Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Computational Physics 193 (2003) 243–259
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

Kinetic energy fix for low internal energy flows

X.Y. Hu [a,*], B.C. Khoo [a,b]

[a] Singapore-MIT Alliance, National University of Singapore, 4 Engineering Drive 3, Singapore 117576, Singapore
[b] Department of Mechanical Engineering, National University of Singapore, Singapore 119260, Singapore

Received 11 October 2002; received in revised form 26 May 2003; accepted 11 August 2003

Abstract

When the kinetic energy of a flow is dominant, numerical schemes employed can encounter difficulties due to negative internal energy. A case study with several commonly used conservative schemes (MUSCL, ENO, WENO and CE/SE) shows that high order schemes may have less ability to preserve positive internal energy (MUSCL and CE/SE), or present less accurate results (WENO and ENO) when the internal energy to kinetic energy ratio is low. By analyzing the positivity property for second-order conservative schemes with large fixed CFL number conditions for time step restriction, this paper proposes the energy consistency conditions for second-order Riemann-solver type schemes and CE/SE method. According to the said energy consistency conditions, a kinetic energy fix method which limits the magnitude of kinetic energy relative to the total energy is introduced. The numerical examples show that the kinetic energy fixed CE/SE method produces reasonable results and keeps positive internal energy for flows with very low internal energy even when a vacuum occurs.
© 2003 Elsevier B.V. All rights reserved.

Keywords: Low internal energy flow; Positivity property; Conservative scheme

1. Introduction

The gas dynamic problems are usually solved by conservative schemes, in which the internal energy is used to determine the pressure by subtracting the kinetic energy component from the total energy. For flows in which the ratio of internal energy to kinetic energy is low, the resulting internal energy obtained may be negative hence giving rise to negative pressure and thereby invalidating the ensuing computation. A useful test case to evaluate the ability of a computational method in handling such low internal energy flow is that of a one-dimensional tube containing a gas having diametrically opposite initial velocities, which is usually referenced as the "1–2–3 problem" [18]. Einfeldt et al. [3] analyzed the characteristics of low internal energy flows and proposed the HLLE scheme [6] to solve the "1–2–3 problem". Toro [18] tested and found that several classic Riemann-solver type schemes can keep the density and internal energy positive and have

* Corresponding author. Tel.: +65-68744797; fax: +65-67752920.
E-mail addresses: smahx@nus.edu.sg (X.Y. Hu), mpekbc@nus.edu.sg (B.C. Khoo).

0021-9991/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/j.jcp.2003.08.007

the so called *positivity property*. Gressier et al. [5] also discussed the positivity conditions of several classical flux vector splitting schemes.

For high order conservative schemes, even for those constructed based on positive first-order schemes, as will be shown by a case study in Section 4, may show less ability to maintain positivity or not being able to obtain results with sufficient accuracy for low internal energy flow problems (or those when the ratio of internal to kinetic energy is lower than 3). On the positivity property, Linde and Roe [9] discussed about the positivity conditions for a second order multi-dimensional MUSCL-type scheme while Perthame [11] and Tang and Xu [17] studied the positivity conditions for second-order kinetic schemes. Perthame and Shu [12] provided a remarkable theorem which states that, given a first-order positive conservation scheme such as Godunov and Lax–Friedrichs schemes, one can always build a higher order positive scheme under the conditions that (a) the cell wall values for numerical flux calculation satisfy positive density and pressure, and (b) sufficiently small CFL number be utilized to constrain the time step. As a fixed CFL number is usually employed in practical computation, (c) additional constraints on the interpolation procedure are needed. Usually, the fixed CFL number permitted in practical computation is considerably small and therefore greatly decreased the efficiency. On the other hand, to increase the accuracy of computation for low internal energy flows, Cocchi et al. [2] suggested a second-order non-conservative formulation for the energy equation to decrease the non-physical temperature increase. However, the non-conservative scheme may result in an exponential error growth, and it is not known that if this non-conservative formula can handle flows with even lower internal to kinetic energy ratios. The CE/SE method [1], a non-Riemann-solver type conservative scheme, introduces less errors when computing for low internal energy flows with reasonably large CFL number for time step restriction. However, as shown in the case study below, it works well provided the internal energy of the flows is not very low.

The motivation of this paper stems from the above mentioned difficulties of maintaining positivity and accuracy preserving for high order schemes when computing for low internal energy flows. We first study the performances of several different high order conservative schemes (MUSCL, ENO, WENO and CE/SE) to determine the difficulties that each scheme may face with the decrease of internal energy in the flow. Then we analyze the positivity properties of second-order Riemann-solver type schemes and the CE/SE method for the cases with large fixed CFL number for high computational efficiency. Based on these analysis, we propose the energy consistency conditions, by which a kinetic energy fix method for a general second-order conservative scheme is introduced. As the CE/SE method is found from the case study to give more accurate results, it is modified with the said kinetic energy fix to compute for several numerical examples for flows with a large range of low internal energy.

2. Euler equations

Assuming the fluid is inviscid and compressible, the flow is described by Euler equations in one dimension as

$$\frac{\partial U}{\partial t}+\frac{\partial F(U)}{\partial x}=0, \tag{1}$$

where $U=(\rho,\rho u,E)^{\mathrm{T}}$ and $F(U)=(\rho u,\rho u^2+p,(E+p)u)^{\mathrm{T}}$. The equation of state is defined as

$$p=(\gamma-1)\rho e, \tag{2}$$

where $\gamma=1.4$ is the heat ratio for an ideal gas. This set of equations describes the conservation of the conservative variables: density ρ, momentum ρu and total energy density $E=\rho e+\frac{1}{2}\rho u^2$, where e is the internal energy per unit mass. The ratio of internal energy to kinetic energy can be defined as

$$K = \frac{2\rho E - (\rho u)^2}{(\rho u)^2}. \tag{3}$$

In a low internal energy flow, the internal energy is close or even smaller than the kinetic energy, i.e. $K < 1$. However, the internal energy and density must always be kept positive, i.e. $K > 0$, to maintain a positive pressure p from Eq. (2) for a real physical state. As the state is derived from the conservative variables when solving Eq. (1), we say that the conservative variables are *energy consistent* when the system possesses a positive internal energy hence giving rise to a positive pressure. If $K < 0$, the energy consistency fails, and negative internal energy and pressure occur.

3. Conservative schemes

3.1. Riemann-solver type

An explicit Riemann-solver type conservation scheme of Eq. (1) on cell j can be written as

$$\bar{U}_j^{n+1} = \bar{U}_j^n - \lambda(\hat{F}_{j+1/2} - \hat{F}_{j-1/2}), \tag{4}$$

where $\bar{U}_j^n$ and $\bar{U}_j^{n+1}$ are the cell average values at nth and $n+1$th time steps, $\hat{F}_{j\pm1/2}$ are numerical fluxes on the respective cell walls. The scheme is stable under a Courant–Friedrich–Lewy (CFL) time step restriction

$$\lambda = \frac{\Delta t}{\Delta x} < \frac{N_{\mathrm{CFL}}}{\max(|u_i| + c_i)}, \quad i = 1, N, \tag{5}$$

where $N_{\mathrm{CFL}} < 1$ is called CFL number. For a first-order conservative scheme, $\hat{F}_{j\pm1/2}$ is defined directly based on the cell average values, i.e.

$$\hat{F}_{j+1/2} = \hat{F}(\bar{U}_j, \bar{U}_{j+1}), \quad \hat{F}_{j-1/2} = \hat{F}(\bar{U}_{j-1}, \bar{U}_j). \tag{6}$$

Several classic conservative schemes, such as Godunov, Lax–Friedrichs, HLLE, have positivity property under a CFL number condition

$$N_{\mathrm{CFL}} < \alpha_0, \tag{7}$$

where $\alpha_0 \leqslant 1$. The positivity property ensures positive density and pressure from $\bar{U}^{n+1}$ when $\bar{U}^n$ satisfies positive density and pressure.

