



# A Characteristic-Featured Shock Wave Indicator for Conservation Laws Based on Training an Artificial Neuron

Yiwei Feng<sup>1</sup> · Tiegang Liu<sup>2</sup> · Kun Wang<sup>3</sup>

Received: 9 May 2019 / Revised: 29 January 2020 / Accepted: 17 March 2020 / Published online: 9 April 2020  
© Springer Science+Business Media, LLC, part of Springer Nature 2020

## Abstract

In this work, we use exact solutions of one-dimensional Burgers equation to train an artificial neuron as a shock wave detector. The expression of the artificial neuron detector is then modified into a practical form to reflect admissible jump of eigenvalues. We show the working mechanism of the practical form is consistent with compressing or intersecting of characteristic curves. In addition, we prove there is indeed a discontinuity inside the cell detected by the practical form, and smooth extrema and large gradient regions are never marked. As a result, we apply the practical form to numerical schemes as a shock wave indicator with its easy extension to multi-dimensional conservation laws. Numerical results are presented to demonstrate the robustness of the present indicator under Runge–Kutta Discontinuous Galerkin framework, its performance is generally compared to TVB-based indicators more efficiently and accurately. To treat the initial inadmissible jumps, including linear contact discontinuities and those evolving into rarefaction waves, a preliminary strategy of combining a traditional indicator in the beginning with the present indicator is suggested. We believe the present indicator can be applied to unstructured mesh in the future.

**Keywords** Hyperbolic conservation laws · Characteristic curve · Discontinuous Galerkin · Troubled-cell indicator · Artificial neural networks

**Mathematics Subject Classification** 65M60 · 68Q32

---

Submitted to the editors on May 9 2019.

---

✉ Tiegang Liu  
liutg@buaa.edu.cn

Yiwei Feng  
fengyw@buaa.edu.cn

<sup>1</sup> LMIB and School of Mathematical Sciences and Shen Yuan Honors College of Beihang University, Beijing, China

<sup>2</sup> School of Mathematical Sciences of Beihang University, Beijing, China

<sup>3</sup> School of Mathematical Sciences and School of Aeronautic Science and Engineering of Beihang University, Beijing, China

# 1 Introduction

At present, high order high resolution schemes, such as Discontinuous Galerkin (DG) [3–7], have been widely applied in numerical simulation of hyperbolic conservation laws. Since even under smooth initial conditions, solutions to conservation laws can develop discontinuities in a finite time [8], spurious oscillations occur in the regions of discontinuities when such high order schemes are applied. There are two steps in handling spurious oscillations. The first one is to detect cells where the solution loses regularity (troubled-cells), the other is to use correction techniques to eliminate spurious oscillations for the troubled-cells. There are several correction techniques in literature, such as artificial limiting [5,6,36], reconstruction techniques [13,24,26,27,31] and artificial viscosity [14,15,34]. Those correction techniques often lead to inefficiencies of numerical simulation and even deterioration of accuracy. In other word, the accuracy and regularity of numerical solution depend largely on effectiveness of the troubled-cell indicator applied. For example, the popular DG methods usually use minmod function [12] to detect troubled-cells, and simultaneously apply it as a limiter to eliminate oscillations, that is  $\tilde{u}_i^{(mod)} = \text{minmod}(\tilde{u}_i, \Delta_+u_i^{(0)}, \Delta_-u_i^{(0)}) =$

$$\begin{cases} s \cdot \min\{\tilde{u}_i, \Delta_+u_i^{(0)}, \Delta_-u_i^{(0)}\} & \text{if } \text{sign}(\tilde{u}_i) = \text{sign}(\Delta_+u_i^{(0)}) = \text{sign}(\Delta_-u_i^{(0)}) = s \\ 0 & \text{otherwise,} \end{cases}$$

where  $u_i^{(0)}$  is the 0<sup>th</sup> moment degree of freedom(dof), and  $\tilde{u}_i$  is the interface collection of higher moment dofs. As one can observe, cells near extrema are indicated as troubled-cells with a minmod limiter. This leads to decreasing of scheme accuracy. In order to preserve high order of the schemes, Total Variation Bounded (TVB) limiter [6] are more popular, that is  $\tilde{u}_i^{(mod)} = \begin{cases} \tilde{u}_i & \text{if } |\tilde{u}_i| \leq Mh^2 \\ \text{minmod}(\tilde{u}_i, \Delta_+u_i^{(0)}, \Delta_-u_i^{(0)}) & \text{otherwise} \end{cases}$  However, the parameter M was found to depend heavily on problems and has a considerable impact on numerical solutions. In [25], Qiu and Shu shared performance of various TVB limiters under different selection of M. To improve efficiency, several procedure such as Weighted Essential Non-Oscillating (WENO) based reconstruction [27], KXRCF shock indicator [16] were proposed to improve the effectiveness of troubled-cell indication. Those indicators or procedure mentioned above mainly detect troubled-cells from aspects of numerical oscillations.

Recently, Ray and Hesthaven proposed artificial neural networks to implement troubled-cell indication under RKDG frameworks [28,29]. They obtained a multi-layer perceptron (MLP) troubled-cell indicator. Numerical results presented in their paper demonstrated the MLP indicator performs generally better than TVB-based indicator. In this work, we employ exact solutions of the Burgers equation to train a concise artificial neuron (AN) with only one linear hidden layer for shock wave detecting, fortunately obtain a convergent AN detector, and it can be modified into a practical form to reflect admissible jump of eigenvalues. The advantage of the practical form is threefold: (a) it has a concise expression involving only cell-averages; (b) it is explicable, and can theoretically guarantee the only detection of discontinuities caused by characteristic curves compressing or intersecting; (c) it can handily be extended to multi-dimensional conservation laws.

The rest of the paper is structured as follows. In Sect. 2, we briefly summarize the theory of characteristic curves for one-dimensional scalar and system conservation laws. In Sect. 3, we illustrate the complete training process of the shock detector and give the expression of the trained AN detector. In Sect. 4, we modify the AN detector into a practical form and use the practical form to analyze its working mechanism. The practical form is then applied as a shock wave indicator and extended to system conservation laws. RKDG frameworks are

briefly introduced in Sect. 5 and we show how TVB-based indicators and present indicator detect troubled-cells under RKDG frameworks. Several numerical results are presented in Sect. 6 to demonstrate the capability of the present shock wave indicator, as compared to TVB indicators. We make a few concluding remarks in Sect. 7.

