u)_\alpha = \max \left(\xi^-, \min \left(\xi^+, \Delta^\# (\rho u)_\alpha \right) \right), \\
\Delta E_\alpha = \frac{1}{\vartheta} \max \left(-E_\alpha + \frac{((\rho u)_\alpha + \vartheta \Delta (\rho u)_\alpha)^2}{2(\rho_\alpha + \vartheta \Delta \rho_\alpha)}, \min \left(E_\alpha - \frac{((\rho u)_\alpha - \vartheta \Delta (\rho u)_\alpha)^2}{2(\rho_\alpha - \vartheta \Delta \rho_\alpha)}, \vartheta \Delta^\# E_\alpha \right) \right),
\end{cases}
\tag{58}
$$

where $\xi^\pm$ are defined as follows

$$
\xi^\pm = u_\alpha \Delta \rho_\alpha \pm \frac{1}{\vartheta} \sqrt{2 \left(E_\alpha - \frac{1}{2} \rho_\alpha (u_\alpha)^2 \right) \left(\rho_\alpha - \frac{(\Delta \rho_\alpha)^2 \vartheta^2}{\rho_\alpha} \right)}.
\tag{59}
$$

6. Numerical experiments

In this section, we test the performance of the unstaggered-central scheme with the proposed extended IRP limiter. Section 6.1 is devoted to examine the MMP property by applying the scheme to scalar conservation laws. In Section 6.2, we examine the PP property by applying the scheme to nonlinear Euler equations with $\gamma = 1.4$.

6.1. Scalar conservation laws

In this section, we apply the proposed scheme to several scalar conservation laws. The numerical accuracies for continuous problems for both linear advection equation and nonlinear Burgers' equation are checked in Examples 1 and 2, respectively. The capturing properties of discontinuous solutions are checked for both linear advection equation and nonlinear Burgers' equation in Examples 1 and 2, respectively. We also check the ability of the convergence of the numerical solutions to the unique physical entropy solution for nonconvex fluxes Examples 3 and 4, respectively.

Example 1. We consider the linear advection equation,

$$
u_t + u_x = 0, \quad -1 \leq x \leq 1
\tag{60}
$$

with some initial condition $u(x, t = 0) = u_0(x)$ and periodic boundary conditions. This type of model problem has the exact solution

$$
u(x, t) = u_0(x - t).
\tag{61}
$$

We first check the influence of the new limiter on the accuracy order when it is applied to continuous initial data. The initial condition is

$$
u_0(x) = \sin(\pi x).
\tag{62}
$$

The simulation is done until the final time $t = 10$. Table 1 shows the L_1, L_2 and L_∞ errors and the corresponding experimental order of convergence, respectively. The expected accuracy orders are achieved.

To check the MMP property of the proposed scheme, we apply the scheme to the example proposed in [44]. This initial condition is highly discontinuous

$$
u_0(x + 0.5) = \begin{cases}
-x \sin \left(\frac{3}{2} \pi x^2 \right), & \text{if } -1 < x < -\frac{1}{3}, \\
|\sin(2\pi x)|, & \text{if } |x| \leq \frac{1}{3}, \\
2x - 1 - \sin(3\pi x)/6, & \text{if } \frac{1}{3} < x \leq 1.
\end{cases}
\tag{63}
$$
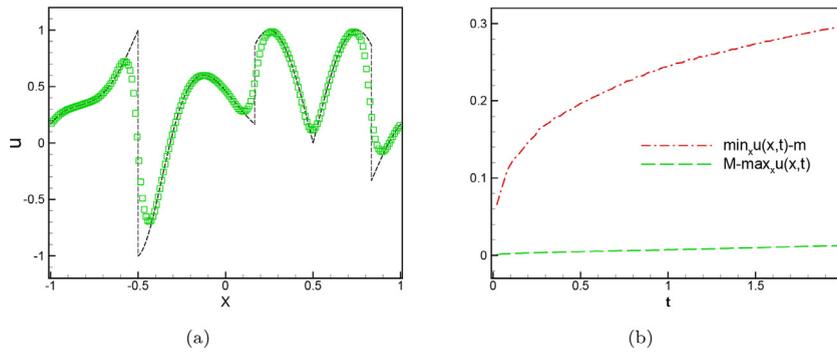
Fig. 4. Example 1: linear advection problem (60) with discontinuous initial condition (63). (a): numerical solution(marked as square) and exact solution(marked as dashed line) at time $t = 2$. (b): the numerical derivations of the minimum and maximum values from the exact ones. The numerical solution is obtained on 200 uniform cells.

Table 2

Example 2: Burgers' equation (64) with initial condition (65). Numerical errors and experimental orders of convergence in different norms at time $t = 0.5$ before the discontinuity comes. The reference solution is obtained on 25,600 meshes.