A higher order scheme is usually constructed based on a positive preserving first-order scheme, such as MUSCL is based on Godunov scheme while ENO or WENO are based on Lax–Friedrichs scheme. The difference is that the $\hat{F}_{j\pm1/2}$ in a high order scheme are defined with the values of the forward and backward conservative variables approximated by a piecewise function $U^n(x)$ on the walls of cell, say $x_{j-1/2}$ and $x_{j+1/2}$. Specifically, the second-order profiles can be written as

$$\rho = \bar{\rho}_j + k_\rho x, \quad \rho u = (\bar{\rho u})_j + k_{\rho u} x, \quad E = \bar{E}_j + k_E x, \tag{8}$$

in which the cell average values also represent the values on the node. The numerical fluxes of a high order scheme are

$$\hat{F}_{j+1/2} = \hat{F}(U^-_{j+1/2}, U^+_{j+1/2}), \quad \hat{F}_{j-1/2} = \hat{F}(U^-_{j-1/2}, U^+_{j-1/2}). \tag{9}$$

On the conservative property, these high order approximations are always conserved pertaining to the local mean on the cell, i.e. $\bar{U}_j^n = (1/\Delta x)\int_{x_{j-1/2}}^{x_{j+1/2}} U^n(x)\mathrm{d}x$. For second-order schemes, the linear interpolation approximations shown in Eq. (8) automatically ensure cell average values are conserved.

3.2. CE/SE method

The CE/SE method [1] is a non-Riemann-solver type conservative scheme and has second-order accuracy in both time and spatial directions. In the CE/SE method, the time integration is performed on a staggered grid and each full time step (satisfying Eq. (5)) is divided into two half time steps. From the initial conditions, smooth regions are defined in the cells near the nodes, such as $(x_{j-1/2}, x_{j+1/2})$, with the linear profiles of Eq. (8). When the first half time step is calculated, the cell is shifted by $\Delta x/2$, which is equivalent to defining for the region between nodes, such as (x_j, x_{j+1}); hence new cell average values $\bar{U}_{j+1/2}^n$ are defined. As the values at the nodes are smooth with first-order derivatives, the fluxes on the new cell walls are physical. Hence, a conservative scheme similar to Eq. (4) can be written as

$$\bar{U}_{j+1/2}^{n+1/2} = \bar{U}_{j+1/2}^n - \frac{\lambda}{2}\left[F(U_{j+1}^{n+1/4}) - F(U_j^{n+1/4})\right], \tag{10}$$

where $\bar{U}_{j+1/2}^{n+1/2}$ are the cell average values after the first half time step , $F(U_{j+1}^{n+1/4})$ and $F(U_j^{n+1/4})$ are the physical fluxes on node points at time $(n+1/4)\Delta t$. After the first half time step, the linear profile near $j+1/2$ is then constructed by a weighted average interpolation approach with the new cell average values $\bar{U}_{j+1/2}^{n+1/2}$ and the new node point values $U_{j+1}^{n+1/2}$ and $U_j^{n+1/2}$ at time $(n+1/2)\Delta t$. The values $U_j^{n+1/4}$ and $U_j^{n+1/2}$ at each node point are defined by the first-order Taylor expansion

$$U_j^{n+\xi} = \bar{U}_j^n + \xi\Delta t\left(\frac{\partial U}{\partial t}\right)_j^n, \tag{11}$$

where $\xi = 1/2$ or $1/4$. With the relation of Eq. (1) and linear approximation, Eq. (11) can be written as

$$U_j^{n+\xi} = \bar{U}_j^n - \xi\lambda\left[F(U_{j+1/2}^-) - F(U_{j-1/2}^+)\right], \tag{12}$$

where $U_{j+1/2}^-$ and $U_{j-1/2}^+$ are the initial cell wall values which are given by a linear profile such as Eq. (8). When the second half time step is calculated, the difference form can be written in a similar way as for Eq. (10), but all the terms are shifted by $\Delta x/2$ spatially. After a full time step, the cell locations are shifted back to the original locations again.

4. Case study of low internal energy flows: "1–2–3 problem"

The "1–2–3 problem" [18] gives a series of Riemann problems for an ideal gas

$$U(x,0) = \begin{cases} U_l & \text{if } x < x_0, \\ U_r & \text{if } x > x_0, \end{cases} \tag{13}$$

where the two constant initial conservative variables are $U_l(1,-2,E_0)$ and $U_r(1,2,E_0)$ and total energy $E_0 > 2$. E_0 can be changed according to the different initial internal energy ratio. According to Smoller [15], if the initial internal to kinetic energy ratio $K_0 \geqslant \frac{1}{7}$ ($E_0 \geqslant \frac{16}{7}$), the theoretical solution consists of two rarefaction waves propagating in opposite directions and a static low density region exists between the rarefaction regions. As K_0 decreases, the density of the static region decreases. If the initial internal energy ratio $K_0 \leqslant \frac{1}{7}$, a vacuum occurs in the solution. However, there arises difficulties for numerical methods when

the initial condition is far less critical than a vacuum solution. Einfeldt et al. [3] found that, if $K_0 \leqslant 4$ ($E_0 \leqslant 10$), the solution is not linearizable and may lead to numerical difficulties. In this paper, four commonly used high order conservative schemes are tested:

- Second-order MUSCL Hancock scheme (MUSCL) [13].
- Third-order ENO-LF scheme (ENO) [14].
- Fifth-order WENO-LF scheme (WENO) [7].
- Second-order CE/SE method (CE/SE) [1].

First, a case which has also been considered by Cocchi et al. [2] is tested. The initial total energy and internal energy ratio are $E_0 = 6$ and $K_0 = 2$. Here, the initial condition is much less critical than that associated with a vacuum solution. While the ENO scheme and the WENO scheme have no difficulty for various CFL numbers used, both the MUSCL scheme and the CE/SE method 'explode' or grow uncontrollably in the presence of negative internal energy when the CFL number of 0.9 is applied. Because of this, lower CFL numbers of 0.5 and 0.8 are used, respectively. Fig. 1 shows the calculated pressure, density,

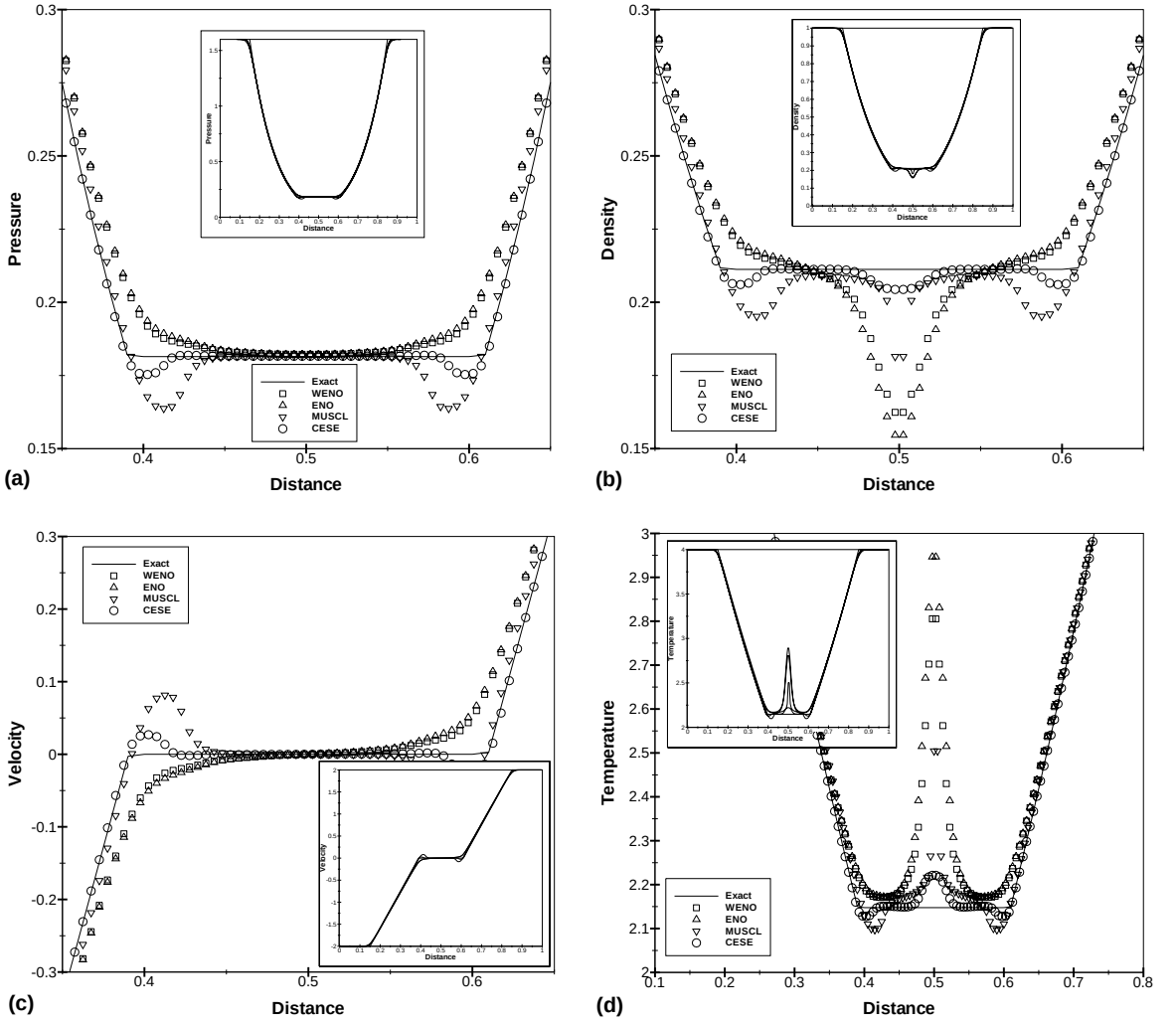

Fig. 1. Pressure, density, velocity and temperature profiles at $t = 0.1$ obtained by various numerical methods.