## 2 Preliminaries of Hyperbolic Conservation Laws

### 2.1 One-Dimensional Scalar Conservation Laws

The one-dimensional scalar conservation law with an initial condition is given by

$$\begin{cases} \frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0 \\ u(x, 0) = u_0(x). \end{cases} \tag{2.1}$$

For Eq. (2.1), we summarize the following useful conclusions without proof. Please refer to [2,35] for more details.

**Lemma 2.1** *Assuming the initial value function  $u(x, 0) = u_0(x)$  is smooth, and denoting  $\lambda(u) = f'(u)$ , then we have following properties for Eq. (2.1):*

- (1) *The characteristic curves in Eq. (2.1) are straight lines with the form of  $x = x_0 + \lambda(u_0(x_0))t$ , and the slope of characteristic curve is  $\lambda(u) = f'(u)$ . The solution remains unchanged along each characteristic curve.*
- (2) *The solution can be obtained by solving “ $u$ ” in the following implicit function  $u = u_0(x - \lambda(u)t)$ .*
- (3) *The solution can evolve into discontinuities even under a smooth initial condition due to the intersection of characteristic lines.*
- (4) *Across the discontinuity, the solution satisfies the Rankine–Hugoniot jump relation  $s[u] = [f]$ , where  $s = \frac{dx_s(t)}{dt}$  represents the speed of the discontinuity and  $[u]$  represents the jump value of variable  $u$  when across the discontinuity.*

It is necessary to deal with non-smooth initial conditions. For the Riemann problem of Eq. (2.1), we list some properties of solution without proof.

**Lemma 2.2** *The Riemann problem for one-dimensional scalar conservation laws is given as follows,*

$$\begin{cases} \frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0 \\ u(x, 0) = \begin{cases} u_l & \text{if } x < 0 \\ u_r & \text{if } x \geq 0. \end{cases} \end{cases} \tag{2.2}$$

*We can conclude that, if  $f(u)$  is convex,*

- (1) *If  $\lambda(u_l) > \lambda(u_r)$ , characteristic curves of two constant regions intersect, and the solution evolves into a discontinuity.*
- (2) *If  $\lambda(u_l) < \lambda(u_r)$ , characteristic curves of two constant regions diverge, and the solution evolves into a centered rarefaction wave connecting two constant regions.*

### 2.2 One-Dimensional System Conservation Laws

One-dimensional system conservation laws with an initial condition are given by

$$\begin{cases} \frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}(\mathbf{U})}{\partial x} = 0 \\ \mathbf{U}(x, 0) = \mathbf{U}_0(x). \end{cases} \tag{2.3}$$

Assuming  $\mathbf{U} \in \mathbb{R}^m$ , denoting  $\mathbf{A}(\mathbf{U}) = \frac{\partial \mathbf{F}}{\partial \mathbf{U}}$ , system (2.3) is called hyperbolic if matrix  $\mathbf{A}(\mathbf{U})$  can be diagonalized,  $\mathbf{A} = \mathbf{R}\Lambda\mathbf{R}^{-1}$  with real eigenvalues, here,  $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_m)$  is the eigenvalue matrix, and  $\mathbf{R} = (\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_m)$  is the corresponding right eigenvector matrix. Each eigenvalue  $\lambda_i$  of  $\mathbf{A}(\mathbf{U})$  defines a characteristic field  $\lambda_i$ -field, if  $\nabla \lambda_i(\mathbf{U}) \cdot \mathbf{R}_i(\mathbf{U}) \neq 0$ , the  $\lambda_i$ -field is genuinely nonlinear; if  $\nabla \lambda_i(\mathbf{U}) \cdot \mathbf{R}_i(\mathbf{U}) = 0$ , the  $\lambda_i$ -field is linearly degenerated.

We summarize some useful conclusions without proof. More details can be found in [35].

**Lemma 2.3** *The solution,  $\mathbf{U}(x, t)$ , holds following properties:*

- (1) *The  $i$ -characteristic curves in system (2.3) are locally straight lines with the form of  $x = x_0 + \lambda_i(\mathbf{U}_0(x_0)) \cdot t$ , and the slope of  $i$ -characteristic curve is the  $i$ -eigenvalue  $\lambda_i$ .*
- (2) *The solution can evolve into discontinuities even under a smooth initial condition due to the intersection of a certain  $i$ -characteristic curves.*
- (3) *Across the discontinuity, the solution  $\mathbf{U}$  satisfies the Rankine-Hugoniot jump relation,  $s[\mathbf{U}] = [\mathbf{F}]$ , where  $s = \frac{dx_s(t)}{dt}$  represents the speed of the discontinuity and  $[\mathbf{U}]$  represents the jump value of  $\mathbf{U}$  across the discontinuity.*

In this work, we pay attention to the Euler equation, the Riemann problem for the Euler equation is given as follows,

$$\begin{cases} \frac{\partial}{\partial t} \begin{bmatrix} \rho \\ \rho u \\ E \end{bmatrix} + \frac{\partial}{\partial x} \begin{bmatrix} \rho u \\ p + \rho u^2 \\ (E + p)\rho \end{bmatrix} = 0 \\ (\rho_0, u_0, p_0)^T(x) = \begin{cases} (\rho_l, u_l, p_l)^T & \text{if } x < 0 \\ (\rho_r, u_r, p_r)^T & \text{if } x \geq 0, \end{cases} \end{cases} \tag{2.4}$$

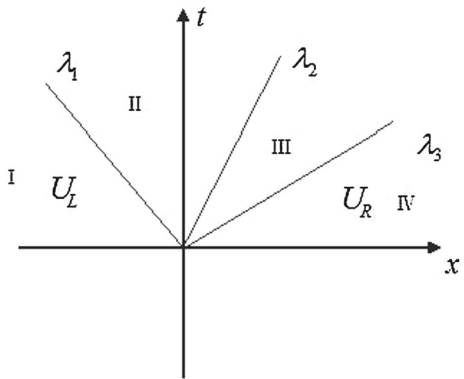
where  $\rho, u, p$  represent the fluid density, velocity, and pressure, respectively. The quantity  $E$  is the total energy per unit volume, which is given as  $E = \rho(e + \frac{u^2}{2})$ , where  $e$  is the specific internal energy given by a caloric equation of state,  $e = e(\rho, p)$ . We choose the equation of state for ideal gas given by  $e = \frac{p}{(\gamma-1)\rho}$ , with  $\gamma = C_p/C_v$  denoting the ratio of specific heats. We set  $\gamma = 1.4$  for all the test cases in this work.