# cells	L_1 error	EOC	L_2 error	EOC	L_∞ error	EOC
25	2.045E-2	-	2.314E-2	-	4.729E-2	-
50	3.724E-3	2.457	5.091E-3	2.184	1.849E-2	1.355
100	9.806E-4	1.925	1.378E-3	1.885	5.927E-3	1.642
200	2.239E-4	2.131	3.310E-4	2.058	1.620E-3	1.872
400	4.703E-5	2.251	6.993E-5	2.243	3.565E-4	2.184
800	9.361E-6	2.329	1.384E-5	2.337	6.480E-5	2.460

The simulation is done on 200 uniform cells until final time $t = 2$. Fig. 4(a) shows the numerical result compared with the exact solution. The numerical solution reproduces the discontinuities correctly. In Fig. 4(b), we study the numerical deviations of the maximum and minimum values from the exact ones. It shows that they are preserved in the sense that $\max_x u(x, t) \leq M$ and $\min_x u(x, t) \geq m$, where M and m are the maximum and minimum values of the solution at the initial time, respectively.

Example 2. We consider the inviscid Burgers' equation

$$u_t + \left(\frac{u^2}{2} \right)_x = 0 \tag{64}$$

in the domain $[-1, 1]$ with the initial condition

$$u_0(x) = \frac{1}{4} + \frac{1}{2} \sin(\pi x) \tag{65}$$

and the periodic boundary conditions. This example is selected from [9]. The solution is smooth up to time $t = \frac{2}{\pi}$, then it develops a moving shock which interacts with a rarefaction wave. At $t = 2$ the interaction between the shock and the rarefaction waves is over, and the solution becomes monotone between the shocks.

The simulation is done on 200 uniform cells. The reference solution is obtained on 25,600 cells. At time $t = 0.5$, we list the errors in different norms in Table 2. We can clearly see the expected numerical accuracy orders. The numerical results at different time $t = \frac{2}{\pi}$, 2 and 4 are plotted in Fig. 5(a), (b) and (c), respectively. The moving shock is captured and resolved. Fig. 5(d) shows the numerical deviations of the maximum and minimum values from the initial ones. The solution is symmetric along the line $y = 0.25$, thus two deviations coincide. The MMP principle is verified.

Example 3. In this test we solve the one-dimensional scalar conservation laws with the nonconvex Buckley-Leverett flux

$$f(u) = \frac{4u^2}{4u^2 + (1 - u)^2} \tag{66}$$

in the domain $[-1, 1]$ with the initial condition

$$u_0(x) = \begin{cases} 1, & -0.5 \leq x \leq 0, \\ 0, & otherwise. \end{cases} \tag{67}$$
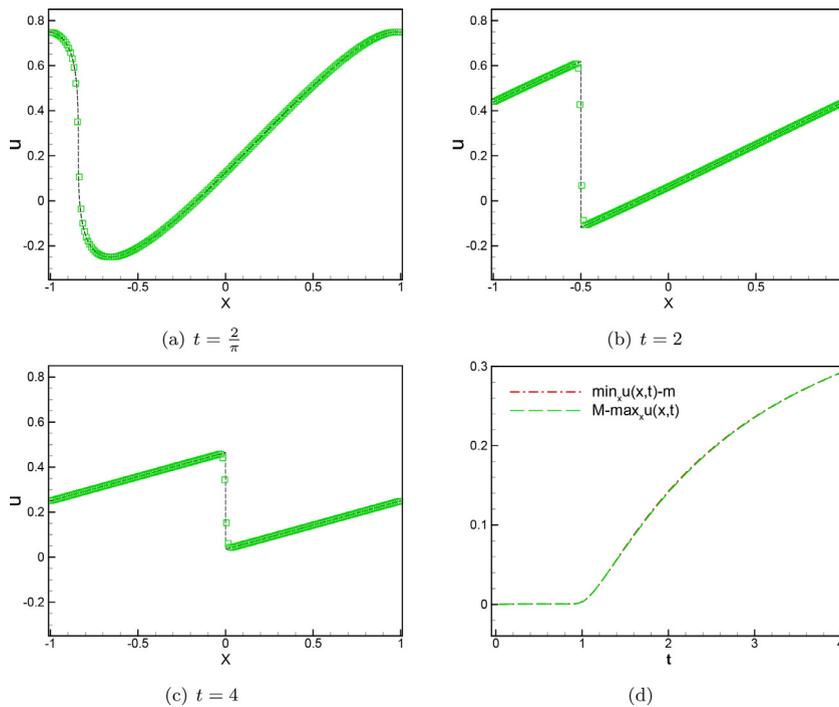
Fig. 5. Example 2: Burgers' equation (64) with initial condition (65). (a–c): numerical solution(marked as square) and referenced solution(marked as dashed line) at different times; (d): the numerical derivations of the minimum and maximum values from the initial ones. The numerical solution is obtained on 200 uniform cells. The reference solution is obtained on 25,600 cells.
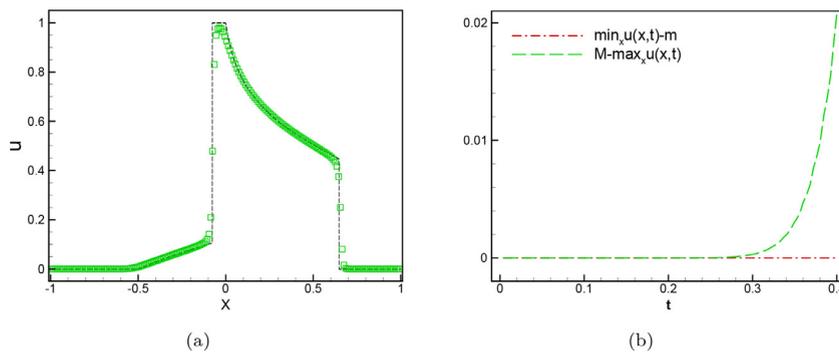
Fig. 6. Example 3: the scalar conservation laws with the nonconvex Buckley-Leverett flux (66) and initial condition (67). (a): numerical solution(marked as square) and referenced solution(marked as dashed line) at time $t = 0.4$; (b): the numerical derivations of the minimum and maximum values from the initial ones. The numerical solution is obtained on 200 uniform cells. The reference solution is obtained on 25,600 cells.