Table 1
Requirement of CFL number for the four numerical schemes with different E_0[a]

E_0/K_0	MUSCL Hancock-2	ENO-LF-3[b]	WENO-LF-5[b]	CE/SE-2
8.0/3.0	0.9	0.9	0.9	0.9
6.0/2.0	0.5	0.9	0.9	0.8
3.0/0.5	–	0.9	0.9	0.6
2.5/0.25	–	0.9	0.9	0.4
2.25/0.125[c]	–	0.7	0.7	0.3

[a] It does not necessarily imply that the solution is accurate compared to analysis.
[b] Usually, a CFL number bigger than 0.8 is not recommended.
[c] A vacuum occurs in exact solution.

velocity and temperature profiles at time $t = 0.1$. For all the flow variables considered, it is found that all the four schemes give reasonable results for the rarefaction regions. By comparing the low temperature and low density static region at the center of the domain with the exact solution, all the four schemes show much larger deviations than for the rarefaction regions. One can still discern the CE/SE method gives the smallest errors, especially for the temperature and density distributions. The ENO and WENO schemes give the largest deviations; further numerical tests on the ENO and WENO show results with smaller CFL number, even at 0.1, the errors cannot be reduced much (not shown here). For the MUSCL scheme, the predicted temperature is comparatively much lesser, however the deviation is still much larger than of the CE/SE method. From the above, one can suggest that the CE/SE method gives the best solution, which predicts the density fairly accurately and considerably reduces the increase of the temperature in the static region. It may be noted that these results by the CE/SE method are already better than that shown by Cocchi et al. [2] with non-conservative modifications.

However, the CE/SE method must be computed with smaller CFL number to keep the internal energy positive, just as for the MUSCL scheme. Table 1 shows the nominally initial conditions of imposed largest CFL number which can be used with the different internal energy ratio for the four schemes such that subsequent computations can continue and still maintain a strictly non-negative internal energy. From Table 1, it is clear that the MUSCL scheme shows the poorest ability to keep the internal energy positive. It 'explodes' when $K_0 = 1$ or smaller, even when the CFL number is decreased to a very small value such as 0.1. The CE/SE method behaves better, but the CFL number also shows a need to be decreased as the initial internal energy ratio decreases. The ENO and WENO schemes are much more robust than both the MUSCL and CE/SE schemes; numerical results show that the ENO and WENO schemes can maintain positivity with much larger CFL number than that in Perthame and Shu [12]. On the other hand, the results of the second-order MUSCL scheme and CE/SE scheme in particular seem to exhibit higher accuracy for such low internal energy flows.

The above results show, for these conservative schemes, positivity and accuracy are difficult to be maintained at the same time. In the following section, the positivity of second-order Riemann-solver type schemes and the CE/SE method are analyzed to show that the difficulties/problems associated with low internal energy is mainly caused by energy consistency failure. A kinetic energy fix method is then proposed to ensure the positivity property while still maintaining reasonably large fixed CFL number and high accuracy at the same time.

5. Analysis of low internal energy flows

As the MUSCL scheme and CE/SE method have no difficulty in maintaining positive internal energy for the "1–2–3 problem" if they degenerate to the first-order forms, and the main difference between the high

order conservative scheme and first-order scheme is the interpolations used to predict the profiles inside the computational cell j concerned, it is quite logical to study how the interpolations affect the positivity property.

5.1. Positivity property of Riemann-solver type schemes

Theorem 1. *Assume the first-order form of a Riemann-solver type conservative scheme satisfies the positivity property under a CFL condition of Eq.* (7). *If the numerical flux is defined by Eq.* (9) *and the cell wall values take on positive density and pressure, the full second-order scheme is positivity preserving under the CFL number condition*

$$N_{\mathrm{CFL}} < \frac{1}{2}\alpha_0. \tag{14}$$

Remark 1. This theorem is similar to Theorem 1 in [12]. However, it further ensures the positivity of a second-order scheme under a fixed CFL number without any other constraint than the cell wall values.

Remark 2. For third or higher order schemes, if the numerical flux is defined with cell wall values, similar results to Theorem 1 in Perthame and Shu [12] can also be obtained. However, the positivity cannot be ensured under the conditions in Theorem 1.

Proof. For a second-order scheme, the conservative variables in the cell j can be approximated by Eq. (8). The cell average values and cell wall values always have the relation

$$\bar{U}_j = \frac{1}{2}(U^-_{j+1/2} + U^+_{j-1/2}). \tag{15}$$

Hence, Eq. (4) can be written as

$$\begin{aligned}\bar{U}^{n+1}_j = {} & \frac{1}{2}\left\{U^-_{j+1/2} - 2\lambda\left[\hat{F}(U^+_{j+1/2}, U^-_{j+1/2}) - \hat{F}(U^-_{j+1/2}, U^+_{j-1/2})\right]\right\} \\ & + \frac{1}{2}\left\{U^+_{j-1/2} - 2\lambda\left[\hat{F}(U^-_{j+1/2}, U^+_{j-1/2}) - \hat{F}(U^+_{j-1/2}, U^-_{j-1/2})\right]\right\}\end{aligned} \tag{16}$$

to give a linear combination of two first-order schemes. As these two first-order schemes have positive property while $U^-_{j+1/2}$ and $U^+_{j-1/2}$ assume positive density and pressure under the condition of Eq. (14), Eq. (16) implies that $\bar{U}^{n+1}_j$ has the same positivity property for pressure, which is indeed a concave function of U. □

5.2. Positivity property of the CE/SE method

Theorem 2. *If the cell wall values of the initial condition satisfy positive density and pressure, the CE/SE method preserves positivity without further time step restriction than Eq.* (5).

Remark 1. As the CE/SE method usually prefers and allows for large CFL number for computation, this theorem implies that the same CFL number as applied in higher internal energy region can be used in the low internal energy region of the same computational domain.

Remark 2. The positivity property for the second order in time discretization is also enforced directly by this theorem.

Proof. As the two half steps of the CE/SE method have the same difference form, we only need to prove the positivity condition for the first half step. One can find the condition is equivalent to ensuring that $\bar{U}_{j+1/2}^{n+1/2}$, $U_j^{n+1/4}$ and $U_j^{n+1/2}$ have positive density and pressure from Eqs. (10) and (12).

For Eq. (10), it can be written as the linear combination of two exact solvers

$$\bar{U}_{j+1/2}^{n+1/2} = \frac{1}{2}\bar{U}_{j+1/2} - \frac{\lambda}{2}\left[F(U_{j+1}^{n+1/4}) - F\left(\frac{1}{2}\bar{U}_{j+1/2}\right)\right] + \frac{1}{2}\bar{U}_{j+1/2} - \frac{\lambda}{2}\left[F\left(\frac{1}{2}\bar{U}_{j+1/2}\right) - F(U_j^{n+1/4})\right]. \tag{17}$$

Since the two exact solvers preserve positivity, Eq. (17) implies that, when $\bar{U}_{j+1/2}$, and $U_j^{n+1/4}$ for every node satisfy positive density and pressure, $\bar{U}_{j+1/2}^{n+1/2}$ has the same property. As Eq. (17) is calculated only for a half time step, the interaction of waves for the exact solvers can be avoided; this indicates that the CE/SE method has positivity property for the first half time step without further CFL number restriction.