Again we summarize some useful conclusions for system (2.4) as follows.

**Lemma 2.4** *The solution holds the following properties:*

- (1) *Based on properties in Lemma 2.3, the solution can be divided into 4 constant regions by 3  $\lambda_i$ -characteristic fields as shown in Fig. 1.*
  - $\lambda_1 = u - a$ ,  $\lambda_1$ -field is genuinely nonlinear, area I and area II are separated by 1-elementary wave (shock wave or centered rarefaction wave), where  $u$  is the fluid velocity and  $a = \sqrt{\frac{\gamma p}{\rho}}$  represents the sound speed.

**Fig. 1** In the 1-D Euler equations, the space is divided into 4 constant regions by 3 characteristic fields



- $\lambda_2 = u$ ,  $\lambda_2$ -field is linearly degenerated, area II and area III are separated by 2-contact discontinuity, the pressure  $p$  and velocity  $u$  keep unchanged across the  $\lambda_2$ -field, the density  $\rho$  allows a jump across the  $\lambda_2$ -field.
- $\lambda_3 = u + a$ ,  $\lambda_3$ -field is genuinely nonlinear; area III and area IV are separated by 3-elementary wave (shock wave or centered rarefaction wave).

(2) According to (1) and Lemma 2.3(2), we can give a criterion of discontinuity for system (2.4) as follows,

$$u_l - a_l > u_r - a_r, \tag{2.5a}$$

$$u_l + a_l > u_r + a_r. \tag{2.5b}$$

The inequality Eq. (2.5a) indicates the solution includes either a 1-shock wave or a 2-contact discontinuity or both. The inequality Eq. (2.5b) indicates the solution consists of either a 3-shock wave or a 2-contact discontinuity or both.

*Remark* Although the indicator to be developed is for shock wave detecting in this work, the Lemma 2.4(2) implies that the shock wave indicator is also capable of detecting contact discontinuities as shown in Sect. 4.

In Sect. 4, we will analyze mechanism of the shock wave detector is consistent with compressing or intersecting of characteristic curves mentioned in this section.

### 3 An Artificial Neuron for Shock Wave Indicating

In this section, our target is to explore an approximation of unknown function which can reflect the characteristics of shock wave discontinuities. Since we have no prior information about the intrinsic features of the indicator, an Artificial Neural Network (ANN) [11] becomes a suitable method to learn these undiscovered features. Generally, an ANN is designed to imitate the learning procedure of the complex biological network. With the rapid development of computing technology, ANNs have gradually evolved into more matured deep neural networks (DNNs) [30]. Except for an input layer and an output layer, the structure of an ANN consists of at least one hidden layer. The nodes of ANN are fully connected linearly, which means that each node in one layer connects with a certain weight to every node in the neighboring layer. For approximating a complex nonlinear function, the neurons adopt nonlinear activation function. Among the activation functions, Rectified Linear Units (ReLU)

function [19] ( $ReLU(x) = \max(0, x)$ ) is popularly used in hidden layers, logistic sigmoid function [23] ( $sigmoid(x) = \frac{1}{1+e^{-x}}$ ) is often used as the activation function of output layer. After an ANN is built, a large amount of tagged data is acquired to train the weights in ANN, the weights are acceptable if the output of ANN and the label of data are close enough under the measure of a certain loss function [21]. Cross entropy function [20] is often selected as the loss function.

In this work, to analyze the working mechanism easily, we initially build an artificial neuron (an ANN with only one linear hidden layer) as a shock wave indicator in conservation laws, and train weights via TensorFlow [1]. The complete training process for the AN model includes generating the training and validation sets, selecting the neuron structure, and training weights using error back propagation algorithm [22]. In this section, we will illustrate the whole process of generating the training and validation data set in detail. The process includes selecting training cases, setting the input stencil of the data, and setting the label for each data.

### 3.1 Training and Validation Set

#### 3.1.1 Selection of Training Cases

We provide several functions as initial conditions to include shock wave discontinuities with various scales of jump values in the one-dimensional nondimensionalized Burgers equation, so that we can obtain a complete data set as well as possible. Notice for the Burgers equation, contact discontinuities are not included in the data set.

- Case-1:  $f(u) = \frac{u^2}{2}, u(x, 0) = \sin(\pi x) + 0.5. \quad x \in [0, 2]$
- Case-2:  $f(u) = \frac{u^2}{2}, u(x, 0) = \begin{cases} u_l & \text{if } x < 0 \\ u_r & \text{if } x \geq 0. \end{cases}$

Solutions to Case-1 consist of smooth extrema, smooth large gradient and discontinuities, we can use Newton Iteration method to solve “u” in nonlinear implicit equation  $u = u_0(x - ut)$  mentioned in Sect. 2. Solutions to Case-2 include a centered rarefaction wave or a shock wave with different selections of  $u_l, u_r$ .

#### 3.1.2 Input Stencil for Training Data

Since the shock wave detector in this work is to detect discontinuities from perspective of characteristic curve rather than from the features of spurious oscillations, we only select the cell-average of the solution  $\bar{u}_e$  and cell-size  $h$  as the input for training. For the Burgers equation (where  $f(u) = \frac{u^2}{2}$ ), the input stencil is  $[\bar{u}_{e-2}, \bar{u}_{e-1}, \bar{u}_e, \bar{u}_{e+1}, \bar{u}_{e+2}, h]$ , where  $\bar{u}_e = \frac{1}{\Delta x} \int_{I_e} u(x, t_n) dx$  is the cell-average of the exact solution.