and out flow boundary conditions. This example is also widely discussed by the literatures, e.g. [9,24,27]. The finial time is $t = 0.4$. We use this test to check the convergence of the numerical solutions to the physically correct entropy solutions.

Because of the nonconvexity of the flux function, it is possible to form the composite waves [27]. A generic case of the composite waves in this example is a shock adjacent to a rarefaction wave. A numerical scheme that uses a compressive limiter, such as the generalized minmod limiter with $\theta > 1$, may result in nonphysical solution. We follow the method proposed in [27] and adapt the uniformly high-order non-oscillatory(UNO) limiter (44).

The reference solution is obtained using 25,600 cells. The computational result on 200 uniform cells is displayed in Fig. 6. Fig. 6(a) shows the numerical solution which is quite satisfactory. The numerical deviation of the maximum and minimum values is shown in Fig. 6(b). The minimum value is exactly the same as the initial minimum value. The maximum value is smaller than the initial maximum value.

Example 4. This is another widely used one-dimensional scalar conservation laws with non-convex flux

$$f(u) = \frac{1}{4}(u^2 - 1)(u^2 - 4). \tag{68}$$
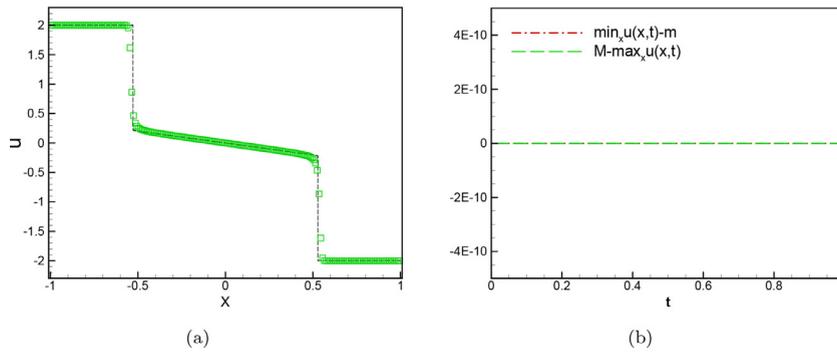
Fig. 7. Example 4: the scalar conservation laws with the nonconvex flux (68) and initial condition (69). (a): numerical solution(marked as square) and referenced solution(marked as dashed line) at time $t = 1$; (b): the numerical derivations of the minimum and maximum values from the initial ones. The numerical solution is obtained on 200 uniform cells. The reference solution is obtained on 25,600 cells.

Table 3
Example 5. Euler equations with initial condition (70). Numerical errors and experimental orders of convergence in different norms at time $t = 0.2$. The reference solution is obtained on 25,600 meshes.

# cells	L_1 error	EOC	L_2 error	EOC	L_∞ error	EOC
25	9.092E-2	-	4.398E-2	-	4.005E-2	-
50	1.084E-2	3.068	6.420E-3	2.776	7.630E-3	2.392
100	6.016E-4	4.171	2.665E-4	4.590	1.506E-4	5.663
200	1.580E-4	1.929	6.999E-5	1.929	3.949E-5	1.931
400	4.043E-5	1.966	1.792E-5	1.966	1.011E-5	1.966
800	1.022E-5	1.984	4.528E-6	1.985	2.555E-6	1.984
1600	2.562E-6	1.996	1.135E-6	1.996	6.405E-7	1.996

We consider the initial condition

$$u_0(x) = \begin{cases} 2, & x \le 0, \\ -2, & x > 0 \end{cases} \tag{69}$$

and nature boundary conditions. The finial time is $t = 1$. There are also composite waves with adjacent shock and rarefaction waves. The adaptive limiter proposed in [27] is used.