For $\bar{U}_{j+1/2}$, with relations as specified in Eq. (8), we always have

$$\bar{U}_{j+1/2} = \frac{1}{4}(\bar{U}_j + U_{j+1/2}^- + \bar{U}_{j+1} + U_{j+1/2}^+), \tag{18}$$

where $U_{j\pm1/2}^{\mp}$ are the two cell wall values from the initial condition. As this equation is a linear combination of conservative variables, we need only to ensure the cell wall values satisfy positive density and pressure. On the other hand, we need $U_j^{n+1/4}$ for every node to satisfy positive density and pressure, and we also require the values $U_j^{n+1/2}$ for every node to have the same property for the linear constructions as used in the second half time step. Similarly, Eq. (12) can also be transformed into the combination of two exact solvers

$$U_j^{n+\xi} = \frac{1}{2}\bar{U}_j - \xi\lambda\left[F(U_{j+1/2}^-) - F\left(\frac{1}{2}\bar{U}_j\right)\right] + \frac{1}{2}\bar{U}_j - \xi\lambda\left[F\left(\frac{1}{2}\bar{U}_j\right) - F(U_{j-1/2}^+)\right]. \tag{19}$$

Eq. (19) has similar form as Eq. (17) and thereby ensures positivity for $U_j^{n+1/2}$ when the cell wall values $U_{j\pm1/2}^{\mp}$ satisfy positive density and pressure. As $U_j^{n+1/4}$ is equivalent to the solution of Eq. (19) at $t = (n + 1/4)\Delta t$, the condition that preserves positivity for $U_j^{n+1/2}$ also satisfies for $U_j^{n+1/4}$ for the smaller time step of $\Delta t/4$.

From the above analysis, we can conclude that (a) the CE/SE has positivity property without further CFL number restriction, (b) the positivity condition is that the cell wall values satisfy positive density and pressure. □

5.3. Energy consistency conditions

For a positive preserving scheme, as the cell average values at the nth time step ensure positive density and pressure, from Eq. (3) these conservative variables are energy consistent; that is $K > 0$. Hence, we have

$$\frac{1}{2}\frac{(\bar{\rho}\bar{u})_j^2}{\bar{\rho}_j\bar{E}_j} < 1. \tag{20}$$

From the analysis of positivity properties, one can find that besides the CFL number constraint, the cell wall values are also required to take on positive density and pressure for preserving positivity on the new cell average values. For second-order schemes, as the density is interpolated from the near cell average values with minmod, weighted average or other limiters, the cell wall density is bounded by these near cell average values and always be positive. Therefore, one only needs to maintain the energy consistency for these cell wall conservative variables, i.e.

$$\frac{1}{2}\frac{(\bar{\rho}\bar{u}_j \pm k_{\rho u}\delta)^2}{(\bar{\rho}_j \pm k_\rho\delta)(\bar{E}_j \pm k_E\delta)} < 1, \tag{21}$$

where $\delta = \Delta x/2$. If the cell wall values are energy consistent with positive density and pressure, the value at location $\pm\varepsilon$, $\delta > \varepsilon > 0$, in the cell can be written as

$$U(\pm\varepsilon) = \frac{\delta - \varepsilon}{\delta}\bar{U}_j + \frac{\varepsilon}{\delta}U^{\mp}_{j\pm1/2}. \tag{22}$$

One can then find that the conservative variables at every location inside the cell have the same property. Hence, the energy consistency condition for the whole cell is satisfied by

$$\frac{1}{2\bar{E}_j}\int_{x_{j-1/2}}^{x_{j+1/2}}\frac{(\rho u)^2}{\rho\Delta x}\mathrm{d}x < 1. \tag{23}$$

The integral on the left-hand side is the cell average kinetic energy, which can be calculated/expressed numerically by different approximations. One of the possible (approximated) energy consistency condition can be simply written as

$$\frac{1}{2}\frac{(\bar{\rho}\bar{u})_j^2 + k_{\rho u}^2\delta^2}{\bar{\rho}_j\bar{E}_j} < 1. \tag{24}$$

By comparing Eq. (24) to (20), one can find that, for a second-order scheme, the whole cell has an additional kinetic energy term, $k_{\rho u}^2\Delta x^2$, compared to that of the cell average value.

6. Kinetic energy fix method

From the discussions in the last section, one can observe that the kinetic energy is non-physical when the energy consistency condition of Eq. (21) or (23) is not satisfied. Hence, the $|k_{\rho u}|$ as utilized in Eq. (8) is larger than that of a real physical profile existing at the cell. Therefore, a kinetic energy fix can be proposed to control the kinetic energy level for a physically reasonable profile. By comparing Eq. (20) to (21) or (23), one can suggest that the magnitude of kinetic energy, density and total energy be fixed/limited to enforce or ensure energy consistency condition required for a positive pressure. In practice, the fixes can be proposed based on energy consistency condition for the cell wall values or the whole cell values. If the kinetic energy fix is based on the the cell wall values, we can limit all the $k_{\rho u}, k_\rho, k_E$ at the same time by introducing a parameter $1 > \alpha > 0$ to Eq. (21). That is, we solve for the equation

$$\frac{1}{2}\frac{(\bar{\rho}\bar{u}_j \pm \alpha k_{\rho u}\delta)^2}{(\bar{\rho}_j \pm \alpha k_\rho\delta)(\bar{E}_j \pm \alpha k_E\delta)} < 1 \tag{25}$$

to obtain α and then fix the slopes via

$$k'_\rho < \alpha k_\rho, \quad k'_{\rho u} < \alpha k_{\rho u,} \quad k'_E < \alpha k_E. \tag{26}$$

If the kinetic energy fix is based on the whole cell values, similarly a parameter $1 > \beta > 0$ can be introduced to Eq. (24) so that

$$\frac{1}{2}\frac{(\bar{\rho}\bar{u})_j^2 + \beta^2 k_{\rho u}^2\delta^2}{\bar{\rho}_j\bar{E}_j} < 1. \tag{27}$$

We solve for β and the slopes are fixed via

$$k'_\rho < \beta k_\rho, \quad k'_{\rho u} < \beta k_{\rho u,} \quad k'_E < \beta k_E. \tag{28}$$

Even as Eqs. (26) and (28) suggest the bounds for the slopes, usually a slightly larger α or β, i.e. $1.1\alpha \sim 1.3\alpha$ or $1.1\beta \sim 1.3\beta$, can be used to achieve even better results. It is noted that the above fixes, Eqs. (26) and (28), are only applied to those cells which did not satisfy the energy consistency conditions caused by interpolations for non-physical profile. The accuracy in the smooth region is still preserved. It is also noted that there may be other kinetic fix methods derived from Eq. (21) and (23); for example one of the alternative fix method is further outlined in Appendix A. However, our numerical results show the present simple fix method is sufficient for a large range of low internal energy flows even when a vacuum occurs.

As the kinetic energy fix methods from Eqs. (25)–(28) are not scheme specific and many second-order conservative schemes have a linear slope prediction step, the said methods may be applicable to different schemes easily. The detailed procedure can be given as follows:

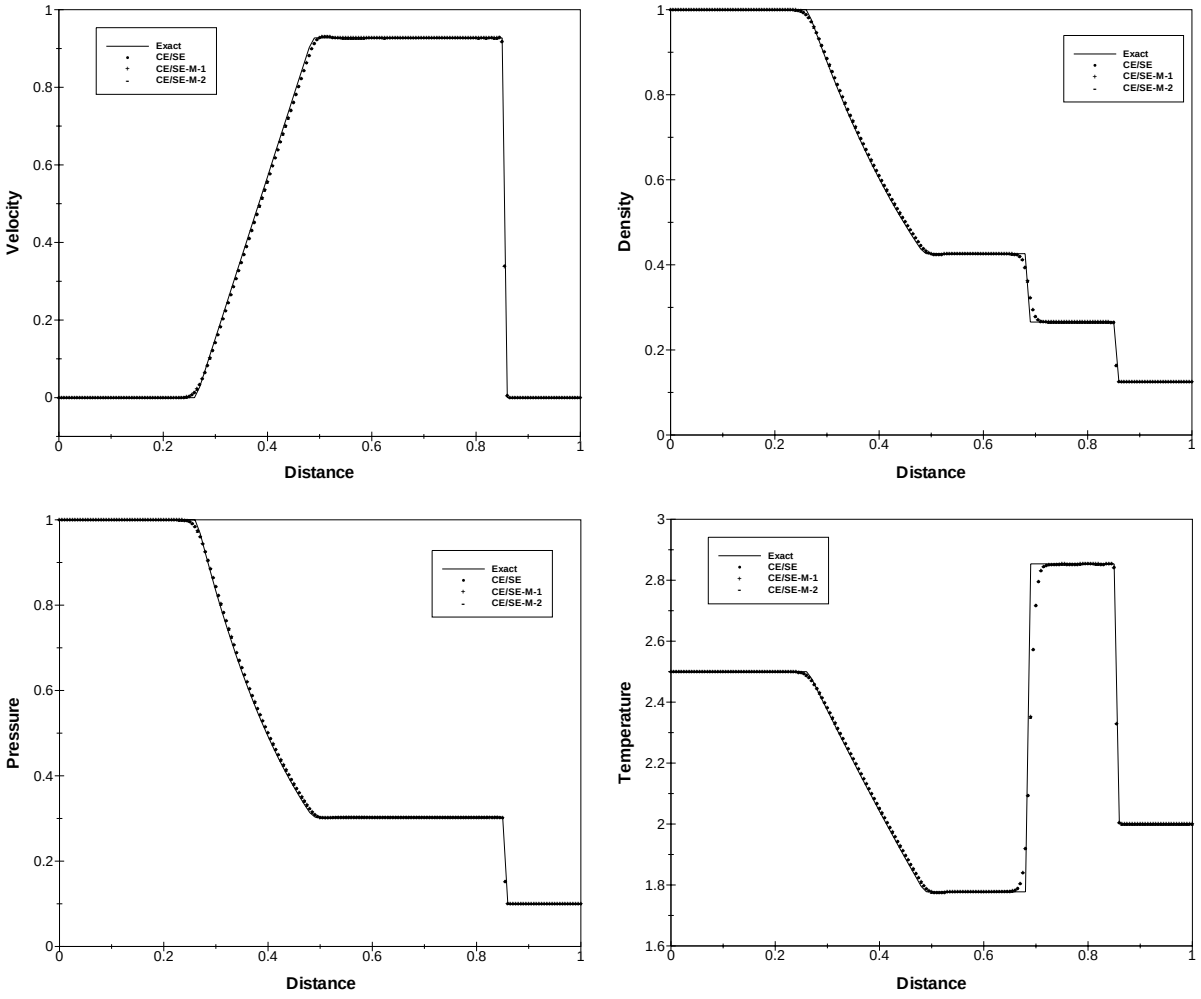

Fig. 2. Sod's problem at $t = 0.2$.

1. Give a further time step restriction by Eq. (14) if needed.
2. Compute the average conservative variables $(\bar{\rho}_j, (\bar{\rho u})_j, \bar{E}_j)$ of the cell, and derivatives k_ρ, $k_{\rho u}$ and k_E at the cell center.
3. Check the energy consistency by Eq. (21) or (24).
4. Modify the derivatives at the cell center by Eq. (26) or (28) if needed.
5. Use the modified derivatives k'_ρ, $k'_{\rho u}$ and k'_E for the new time or half time step.

7. Numerical examples

As an application, the kinetic energy fix method is implemented into the CE/SE method. The following examples, including Sod's problem, Lax problem and "1–2–3 problem" series, illustrate the compatibility of the kinetic energy fixed CE/SE method (denoted as CE/SE-M-1 for Eq. (26) and CE/SE-M-2 for Eq. (28)) with its original form and further demonstrate the ability of the said method to handle the potential difficulties of low internal energy flows. For all the test cases, the number of grid points is 200 and the referenced exact solution is sampled on 100 grid points.

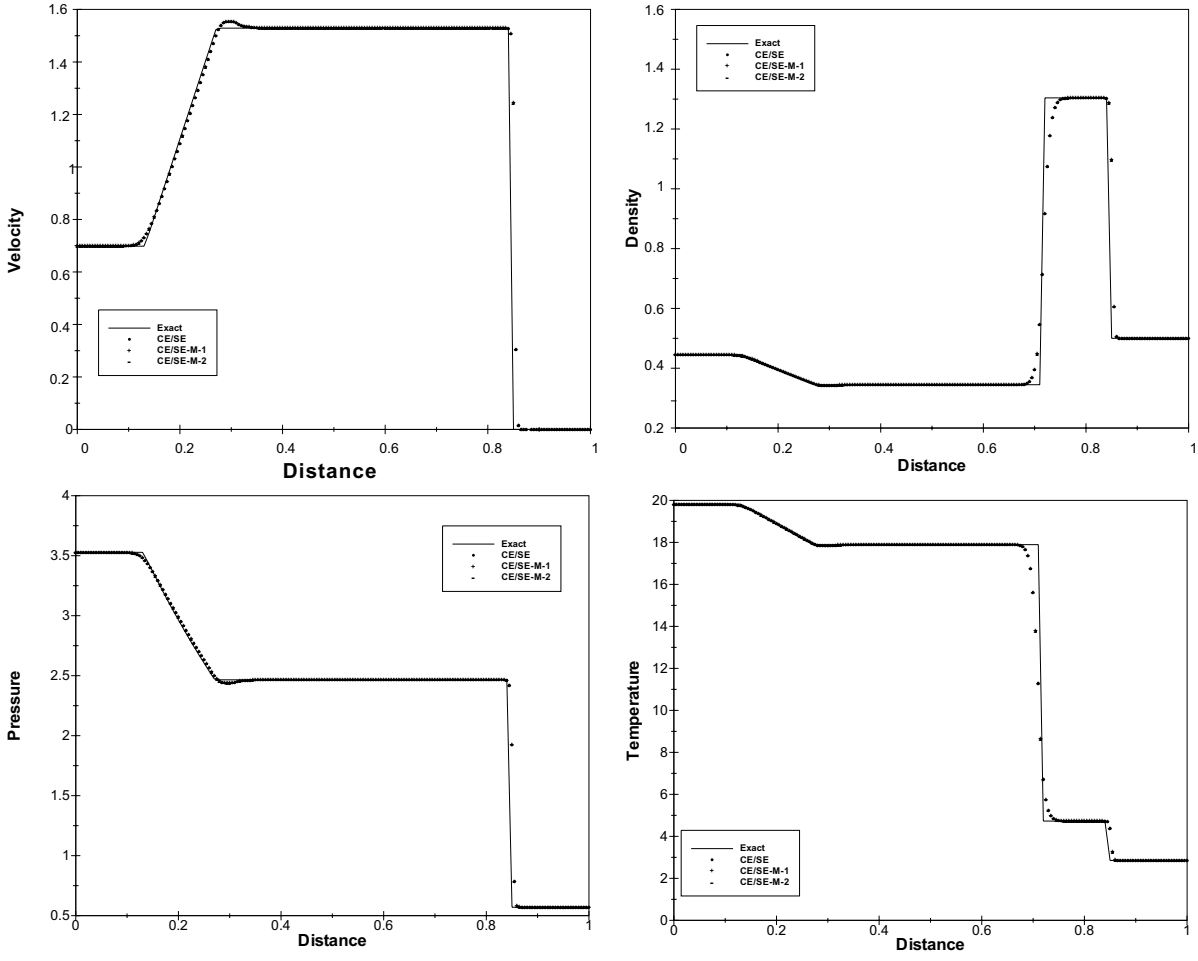

Fig. 3. Lax's problem at $t = 0.14$.