#### 3.1.3 Labels for Each Data

Given a cell  $I_e$  and its input information  $[\bar{u}_{e-2}, \bar{u}_{e-1}, \bar{u}_e, \bar{u}_{e+1}, \bar{u}_{e+2}, h]$ , we label  $I_e$  as a troubled-cell if a discontinuity of exact solution exists inside  $I_e$ , the label is set to 1. Otherwise, the label is set to 0 to represent  $I_e$  is a good cell.

Based on the process mentioned in Sect. 3.1, the algorithm to generate the training and validation set for the Burgers equation is given as follows.

**Algorithm 3.1** generate the training and validation data set

---

```

Define the training and validation set:  $\mathcal{R}$ 
Select Case-I in Sect. 3.1.1,  $I=1,2$ 
while the amount of data obtained from Case-I is larger than  $N_I$  do
  Step-1: Set  $u_l, u_r$  (Case - 2),  $t$ .
  Step-2: Calculate the exact solution  $u(x, t)$  under the selection of  $u_l, u_r$ , and  $t$ .
  Step-3: Set the cell-size  $h$  and discrete space domain  $[a, b] = \bigcup_{e=1}^N I_e$ , and calculate the cell-average of  $u(x,t)$  in each  $I_e$ .
  Step-4: Set the input stencil of each data and label each data by Sects. 3.1.2 and 3.1.3.
  Step-5: Add the same amount of data labeled "0" and "1" into the data set  $\mathcal{R}$ .
end while
Change the Case-I and repeat steps above until I traversing 1,2
return  $\mathcal{R}$ 

```

---

**Notice**

- (1) The reason we choose 5-cell-stencil is a balance for accuracy and compactness. 3-cell-stencil might not be enough to capture the trend of the solution in cell-averages, and multi-cell-stencil might increase the computational cost greatly and lose the compactness of the scheme as well. To do so, we can largely maintain the local compactness if applied to DG scheme.
- (2) Although the obtained detector is applied under DG scheme in this work, it can be applied to high-order finite difference or finite volume schemes.
- (3) Contact discontinuities are not included in data set. Theoretically, contact discontinuities might not be detected by the resulting model for the scalar conservation law.
- (4) The amount of data labeled "0" should be almost equal to that labeled "1" in order to ensure unbiased data set.

**3.2 Parameters in AN Training**

In this part, we seek a suitable AN for constructing a shock wave detector and suitable conditions for AN training. Only one linear hidden layer with 6 neurons lies in the AN, and is connected with the input layer linearly, the sigmoid activation function is used in the output layer. The structure is given as follows (Fig. 2):

The weight vectors in AN to be determined include weights  $\mathbf{W}$  and biases  $\mathbf{B}$ . Here,  $\mathbf{W} = \{\mathbf{W}_1, \mathbf{W}_2\}$  with  $\mathbf{W}_1 \in \mathbb{R}^{6 \times 6}$ ,  $\mathbf{W}_2 \in \mathbb{R}^{6 \times 1}$  and  $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$  with  $\mathbf{b}_1 \in \mathbb{R}^6$ ,  $\mathbf{b}_2 \in \mathbb{R}^1$ .

Following the Algorithm 3.1 in Sect. 3.1, we generate the training and validation data set with 1600 amounts of supervised data including 400 from Case-1, 1200 from Case-2.

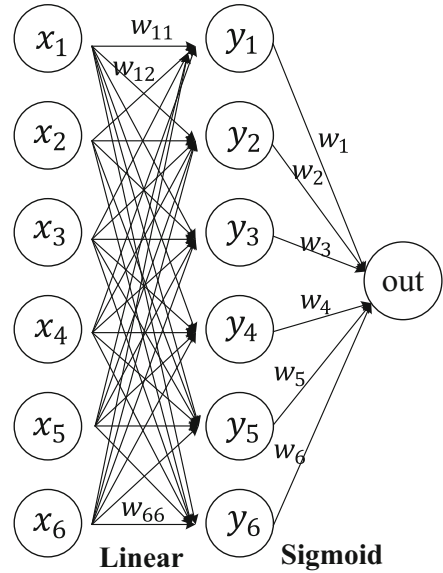
Next, the AN is trained under following conditions. The weight parameters in AN are initialized based on a normal distribution. We use stochastic gradient descent (SGD) method for optimizing with a mini-batch size of 16. The learning algorithm converges fast and it takes only few minutes.

**3.3 Training Weights Results**

The convergent AN result is given as follows,

$$out_e = \frac{1}{1 + e^{-(3.86\bar{u}_{e-2} + 13.44\bar{u}_{e-1} + 0.17\bar{u}_e - 11.88\bar{u}_{e+1} - 5.68\bar{u}_{e+2} - 9.60h - 4.22)}}. \tag{3.1}$$

Fig. 2 The AN structure



For the Burgers equation, the primitive variable  $u$  is precisely the characteristic variable  $\lambda(u)$ . Thus, using the exact solution as input is equivalent to using the exact eigenvalue (slope of characteristic curve) as input for the Burgers equation. In order to illustrate the mechanism of the expression (3.1) from the perspective of characteristic curves, we use the characteristic variable  $\lambda(u)$  for subsequent analysis in Sect. 4.

In the next section, we will show that expression (3.1) implies the information of shock wave discontinuities.

### 4 The Shock Wave Indicator and Its Working Mechanism

In this section, we shall analyze the working mechanism of expression (3.1) obtained in Sect. 3 so that it can be used generally to numerical simulation of conservation laws. Before analysis, we firstly give two definitions.

**Definition 4.1** For cell-average of  $u$  inside  $I_e$  and its neighbor cells  $\bar{u}_{e-1}, \bar{u}_e, \bar{u}_{e+1}$ , we call  $\bar{u}_L = \sum_{i=0}^1 w_L^i \bar{u}_{e-i}$  is a type of left-side weighted average of variable  $u$  in  $I_e$  reconstructed by  $\bar{u}_{e-1}, \bar{u}_e$  if  $w_L^i > 0$  and  $\sum_{i=0}^1 w_L^i = 1$ ; we call  $\bar{u}_R = \sum_{i=0}^1 w_R^i \bar{u}_{e+i}$  is a type of right-side weighted average of variable  $u$  in  $I_e$  reconstructed by  $\bar{u}_e, \bar{u}_{e+1}$  if  $w_R^i > 0$  and  $\sum_{i=0}^1 w_R^i = 1$ .  $w_L^i, w_R^i$  are called weight coefficient.