The reference solution is obtained using 25,600 cells. The computational result on 200 uniform cells is displayed in Fig. 7. Fig. 7(a) shows the numerical solution which is quite satisfactory. The numerical deviations of the maximum and minimum values is shown in Fig. 7(b). The maximum and minimum values are exactly same with initial ones.

6.2. Nonlinear Euler equations

In this section, we examine the proposed scheme for compressible Euler equations. The numerical accuracy for continuous problem with very low density and low pressure is checked in Example 5. We apply the scheme to the Sod shock tube problem in Example 6 to check higher resolution property. Then a new 123 problem which generates vacuum state in the middle in Example 7 is designed to check the PP property. Next, a Leblanc shock tube problem in Example 8 is considered to check the property of our scheme to resolve the waves with large ratios of density and pressure. Finally, a Sedov blast wave problem in Example 9 which is a typical low density problem involving strong shocks is given to demonstrate the ability to capture shock.

Example 5. Consider a low density and low pressure problem. The initial condition is

$$\rho_0(x) = 1 + 0.99 \sin(x), \quad u_0(x) = 1, \quad p_0(x) = 0.01. \tag{70}$$

The domain is $[0, 2\pi]$ and the boundary conditions are periodic. The final time is $t = 0.2$.

The L_1, L_2 and L_∞ errors and the corresponding experimental orders of convergence are listed in Table 3, respectively. We observe the designed order of the accuracy of this problem with low density and low pressure.

This indicates that the practical time step selection mentioned in Remark 3.8 maintains the numerical accuracy.

Example 6. This is the famous Sod shock tube problem introduced in [45]. The initial condition is

$$(\rho, u, p) = \begin{cases} (1, \ 0, \ 1), & \text{if } x < 0.5, \\ (0.125, \ 0, \ 0.1), & \text{if } x > 0.5. \end{cases} \tag{71}$$
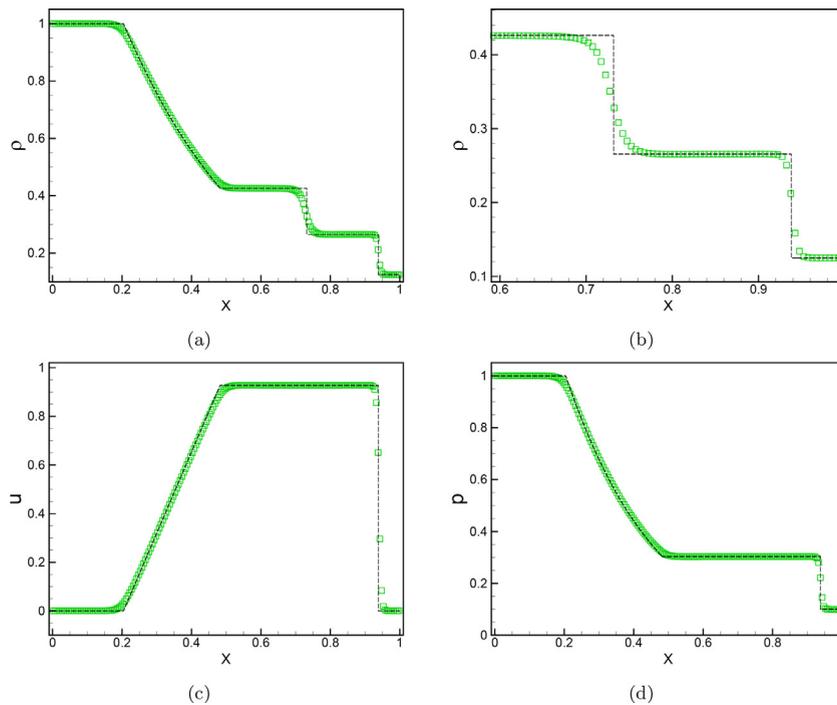
Fig. 8. Example 6. Numerical solution (marked as square) and referenced solution (marked as dashed line) of Euler equations with initial condition (71) at time $t = 0.25$: (a) density; (b) zoomed density; (c) velocity; (d) pressure. The solution is obtained on 200 uniform meshes.

The domain is [0,1] and the zero extending boundary condition is applied. The final time is $t = 0.25$. The solution consists of a left rarefaction, a contact and a right shock.

We do the simulation on 200 uniform meshes. The numerical solutions of density, velocity, pressure are shown in Fig. 8. We can see that all three waves are captured. In Fig. 8(b), we plot the zoomed figure of the density in the region [0.59,1] from where no oscillation is observed across the discontinuities including both contact and shock.