7.1. Compatibility of the kinetic energy fixed method

Here, we show the results for the two well-known problems with shock waves, rarefaction waves and contact discontinuities. This is to test the compatibility of the kinetic energy fixed CE/SE method. The first problem is due to Sod [16]. The initial conservative variables are

$$U(x,0) = \begin{cases} (1,0,2.5) & \text{if } x < 0.5, \\ (0.125,0,0.3125) & \text{if } x > 0.5. \end{cases} \tag{29}$$

The second one is the Riemann problem proposed by Lax [8]. The initial conservative variables are

$$U(x,0) = \begin{cases} (10.445,0.311,8.928) & \text{if } x < 0.5, \\ (0.5,0,1.4275) & \text{if } x > 0.5. \end{cases} \tag{30}$$

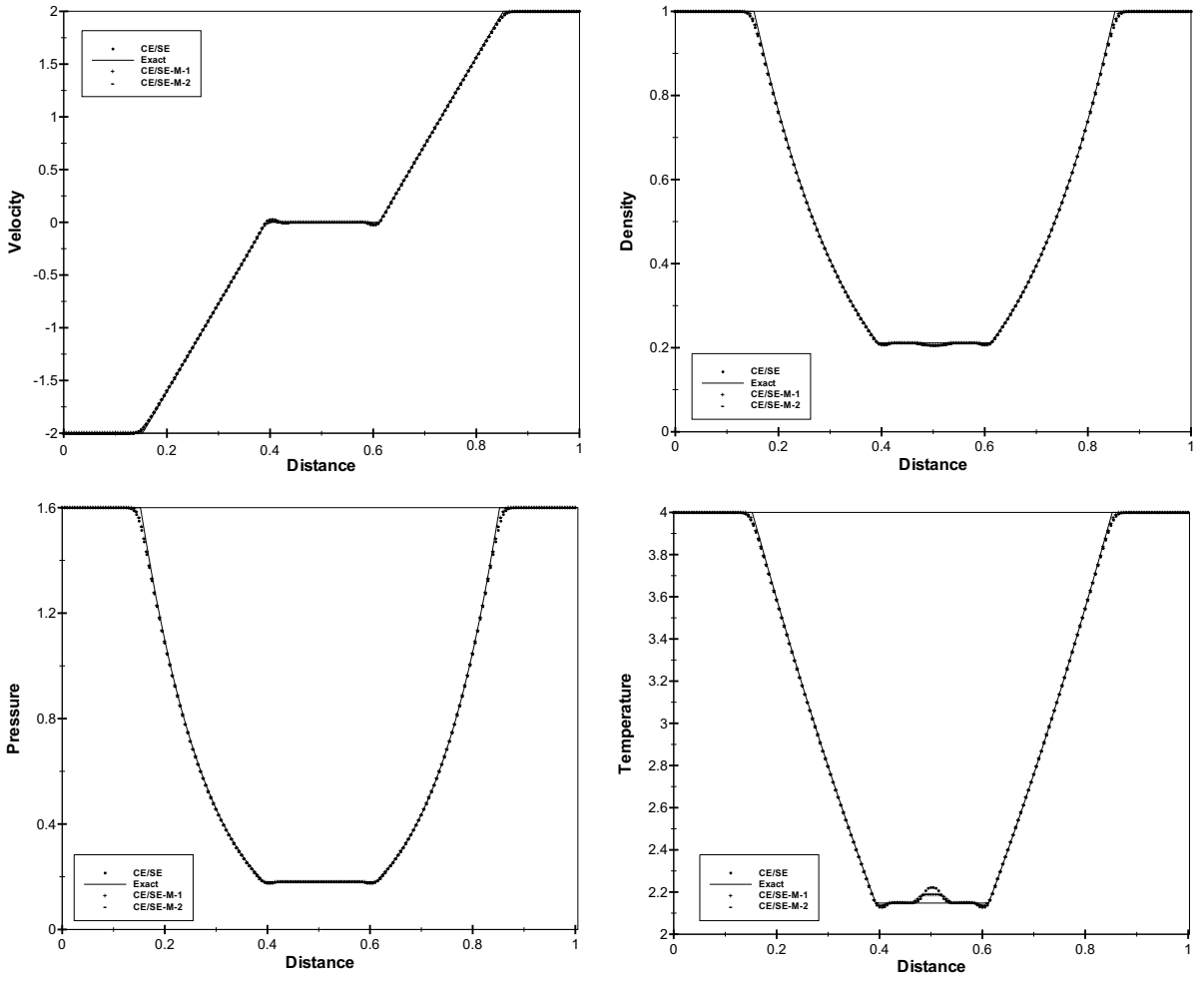

Fig. 4. $E_0 = 6/K_0 = 2$ at $t = 0.1$.

These two problems, especially Lax's problems, are considered as benchmark tests. The numerical results for both the CE/SE method with and without kinetic energy fix, denoted by CE/SE, CE/SE-M-1 and CE/SE-M-2, are presented in Figs. 2 and 3. Even for such non-characteristic-based schemes, one can see that all the methods give identical solutions with good and reasonable accuracy.

We have also tested Woodward and Colella's case [19], Shu–Osher's case [14] and the slowly moving shock problem of Quirk [13]. As with the high internal energy ratio in these problems, CE/SE, CE/SE-M-1 and CE/SE-M-2 always give identical and correct solutions. These results are not presented here.

7.2. Low internal energy flows

As in Section 4, we also use the "1–2–3 problem" series to test the ability of the kinetic fixed CE/SE method for low internal energy flows. Among the MUSCL, ENO, WENO and CE/SE schemes without kinetic energy fix, the CE/SE method is found to be the most effective. Therefore, the comparisons presented next are mainly confined to between the solutions of the fixed and original CE/SE method.

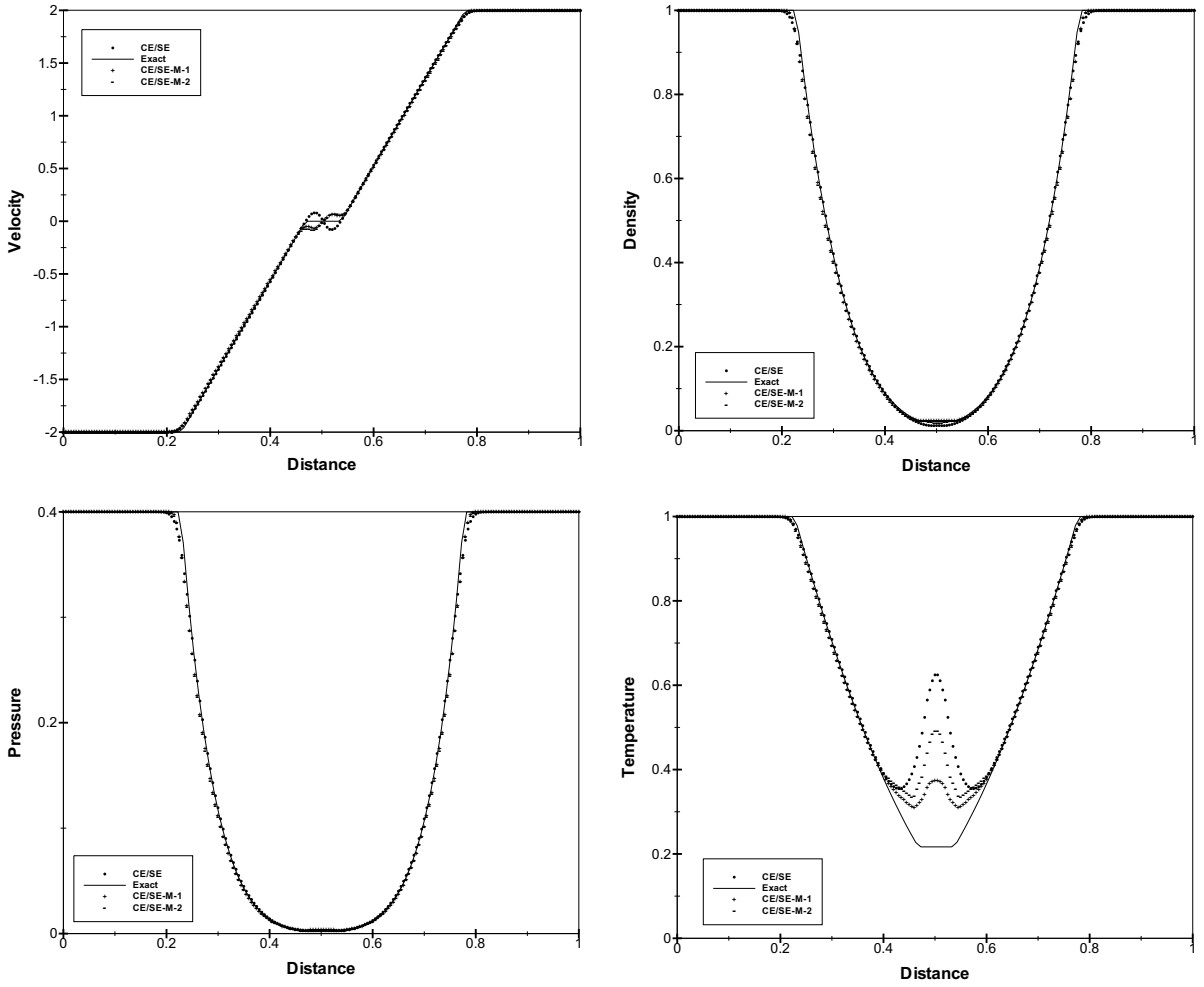

Fig. 5. $E_0 = 3/K_0 = 0.5$ at $t = 0.1$.