In Sect. 4.2, we will show some properties of  $\bar{u}_L$  and  $\bar{u}_R$ .

**Definition 4.2** If a certain eigenvalue  $\lambda(u)$  (slope of characteristic curve) has a jump somewhere, and this jump leads to characteristic curve compressing or intersecting, we call this jump of  $\lambda(u)$  is **admissible**. Otherwise, this jump is called **inadmissible**.

Definition 4.2 is well-defined in the physical sense. According to the Lemma 2.1 and Lemma 2.3(2), the compressing or intersecting of characteristic curves leads to the solution



evolving into a discontinuity, this discontinuity is admissible in weak solution, while the divergence of characteristic curves leads to the solution evolving into a rarefaction wave.

*Remark* The inadmissible jumps include linear contact discontinuities, across which the characteristic curves are parallel, and those initial jumps, which evolve into rarefaction waves.

In this section, we will construct a shock wave indicator in numerical simulation of conservation laws based on the AN model (3.1), and analyze its working mechanism based on the assumption “step-size  $h$  is sufficiently small” and Definitions 4.1, 4.2. This section consists of three subsections. In Sect. 4.1 we modify expression (3.1) into a more practical form for shock wave detecting. In Sect. 4.2, we apply the detector to numerical schemes as a shock wave indicator, and analyze its working mechanism. The indicator is extended to system conservation laws in Sect. 4.3 by directly applying it to each genuinely nonlinear characteristic field.

### 4.1 The Shock Wave Indicator

In order to make the output of artificial neuron model compact and explicable, we are going to modify the initial indicator (3.1) into a practical form on the assumption of small mesh cell and ignoring certain small terms. The results are presented in the following two lemmas.

**Lemma 4.3** *Under the assumption that the discontinuity only locates on  $I_e$ , if the space step-size  $h$  is sufficiently small, the form (3.1) can be approximately modified as the form of*

$$out_e \approx \frac{1}{1 + e^{-(W_L \bar{\lambda}_L - W_R \bar{\lambda}_R) + M_1 h + M_2}} := out'_e.$$

Here,  $\bar{\lambda}_L = w_L^0 \lambda(\bar{u}_e) + w_L^1 \lambda(\bar{u}_{e-1})$  and  $\bar{\lambda}_R = w_R^0 \lambda(\bar{u}_e) + w_R^1 \lambda(\bar{u}_{e+1})$  are the left and right side weighted averages of eigenvalue  $\lambda(u)$  respectively, and  $W, M_1, M_2$  are constant.

**Proof** Based on the assumption and Taylor expansion, we have  $|\bar{u}_{e-2} - \bar{u}_{e-1}| = O(h)$ ,  $|\bar{u}_{e+2} - \bar{u}_{e+1}| = O(h)$ , thus we have

$$\begin{aligned} \frac{3.86\lambda(\bar{u}_{e-2}) + 13.44\lambda(\bar{u}_{e-1}) + 0.17\lambda(\bar{u}_e)}{W_L} &\approx \frac{17.30\lambda(\bar{u}_{e-1}) + 0.17\lambda(\bar{u}_e)}{W_L} =: \bar{\lambda}_L, \\ \frac{0\lambda(\bar{u}_e) + 11.88\lambda(\bar{u}_{e+1}) + 5.68\lambda(\bar{u}_{e+2})}{W_R} &\approx \frac{0\lambda(\bar{u}_e) + 17.56\lambda(\bar{u}_{e+1})}{W_R} =: \bar{\lambda}_R. \end{aligned} \tag{4.1}$$

As a result, the expression (3.1) can be approximately rewritten as

$$out_e \stackrel{(4.1)}{\approx} \frac{1}{1 + e^{-(W_L \bar{\lambda}_L - W_R \bar{\lambda}_R) + M_1 h + M_2}} =: out'_e, \tag{4.2}$$

where  $W_L = 3.86 + 13.44 + 0.17 = 17.47$ ,  $W_R = 0 + 11.88 + 5.68 = 17.56$ ,  $M_1 = 9.60$ ,  $M_2 = 4.22$ , and  $\bar{\lambda}_L, \bar{\lambda}_R$  are left and right side weighted averages of eigenvalue  $\lambda(u)$  defined in (4.1) presented as

$$\begin{aligned} \bar{\lambda}_L &= w_L^0 \lambda(\bar{u}_e) + w_L^1 \lambda(\bar{u}_{e-1}) = 0.01\lambda(\bar{u}_e) + 0.99\lambda(\bar{u}_{e-1}) \\ \bar{\lambda}_R &= w_R^0 \lambda(\bar{u}_e) + w_R^1 \lambda(\bar{u}_{e+1}) = 0.0\lambda(\bar{u}_e) + 1.0\lambda(\bar{u}_{e+1}). \end{aligned} \tag{4.3}$$

□

One may note that  $W_L \approx W_R$  in (4.2), we find the expression above  $out'_e$  can be further modified into a more practical form, which is convenient for mechanism analyzing. Specifically, replacing  $W_L, W_R$  by a constant number of  $W = 17.5$  and ignoring the impacts of small terms  $\Delta W_L := W_L - W, \Delta W_R := W_R - W$ , we have the following lemma.

**Lemma 4.4** *If ignoring the relatively small difference of weight coefficients  $O(\Delta W_L)$  and  $O(\Delta W_R)$ , the form (4.2) can be approximately modified into the following practical form of*

$$out'_e \approx \frac{1}{1 + e^{-W(\bar{\lambda}_L - \bar{\lambda}_R) + M_1 h + M_2}} := \widetilde{out}_e.$$

Here,  $\bar{\lambda}_L$  and  $\bar{\lambda}_R$  are the left and right side weighted averages of eigenvalue  $\lambda(u)$  defined in (4.3), and  $W = 17.5$ ,  $M_1 = 9.60$ ,  $M_2 = 4.22$  remain constant.