Example 7. We consider 123 problem in [45] which has two rarefaction waves. Different from the traditional case, the initial condition will generate vacuum state in the middle. The initial condition is

$$(\rho, u, p) = \begin{cases} (1, \ -2, \ 0.15), & \text{if } x < 0.5, \\ (1, \ 2, \ 0.15), & \text{if } x > 0.5. \end{cases} \tag{72}$$

The domain is [0,1] and the zero extending boundary condition is applied. The final time is $t = 0.15$.

We do the simulation on 200 uniform meshes. The numerical solutions of density, velocity, pressure are shown in Fig. 9. The two rarefaction are captured. The low density and low pressure in the middle region near the origin (Vacuum) are also captured by proposed scheme. We also observe that the new origin of the velocity is also captured. We also plot the minimums of both density and pressure as functions of time in Fig. 9(d), which shows that the PP property is maintained in the whole simulation.

Example 8. We consider a Leblanc shock tube problem which is selected from references [10,46]. Similar discussion is done in [47]. The initial condition is

$$(\rho, u, p) = \begin{cases} (1000, \ 0, \ 1000), & \text{if } x < 0, \\ (1, \ 0, \ 1), & \text{if } x > 0. \end{cases} \tag{73}$$

The domain is $[-10, 10]$ and the zero extending boundary condition is applied. The final time is $t = 10^{-4}$. The solution consists of two strong rarefaction waves and a trivial stationary contact discontinuity.