7.2.1. $E_0 = 6/K_0 = 2$

In Section 4, one can find the original CE/SE works with CFL number 0.8. With a CFL number of 0.9, the results of the fixed method are shown in Fig. 4.

By comparing the solution of the original CE/SE, one can find the solutions of the both original and fixed methods are almost identical. Furthermore, one can also find that both the CE/SE-M-1 and CE/SE-M-2 give better solutions, such as the less non-physical temperature increase at the center of the domain.

7.2.2. $E_0 = 3/K_0 = 0.5$

This is a test of increasing difficulty with extreme condition where the MUSCL scheme 'explodes' even with a very small CFL number of 0.1. For the original CE/SE method, it can still compute at a CFL number of 0.6, but gives rise to serious oscillations for the velocity profile. Only by reducing the CFL number to 0.4, the oscillations are decreased to a more manageable situation, as shown in Fig. 5. For the CE/SE-M-1 and CE/SE-M-2, the problem can be solved with a CFL number of 0.9. The numerical results

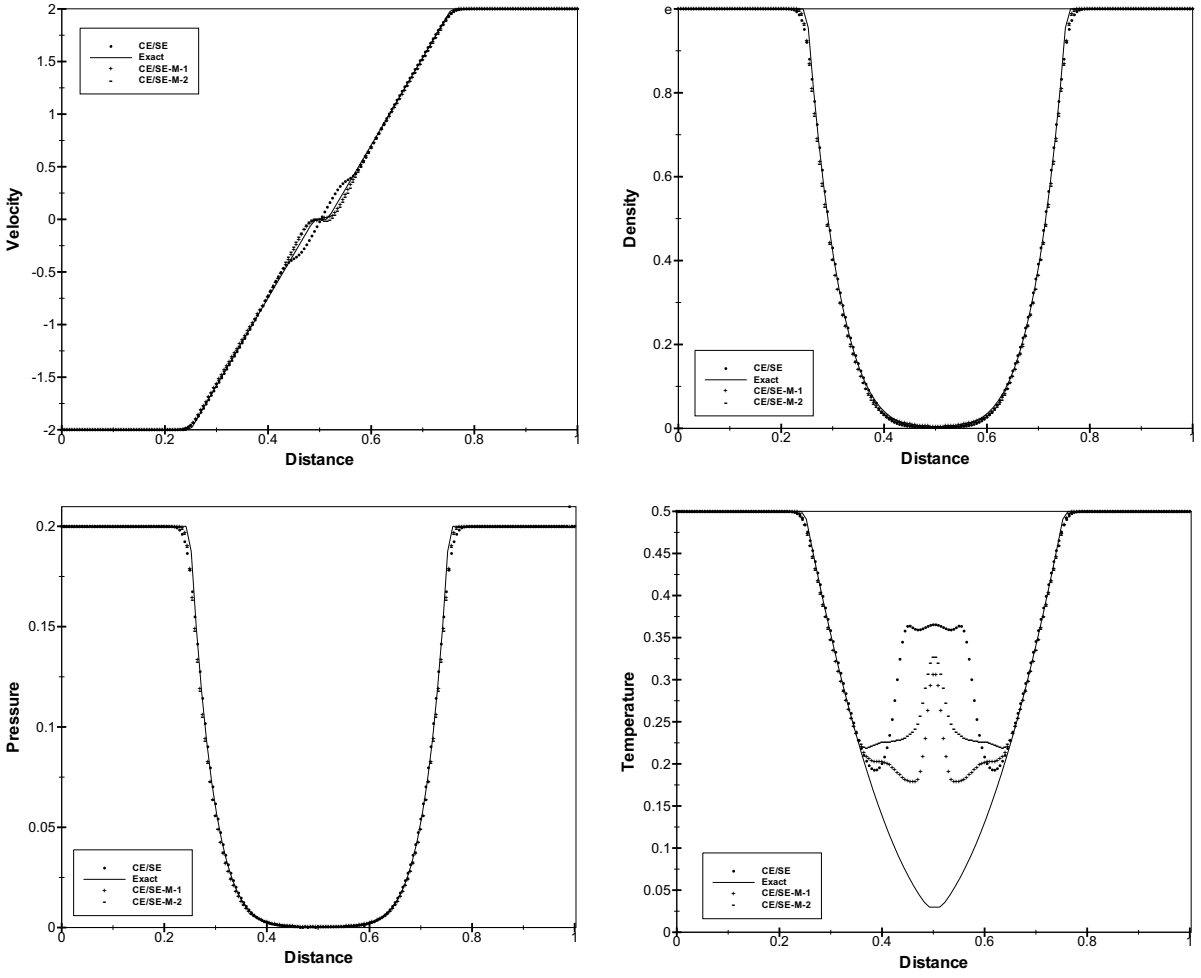

Fig. 6. $E_0 = 2.5/K_0 = 0.25$ at $t = 0.1$.

in Fig. 5 show that the fixed methods produce accurate density profile as compared to the over-prediction by the original method. It may be noted that the calculated density profiles with kinetic energy fix are even better that those obtained by second-order positivity preserving kinetic schemes [4,10]. The velocity profiles are computed without oscillation in contrast to the result by the original CE/SE method. However, as this case is considered as one with extreme condition, the fixed CE/SE method still predicts a high non-physical temperature profile in the central region; it presents a peak temperature magnitude of close to 0.4 and 0.5 for the CE/SE-M-1 and CE/SE-M-2, respectively. Even then, these results are better than that of the ENO-LF and WENO-LF schemes, in which the temperature is higher than 0.8 (not shown here).

7.2.3. $E_0 = 2.5/K_0 = 0.25$

As the initial internal energy ratio is 0.25, the exact solution for this problem is getting very close to that with a vacuum solution. A CFL number of 0.4 is needed for the original CE/SE method while the fixed method can still compute with a CFL number of 0.9. The results in Fig. 6 show a more accurate velocity profile without oscillation and a smaller non-physical temperature increase are obtained with both the CE/SE-M-1 and CE/SE-M-2. On the other hand, as larger time step is permitted, the fixed method is able to