**Proof** Based on the assumption, we have

$$\begin{aligned} W_L \bar{\lambda}_L - W_R \bar{\lambda}_R &= W(\bar{\lambda}_L - \bar{\lambda}_R) + \Delta W_L \bar{\lambda}_L + \Delta W_R \bar{\lambda}_R \\ &= W(\bar{\lambda}_L - \bar{\lambda}_R) + O(\Delta W_L) + O(\Delta W_R) + O(h), \end{aligned} \tag{4.4}$$

with  $W = 17.5$ ,  $\Delta W_L = -0.03$  and  $\Delta W_R = 0.06$ .

Although  $\Delta W_L$ ,  $\Delta W_R$  does not scale with  $h$ , we regard these as respectively small terms due to  $\frac{\Delta W_L}{W}, \frac{\Delta W_R}{W} \sim 1e - 3 \ll 1$  and in addition drop the term  $O(\Delta W_L) + O(\Delta W_R)$  in (4.4) so that the expression (4.4) is convenient to be explicable.

As a result, the expression (3.1) can be further approximately modified as

$$out_e \stackrel{(4.1)}{\approx} out'_e \stackrel{(4.4)}{\approx} \frac{1}{1 + e^{-W(\bar{\lambda}_L - \bar{\lambda}_R) + M_1 h + M_2}} =: \widetilde{out}_e. \tag{4.5}$$

□

Thus, according to the Lemmas 4.3 and 4.4, the indicator (3.1) can be approximately rewritten as the concise and universal form of

$$\widetilde{out}_e = \frac{1}{1 + e^{-W(\bar{\lambda}_L - \bar{\lambda}_R) + M_1 h + M_2}}. \tag{4.6}$$

Note that the approximation made in (4.1) is to recover the “local” property of DG scheme, the approximation made in (4.4) is for the convenience of using characteristic compressing to analyze the working mechanism of (4.6). With the above modifications, we are able to discover the properties of the practical form (4.6) when applied to be a shock wave indicator in the next subsection.

### 4.2 Working Mechanism of the Shock Wave Indicator

In this subsection we will prove the practical form (4.6) can be used to detect shock wave discontinuities in numerical simulation of conservation laws.

We give the following assumption for discontinuity indicating in numerical simulation.

**Assumption 4.1** Assuming  $h$  is sufficiently small, we consider the  $O(1)$  jump between  $\bar{u}_L$  and  $\bar{u}_R$  is a discontinuity in numerical simulation of variable “ $u$ ”, where  $\bar{u}_L$  and  $\bar{u}_R$  are side weighted averages of  $u$  defined in Definition 4.1.

All analysis below is done under the Assumption 4.1.

By denoting the local region  $I_E = \bigcup_{i=e-1}^{e+1} I_i$ , we obtain the following lemma to explain the role  $\bar{\lambda}_L$  and  $\bar{\lambda}_R$  playing in numerical simulation of conservation laws.

**Lemma 4.5** *If the solution in region  $I_E$  is differentiable,  $|\bar{\lambda}_L - \bar{\lambda}_R| \leq M_e h \sim O(h)$ . If there exists a discontinuity in  $I_e$ ,  $\bar{\lambda}_L$  and  $\bar{\lambda}_R$  satisfy  $\bar{\lambda}_R - \bar{\lambda}_L = C[\lambda(u)] + O(h) \sim O(1)$ , where  $C < 1$  and depends on the weight coefficients  $w_L^0, w_R^0$ . Specially, if  $w_L^0, w_R^0 \rightarrow 0$ ,  $\bar{\lambda}_R - \bar{\lambda}_L \rightarrow [\lambda(u)] + O(h)$ , where  $[\lambda(u)]$  represents jump of a certain eigenvalue.*

**Proof** (1) If solution  $u$  is differentiable in region  $I_E$ , denoting  $\bar{\lambda}_e = \lambda(\bar{u}_e)$ , by Taylor expansion we can conclude  $\bar{\lambda}_L = \bar{\lambda}_{e-1} + w_L^0(\bar{\lambda}_e - \bar{\lambda}_{e-1}) = \bar{\lambda}_{e-1} + w_L^0\lambda'(\xi_1)O(h) = \bar{\lambda}_{e-1} + O(h)$ , similarly  $\bar{\lambda}_R = \bar{\lambda}_{e+1} + O(h)$ . Then  $\bar{\lambda}_L$  and  $\bar{\lambda}_R$  satisfy  $|\bar{\lambda}_L - \bar{\lambda}_R| \leq M_e h \sim O(h)$ .

(2) If the solution consists of a discontinuity inside  $I_e$ ,  $\bar{\lambda}_R - \bar{\lambda}_L = w_R^0\bar{\lambda}_e + (1 - w_R^0)\bar{\lambda}_{e+1} - w_L^0\bar{\lambda}_e - (1 - w_L^0)\bar{\lambda}_{e-1} = (1 - w_L^0 - w_R^0)(\bar{\lambda}_{e+1} - \bar{\lambda}_{e-1}) + w_L^0(\bar{\lambda}_{e+1} - \bar{\lambda}_e) + w_R^0(\bar{\lambda}_e - \bar{\lambda}_{e-1}) + O(h) \triangleq w_1(\bar{\lambda}_{e+1} - \bar{\lambda}_{e-1}) + w_2(\bar{\lambda}_{e+1} - \bar{\lambda}_e) + w_3(\bar{\lambda}_e - \bar{\lambda}_{e-1})$  with  $w_1 + w_2 + w_3 = 1$ .

Note that on the left side of the discontinuity located in  $I_e$ , we have  $\lambda(u(x)) = \bar{\lambda}_{e-1} + O(h)$ . Similarly, on the right side of the discontinuity we have  $\lambda(u(x)) = \bar{\lambda}_{e+1} + O(h)$ . Thus, we have

$$\bar{\lambda}_{e+1} - \bar{\lambda}_{e-1} = [\lambda(u)] + O(h). \tag{4.7}$$

Futhermore, we can always find a  $0 \leq \mu \leq 1$  which only depends on the cell-size  $h$  and the location of the discontinuity in  $I_e$ , such that

$$\bar{\lambda}_e = \mu\bar{\lambda}_{e-1} + (1 - \mu)\bar{\lambda}_{e+1} + O(h). \tag{4.8}$$

Using (4.7) and (4.8), we have

$$\bar{\lambda}_{e+1} - \bar{\lambda}_e = \mu[\lambda(u(x))] + O(h), \quad \bar{\lambda}_e - \bar{\lambda}_{e-1} = (1 - \mu)[\lambda(u(x))] + O(h).$$

Therefore,  $\bar{\lambda}_R - \bar{\lambda}_L = C[\lambda(u)] + O(h) \sim O(1)$ , where  $C = w_1 + \mu w_2 + (1 - \mu)w_3 < 1$ . Specially, if  $w_L^0, w_R^0 \rightarrow 0, C \rightarrow 1$ , we can obtain  $\bar{\lambda}_R - \bar{\lambda}_L \rightarrow [\lambda(u)] + O(h)$ .  $\square$

The following lemma show a relationship between  $\widetilde{out}_e$  and  $\bar{\lambda}_L - \bar{\lambda}_R$ .