For such shock tube problem with initial high pressure and high density ratios, the popular high resolution schemes worked very inefficiently. The increasing scheme accuracy did not improve the results much and the numerical smearing at the contact discontinuity did not attribute much to this defect [46]. In order to better solve the problem, the authors use the exact solution of $\mathbf{u}(x, t_0) = \mathbf{u}^R(x, t_0)$ as the initial condition at t_0, which is bigger than the critical time t_{cr}. If the exact solution of $t_0 < t_{cr}$ is used as the initial condition, the high resolution schemes are unable to provide the correct shock
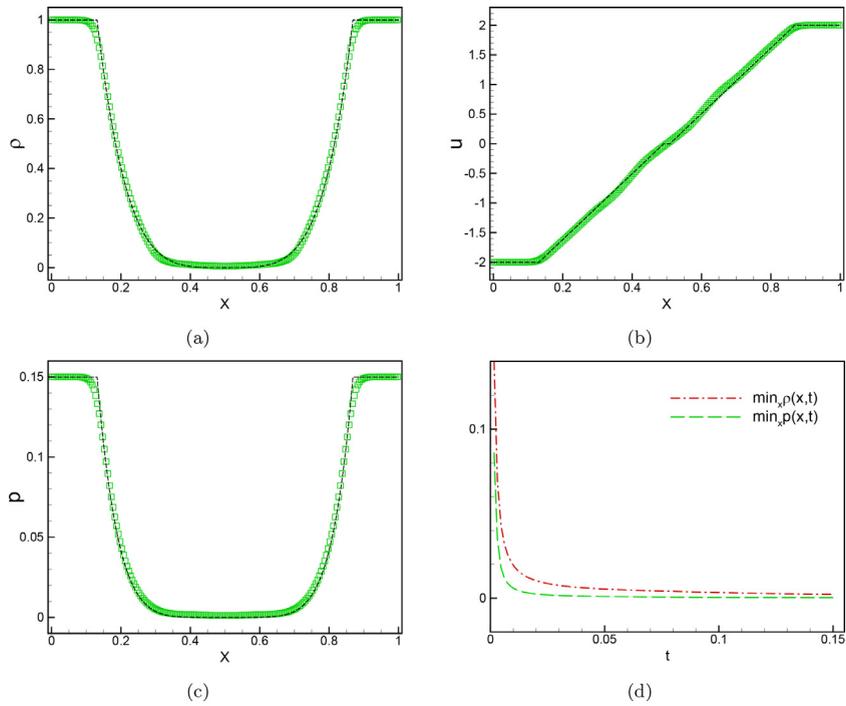
Fig. 9. Example 7. Numerical solution (marked as square) and referenced solution (marked as dashed line) of Euler equations with initial condition (72) at time $t = 0.15$: (a) density; (b) velocity; (c) pressure; (d) minimums of density and pressure. The solution is obtained on 200 uniform meshes.
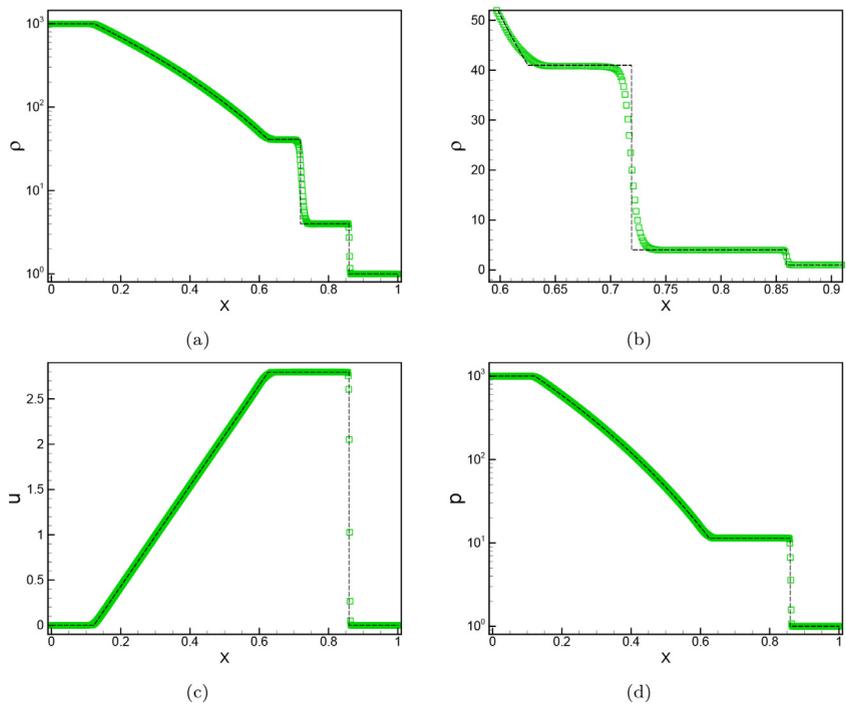
Fig. 10. Example 8. Numerical solution (marked as square) and referenced solution (marked as dashed line) of Euler equations with initial condition (73) at time $t = 10^{-4}$: (a) density; (b) zoomed density; (c) velocity; (d) pressure. The solution is obtained on 800 uniform meshes.
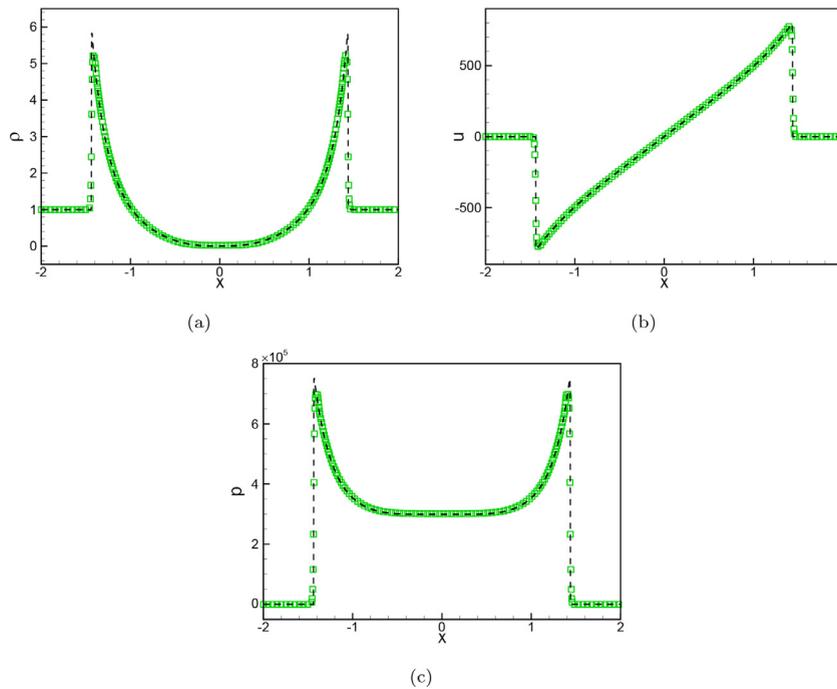
Fig. 11. Example 9. Numerical solution (marked as square) and referenced solution (marked as dashed line) of Euler equations at time $t = 0.001$: (a) density; (b) velocity; (c) pressure. The solution is obtained on 800 meshes.

location. In this example, we take $t_0 = 3 \times 10^{-6}$ as the critical time. We do the simulation on 800 uniform meshes and the results are shown in Fig. 10.

Example 9. The Sedov blast wave is chosen from reference [10,48] and it is a typical low density problem involving two very strong shocks. For the initial condition, the density is 1, velocity is zero, total energy is 10^{-12} everywhere except that the energy in the center cell is the constant $\frac{E_0}{\Delta x}$ with $E_0 = 3200000$ (emulating a $\delta-$function at the center). The final time is $t = 0.001$. We do the simulation on 800 uniform meshes and the results are shown in Fig. 11, the minimum density in Fig. 11(a) is 2.49×10^{-2}. We can see the shock is captured very well even with the time step proposed in Remark 3.8. The results are comparable to the results in references [10,48].