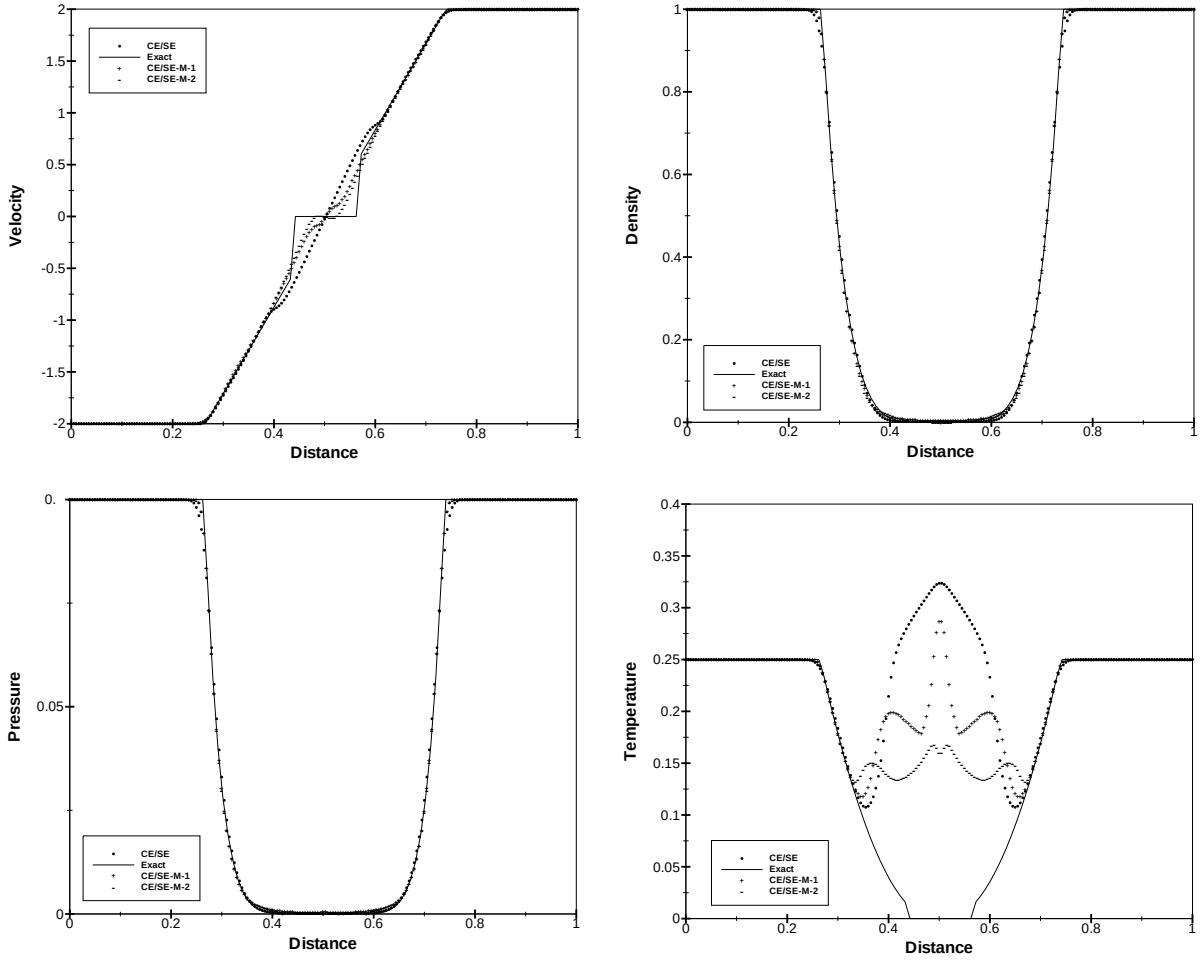

Fig. 7. $E_0 = 2.25/K_0 = 0.125$ at $t = 0.1$.

compute for the rarefaction waves more accurately especially near the rarefaction wave ends where density and pressure decrease with large gradient (see Fig. 6).

7.2.4. $E_0 = 2.25/K_0 = 0.125$

This is the most critical test case among all the cases presented here. As the initial energy ratio is only 0.125, a vacuum occurs in the solution. Without kinetic energy fix, this case can only be computed with a CFL number of 0.3 by the original CE/SE method, and the computation can even not proceed with a CFL number of 0.9 for the most robust ENO and WENO schemes. For the fixed method, the computation can proceed without difficulty at the CFL number of 0.9. The results, as seen in Fig. 7, still enable a more accurate density and pressure profiles in the rarefaction wave regions. However, as a vacuum occurs, the solution at the center of the domain have strictly no physical meaning. For example, the zero velocity at the domain center results from the absence of gaseous material rather than the flow having a zero velocity.

8. Concluding remarks

For kinetic flows with low internal energy ratio, the high order conservative schemes may not ensure positive internal energy as do for their first-order forms or produce inaccurate results in contrast to flows with higher internal energy. Based on analysis of the positivity property for high order schemes for the cases which permit fixed large CFL number restriction, we have obtained the energy consistency conditions required for the existence of positive pressure for second-order Riemann-solver type schemes and CE/SE method. A kinetic energy fix method applicable for a general second-order scheme is proposed to ensure compliance of the energy consistency conditions for the flow problems with low internal energy. The results of the numerical examples show that the kinetic energy fix CE/SE method exhibits greater robustness and accuracy when computing flows with very low internal energy or even a vacuum. As the kinetic energy fix is only employed when the energy consistency conditions are violated for the particular cell wall or whole cell values, it is able to preserve uniform second-order accuracy for other regions.

Appendix A. An alternative kinetic energy fix method

If the energy consistency condition of Eq. (23) is not satisfied thereby indicating a non-physical profile, one can suggest a parameter

$$W = \frac{1}{2\bar{E}_j}\int_{x_{j-1/2}}^{x_{j+1/2}} \frac{(\rho u)^2}{\rho \Delta x}\mathrm{d}x \tag{A.1}$$

to determine the extent or measurement of this departure when $W \geqslant 1$. Hence W can be used for limiting the kinetic energy quantity. Here, the kinetic energy fix is proposed as

$$k_{\mathrm{fix}} = \frac{k}{1+\alpha' W^{\beta'}}, \tag{A.2}$$

where k and k_{fix} are the original and the fixed slope for interpolation, α' and β' are two positive parameters. The kinetic energy fix makes the effective $|k|$ smaller to satisfy a possible real physical profile especially when W is close to 1 or larger. The parameter α' determines the influence of W on $|k|$ while the parameter β' controls the fixing rate. For many flow phenomena, the ratio of kinetic energy to total energy W is much smaller than 1. For such cases, it is found that $|k|$ is almost not affected by Eq. (A.2). Usually, only $k_{\rho u}$ is

fixed by Eq. (A.2) and is sufficient. Sometimes, such as when a near vacuum solution occurs, k_ρ, $k_{\rho u}$ and k_E are all fixed simultaneously in order to achieve even better energy consistency properties. For the CE/SE method, when Eq. (A.1) is approximated as

$$W = \frac{1}{2}\frac{(\bar{\rho}\bar{u})_j^2 + ak_{\rho u}^2\delta^2}{\bar{\rho}_j\bar{E}_j}, \tag{A.3}$$

where $a = 1/2$ and $\delta = \Delta x/2$, it is found that setting $\alpha' = 0.1$ and $\beta' = 3.0$ can give reasonable results for a large range of internal to kinetic energy ratio and applicable to those examples in Section 7 (results not shown here).

References

[1] S.C. Chang, The method of space–time conservation element and solution element – a new approach for solving the Navier–Stokes and Euler equations, J. Comput. Phys. 119 (1995) 295.
[2] J.P. Cocchi, R. Saurel, J.C. Loraud, B. Sjogreen, Some remarks about the resolution of high velocity flows near low densities, Shock Waves 8 (1998) 119.
[3] B. Einfeldt, C.D. Munz, P.L. Roe, B. Sjogreen, On Godunov-type methods near low densities, J. Comput. Phys. 92 (1991) 273.
[4] J.L. Estivalezes, P. Villedieu, High-order positivity-preserving kinetic schemes for compressible Euler equations, SIAM J. Numer. Anal. 33 (1996) 2050.
[5] J. Gressier, P. Villedieu, J.M. Moschetta, Positivity of flux vector splitting schemes, J. Comput. Phys. 155 (1999) 199.
[6] A. Harten, P.D. Lax, B. van Leer, On upstream differencing and Godunov-type schemes for hyperbolic conservation laws, SIAM Rev. 23 (1983) 35.
[7] G.S. Jiang, C.W. Shu, Efficient implementation of weighted ENO schemes, J. Comput. Phys. 126 (1996) 202.
[8] P.D. Lax, Weak solution of non-linear hyperbolic equations and their numerical computations, Commun. Pure Appl. Math. 7 (1954) 235.
[9] T. Linde, P.L. Roe, On multidimensional positively conservative high-resolution schemes, in: V. Venkatakrishnan, M.D. Salas, S.R. Chakravarthy (Eds.), Barriers and Challenges in Computational Fluid Dynamics, Kluwer Academic Publishers, Dordrecht Hardbound, 1997.
[10] B. Khobalatte, B. Perthame, Maxmum principle on the entropy and second-order kinetic schemes, Math. Comput. 62 (1994) 119.
[11] B. Perthame, Second-order Boltzmann schemes for compressible Euler equations in one and two space dimensions, SIAM J. Numer. Anal. 29 (1992) 1.
[12] B. Perthame, C.W. Shu, On positivity preserving finite volume schemes for Euler equations, Numer. Math. 73 (1996) 119.
[13] J.J. Quirk, An alternative to unstructured grids for computing gas dynamics flows around arbitrarily complex two dimensional bodies, Comput. & Fluids 23 (1994) 125.
[14] C.W. Shu, S. Osher, Efficient implementation of essentially non-oscillatory shock-capturing schemes, II, J. Comput. Phys. 83 (1989) 32.
[15] J. Smoller, Shock Waves and Reaction–diffusion Equations, Springer, Berlin, 1983.
[16] G.A. Sod, A survey of several finite difference methods for systems of non-linear hyperbolic conservation laws, J. Comput. Phys. 27 (1978) 1.
[17] T. Tang, K. Xu, Gas-kinetic schemes for the compressible Euler equations: Positivity-preserving analysis, Z. Angew. Math. Phys. 50 (1999) 258.
[18] E.F. Toro, Riemann Solvers and Numerical Methods for Fluid Dynamics, Springer, Berlin, 1997.
[19] P. Woodward, P. Colella, The numerical simulation of two-dimensional fluid flow with strong shocks, J. Comput. Phys. 54 (1984) 115.