**Lemma 4.6** *In  $I_e$ ,*

- (1)  $\widetilde{out}_e < 0.0145 \ll 0.5 \iff \bar{\lambda}_L - \bar{\lambda}_R < 0$  or  $|\bar{\lambda}_L - \bar{\lambda}_R| \sim \frac{1.0e-3}{W} + O(h)$ .
- (2)  $\widetilde{out}_e \geq 0.5 \iff \bar{\lambda}_L - \bar{\lambda}_R \geq 0$  and  $\bar{\lambda}_L - \bar{\lambda}_R \geq \frac{M_1h+M_2}{W} \sim O(1)$ , and  $\widetilde{out}_e \rightarrow 1 \iff \bar{\lambda}_L - \bar{\lambda}_R \rightarrow +\infty$ .

**Proof** According to (4.6), we can derive,

$$\widetilde{out}_e \geq s \iff e^{-W(\bar{\lambda}_L - \bar{\lambda}_R) + M_1h + M_2} \leq \frac{1}{s} - 1 \iff \bar{\lambda}_L - \bar{\lambda}_R \geq C_0. \tag{4.9}$$

Here,  $s$  is a number belonging to  $(0, 1]$ , and  $C_0 = \frac{1}{W}(M_1h + M_2 - \ln(\frac{1}{s} - 1))$ .

(1) If  $\bar{\lambda}_L - \bar{\lambda}_R < 0$ ,  $\widetilde{out}_e < \frac{1}{1 + e^{M_1h + M_2}} \approx 0.0145 \ll 0.5$ .

If  $\bar{\lambda}_L - \bar{\lambda}_R > 0$  and  $\bar{\lambda}_L - \bar{\lambda}_R \sim \frac{1.0e-3}{W} + O(h)$ ,  $\widetilde{out}_e \approx 0.0145 \ll 0.5$ .

Conversely, if  $\widetilde{out}_e < 0.0145$ ,  $\bar{\lambda}_L, \bar{\lambda}_R$  should satisfy  $\bar{\lambda}_L - \bar{\lambda}_R < 0$  or  $|\bar{\lambda}_L - \bar{\lambda}_R| \sim \frac{1.0e-3}{W} + O(h)$ .

(2) Denoting  $s = 0.5$  in (4.9), we can obtain  $C_0 = \frac{M_2}{W} + O(h) \sim O(1)$ . Therefore,  $\bar{\lambda}_L - \bar{\lambda}_R \geq \frac{M_1h+M_2}{W} \sim O(1) \iff \widetilde{out}_e \geq 0.5$ .

As  $\widetilde{out}_e$  increases from 0.5 to 1,  $\bar{\lambda}_L - \bar{\lambda}_R$  also increases from  $\frac{M_1h+M_2}{W}$  to  $+\infty$ . Therefore,  $\widetilde{out}_e \rightarrow 1 \iff \bar{\lambda}_L - \bar{\lambda}_R \rightarrow +\infty$ .  $\square$

**Remark** In numerical simulation,  $h$  is generally much larger than  $\frac{1.0e-3}{W} \sim 10^{-5}$ , therefore we can assume  $|\bar{\lambda}_L - \bar{\lambda}_R| \sim \frac{1.0e-3}{W} + O(h)$  in Lemma 4.6 is equivalent to  $|\bar{\lambda}_L - \bar{\lambda}_R| \sim O(h)$  in subsequent analysis.

**Theorem 4.7** *The output of (4.6) in  $I_e$   $\widetilde{out}_e$  satisfies  $\widetilde{out}_e < 0.0145 \ll 0.5$  if  $I_e$  is a cell near a smooth extrema or an inadmissible large-gradient/jump.*

**Proof** (1) For a smooth extrema. If  $I_e$  is a cell near a smooth extrema, under the smooth assumption, the derivative of  $u$  at the extreme point is 0, then we can obtain  $|\bar{\lambda}_L - \bar{\lambda}_R| < d_0|\lambda'(\xi)u'(\eta)|h < D_0h^2$  by Taylor expansion. Immediately by Lemma 4.6,  $\widetilde{out}_e = \frac{1}{1+e^{M_2+O(h)}} \approx 0.0145 \ll 0.5$ .

(2) For a inadmissible large gradient/jump. If  $I_e$  is a cell near an inadmissible large-gradient/jump,  $\bar{\lambda}_L - \bar{\lambda}_R < 0$  by Definition 4.2, immediately we can obtain directly  $\widetilde{out}_e \ll 0.5$  by Lemma 4.6. □

Therefore, the indicator (4.6) can never indicate cells near a smooth extrema or an inadmissible large-gradient/jump as troubled-cells.

In theory, a solution discontinuity occurs only inside one cell. In numerical simulation, due to numerical dissipation, a discontinuity might span a few cells of the numerical solution. Therefore, these cells will be detected as the troubled-cells by the indicator (4.6), due to their  $O(1)$  magnitude admissible jump value of  $\lambda(u)$ .

**Theorem 4.8** *In the numerical simulation of system (2.1), a discontinuity of numerical solution caused by compressing of characteristic curves exists indeed inside  $I_e$  if  $\widetilde{out}_e \geq 0.5$ .*

**Proof** As one can obtain in Lemma 4.6,  $\widetilde{out}_e \geq 0.5 \iff \bar{\lambda}_L - \bar{\lambda}_R \geq \frac{M_1h+M_2}{W} \sim O(1)$ . By the Definition 4.2 there exists an admissible jump value of  $\lambda(u)$  and the admissible jump value has a magnitude of  $O(1)$ . Under the Assumption 4.1 there exists an occurrence of characteristic curves intersection inside  $I_e$ , consequently a discontinuity of numerical solution exists indeed inside  $I_e$ .