7. Conclusion

In present paper, the emphasis is on an applicable stability analysis of invariant-region-preserving (IRP) principle. The proof is given by a so called forward-backward splitting method. Specifically, the evolved solution is decomposed into several terms, which is a convex combination of the extrapolated states and other states in the forms of the Lax-Friedrichs scheme. These states are obtained by solving the forward and backward processes of the original PDE. With the extended IRP limiter applied to the reconstructed slopes in the forward projection step, a relatively relaxed CFL number is obtained, which also implies a larger time step. The extended IRP limiter and method of the stability analysis are available to both scalar equation and general nonlinear systems. The main properties of the present scheme are checked by a series of numerical experiments. Here we want to illustrate that the stability condition may not be optimal, which needs to be clarified in the future. The next work is to extend the present IRP scheme to multi-dimensional cases and apply it to solve other nonlinear systems.

Data availability

No data was used for the research described in the article.

Acknowledgment

References

[1] P.D. Lax, Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves, SIAM, 1973.

[2] K.N. Chueh, C.C. Conley, J.A. Smoller, Positively invariant regions for systems of nonlinear diffusion equations, Indiana Univ. Math. J. 26 (2) (1977) 373–392.

[3] Y. Jiang, H. Liu, Invariant-region-preserving DG methods for multi-dimensional hyperbolic conservation law systems, with an application to compressible Euler equations, J. Comput. Phys. 373 (2018) 385–409.

[4] K. Wu, Minimum principle on specific entropy and high-order accurate invariant region preserving numerical methods for relativistic hydrodynamics, SIAM J. Sci. Comput. 43 (6) (2021) B1164–B1197.

[5] J.-L. Guermond, M. Nazarov, B. Popov, I. Tomas, Second-order invariant domain preserving approximation of the Euler equations using convex limiting, SIAM J. Sci. Comput. 40 (5) (2018) A3211–A3239.

[6] J.-L. Guermond, B. Popov, Invariant domains and first-order continuous finite element approximation for hyperbolic systems, SIAM J. Numer. Anal. 54 (4) (2016) 2466–2489.

[7] C. Berthon, F. Marche, A positive preserving high order VFRoe scheme for shallow water equations: a class of relaxation schemes, SIAM J. Sci. Comput. 30 (5) (2008) 2587–2612.

[8] B. Perthame, C.-W. Shu, On positivity preserving finite volume schemes for Euler equations, Numer. Math. 73 (1) (1996) 119–130.

[9] X. Zhang, C.-W. Shu, On maximum-principle-satisfying high order schemes for scalar conservation laws, J. Comput. Phys. 229 (9) (2010) 3091–3120.

[10] X. Zhang, C.-W. Shu, On positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations on rectangular meshes, J. Comput. Phys. 229 (23) (2010) 8918–8934.

[11] Y. Cheng, F. Li, J. Qiu, L. Xu, Positivity-preserving DG and central DG methods for ideal MHD equations, J. Comput. Phys. 238 (2013) 255–280.

[12] K. Wu, H. Tang, High-order accurate physical-constraints-preserving finite difference WENO schemes for special relativistic hydrodynamics, J. Comput. Phys. 298 (2015) 539–564.

[13] K. Wu, Y. Xing, Uniformly high-order structure-preserving discontinuous Galerkin methods for Euler equations with gravitation: positivity and well-balancedness, SIAM J. Sci. Comput. 43 (1) (2021) A472–A510.

[14] S. Godunov, I. Bohachevsky, Finite difference method for numerical computation of discontinuous solutions of the equations of fluid dynamics, Matematičeskij sbornik 47 (3) (1959) 271–306.

[15] A. Harten, P.D. Lax, B. van Leer, On upstream differencing and Godunov-type schemes for hyperbolic conservation laws, SIAM Rev. 25 (1) (1983) 35–61.

[16] B. Einfeldt, On Godunov-type methods for gas dynamics, SIAM J. Numer. Anal. 25 (2) (1988) 294–318.

[17] E.F. Toro, Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction, Springer Science & Business Media, 2013.

[18] W. Tong, R. Yan, G. Chen, A sharp stability analysis of the MUSCL-Hancock scheme for scalar hyperbolic conservation laws, submitted (2022).

[19] G. Chen, S. Noelle, A new hydrostatic reconstruction scheme based on subcell reconstructions, SIAM J. Numer. Anal. 55 (2) (2017) 758–784.

[20] G. Chen, S. Noelle, A unified surface-gradient and hydrostatic reconstruction scheme for the shallow water equations, J. Comput. Phys. 467 (2022) 111463.

[21] H. Nessyahu, E. Tadmor, Non-oscillatory central differencing for hyperbolic conservation laws, J. Comput. Phys. 87 (2) (1990) 408–463.

[22] R. Sanders, A. Weiser, High resolution staggered mesh approach for nonlinear hyperbolic systems of conservation laws, J. Comput. Phys. 101 (2) (1992) 314–329.

[23] G.-S. Jiang, E. Tadmor, Nonoscillatory central schemes for multidimensional hyperbolic conservation laws, SIAM J. Sci. Comput. 19 (6) (1998) 1892–1917.

[24] G.-S. Jiang, D. Levy, C.-T. Lin, S. Osher, E. Tadmor, High-resolution nonoscillatory central schemes with nonstaggered grids for hyperbolic conservation laws, SIAM J. Numer. Anal. 35 (6) (1998) 2147–2168.

[25] A. Kurganov, E. Tadmor, New high-resolution central schemes for nonlinear conservation laws and convection–diffusion equations, J. Comput. Phys. 160 (1) (2000) 241–282.

[26] A. Kurganov, G. Petrova, A third-order semi-discrete genuinely multidimensional central scheme for hyperbolic conservation laws and related problems, Numer. Math. 88 (4) (2001) 683–729.

[27] A. Kurganov, G. Petrova, B. Popov, Adaptive semidiscrete central-upwind schemes for nonconvex hyperbolic conservation laws, SIAM J. Sci. Comput. 29 (6) (2007) 2381–2401.

[28] B. Popov, Y. Hua, Invariant domain preserving central schemes for nonlinear hyperbolic systems, Commun. Math. Sci. 19 (2) (2021) 529–556.

[29] X.-D. Liu, E. Tadmor, Third order nonoscillatory central scheme for hyperbolic conservation laws, Numer. Math. 79 (3) (1998) 397–425.

[30] R. Touma, P. Arminjon, Central finite volume schemes with constrained transport divergence treatment for three-dimensional ideal MHD, J. Comput. Phys. 212 (2) (2006) 617–636.

[31] R. Touma, U. Koley, C. Klingenberg, Well-balanced unstaggered central schemes for the euler equations with gravitation, SIAM J. Sci. Comput. 38 (5) (2016) B773–B807.

[32] R. Touma, Central unstaggered finite volume schemes for hyperbolic systems: applications to unsteady shallow water equations, Appl. Math. Comput. 213 (1) (2009) 47–59.

[33] R. Touma, C. Klingenberg, Well-balanced central finite volume methods for the Ripa system, Appl. Numer. Math. 97 (2015) 42–68.

[34] R. Touma, Well-balanced central schemes for systems of shallow water equations with wet and dry states, Appl. Math. Model. 40 (4) (2016) 2929–2945.

[35] R. Touma, F. Kanbar, Well-balanced central schemes for two-dimensional systems of shallow water equations with wet and dry states, Appl. Math. Model. 62 (2018) 728–750.

[36] R. Touma, M. Saleh, Well-balanced central schemes for pollutants transport in shallow water equations, Math. Comput. Simulat. 190 (2021) 1275–1293.

[37] Y. Liu, Central schemes on overlapping cells, J. Comput. Phys. 209 (1) (2005) 82–104.

[38] Y. Liu, C.-W. Shu, E. Tadmor, M. Zhang, Central discontinuous Galerkin methods on overlapping cells with a nonoscillatory hierarchical reconstruction, SIAM J. Numer. Anal. 45 (6) (2007) 2442–2467.

[39] R. Yan, W. Tong, G. Chen, An improved central scheme for nonlinear shallow water equations, submitted (2022).

[40] P.D. Lax, Weak solutions of nonlinear hyperbolic equations and their numerical computation, Commun. Pur. Appl. Math. 7 (1) (1954) 159–193.

[41] C. Berthon, Stability of the MUSCL schemes for the Euler equations, Commun. Math. Sci. 3 (2) (2005) 133–157.

[42] A. Harten, S. Osher, Uniformly high order accurate nonoscillatory schemes. I, SIAM J. Numer. Anal. 24 (1982) 279–309.

[43] B. van Leer, Towards the ultimate conservative difference scheme. V. A second-order sequel to Godunov's method, J. Comput. Phys. 32 (1) (1979) 101–136.

[44] G.-S. Jiang, C.-W. Shu, Efficient implementation of weighted ENO schemes, J. Comput. Phys. 126 (1) (1996) 202–228.

[45] G.A. Sod, A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws, J. Comput. Phys. 27 (1) (1978) 1–31.

[46] H. Tang, T. Liu, A note on the conservative schemes for the Euler equations, J. Comput. Phys. 218 (2) (2006) 451–459.

[47] J. Li, Q. Li, K. Xu, Comparison of the generalized Riemann solver and the gas-kinetic scheme for inviscid compressible flow simulations, J. Comput. Phys. 230 (12) (2011) 5080–5099.

[48] X. Zhang, C.-W. Shu, Positivity-preserving high order finite difference WENO schemes for compressible Euler equations, J. Comput. Phys. 231 (5) (2012) 2245–2258.