Therefore,  $\widetilde{out}_e \geq 0.5 \iff \bar{\lambda}_L - \bar{\lambda}_R \geq \frac{M_1h+M_2}{W} \sim O(1) \implies$  a discontinuity of numerical solution occurs inside  $I_e$ . □

*Remark*

- (1) Since  $I_e$  is detected as a troubled-cell if  $\widetilde{out}_e \geq 0.5$ , this correspondingly results in  $\bar{\lambda}_L - \bar{\lambda}_R \geq 0.241 + 0.549h$ . Therefore, if the jump strength of the discontinuity is less than 0.241, the present indicator would not treat it.
- (2) The scalar indicator (4.6) would not detect the initial inadmissible jumps (such as the initial linear contact discontinuities and those evolving into rarefaction waves). As a remedy, we suggest a simple strategy to treat the initial inadmissible jumps as for the 1-2-3 case to be given in the section of numerical simulation.

Next, we will extend the indicator (4.6) to one-dimensional system conservation laws.

**4.3 Extension to System Conservation Laws**

For system (2.3), according to Lemma 2.3(2), the solution can evolve into shock waves due to the intersection of a certain  $i$ -characteristic curves in genuinely nonlinear characteristic field, therefore, we only consider the eigenvalues in genuinely nonlinear characteristic field. By the Entropy Condition which is shown in [35], the intersection of a certain  $i$ -characteristic curves implies  $\lambda_i^- > \lambda_i^+$ , where  $\lambda_i^\pm = \lim_{\epsilon \rightarrow 0^+} \lambda_i(u(x \pm \epsilon))$ .

For system conservation laws, we do not need to retrain the shock wave indicator, and can directly apply the indicator (4.6) to each genuinely nonlinear  $\lambda_i$ -characteristic field. Following the idea, we can extend indicator (4.6) to system conservation laws by two steps, the first one is to calculate the outputs of indicator (4.6) for each eigenvalue  $\lambda_i$ ; the other is to select the maximum outputs among them.

The expression of indicator for the system is then given as follows,

$$\widehat{out}_e = \max_i \{ \widetilde{out}_e(\bar{\Lambda}_i, h) \}. \tag{4.10}$$

Here,  $\bar{\Lambda}_i = (\bar{\lambda}_{i,e-1}, \bar{\lambda}_{i,e}, \bar{\lambda}_{i,e+1})$ ,  $\widetilde{out}_e(\bar{\Lambda}_i, h)$  is defined by the expression (4.6).

**Lemma 4.9** *For the 1-D Euler equations, the indicator (4.10) becomes*

$$\widehat{out}_e = \max \{ \widetilde{out}_e(\bar{u} - \bar{a}, h); \widetilde{out}_e(\bar{u} + \bar{a}, h) \}. \tag{4.11}$$

Here,  $\bar{u} \pm \bar{a} = (\bar{u}_{e-1} \pm \bar{a}_{e-1}, \bar{u}_e \pm \bar{a}_e, \bar{u}_{e+1} \pm \bar{a}_{e+1})$ .

For system (2.3), indicator (4.10) possesses similar indicating properties as indicator (4.6). Therefore, we give the following two conclusions without proof.

**Theorem 4.10** *The output of (4.10) in  $I_e$  satisfies  $\widehat{out}_e \ll 0.5$  if  $I_e$  is a cell near a smooth extrema or an inadmissible large-gradient/jump.*

**Theorem 4.11** *A discontinuity with admissible eigenvalue jump value exists indeed inside the  $I_e$  if  $\widehat{out}_e \geq 0.5$ .*

**Corollary 4.12** *For the Euler equations, a shock wave discontinuity or a contact discontinuity exists indeed inside the  $I_e$  if  $\widehat{out}_e \geq 0.5$ .*

**Proof** According to the Lemma 2.4, an admissible jump occurs to one of the non-linear eigenvalues if there is a shock wave. In addition, across the 2-contact discontinuity,  $u, p$  remain unchanged while  $\rho$  has a jump, leading to a jump occurrence for the sound speed  $a$  and the satisfaction of either  $u_l - a_l > u_r - a_r$  or  $u_l + a_l > u_r + a_r$  across the contact discontinuity. As a result, additionally there is always an admissible jump occurring to the 1- or 3- eigenvalue when across the 2-contact discontinuity. According to Theorem 4.8, if  $\widehat{out}_e \geq 0.5$ , a shock wave discontinuity or a contact discontinuity exists inside the  $I_e$ .  $\square$

*Remark* The Lemma 4.9 and Corollary 4.12 imply that besides the shock wave discontinuities, the present indicator (4.11) is able to detect the 2-contact discontinuities, although the 2-characteristic is never be traced. The present indicator maintains the mathematical property shown in Lemma 2.4.

### 5 Runge–Kutta Discontinuous Galerkin (RKDG) Framework

In this section we take system (2.1) as an example to review the RKDG framework, and the indicator (4.6) will be applied to RKDG framework compared to TVB-based indicator.

To approximate the solution of system (2.1), the computational domain  $[a, b]$  is discretized into  $N$  non-overlapping cells,  $[a, b] = \bigcup_{e=1}^N I_e$ ,  $I_e = [x_{e-\frac{1}{2}}, x_{e+\frac{1}{2}}]$ , where  $a = x_{\frac{1}{2}} < x_{\frac{3}{2}} < \dots < x_{N+\frac{1}{2}} = b$ ,  $x_e$  is the center of  $I_e$ ,  $h$  is the cell-size of each cell. The approximate solution  $u_h$  is an element in the space of broken polynomials  $V_h^k = \{u \in L^2[a, b] : u|_{I_e} \in P^k(I_e)\}$ , where  $P^k(I_e)$  is the space of polynomials with degree  $d \leq k$  in the cell  $I_e$ . In each  $I_e$ , the approximate solution  $u_h$  can be represented by  $u_h = \sum_{i=0}^k u_e^{(i)}(t) \cdot \phi_{e,i}(x)$ , where  $\{\phi_{e,i}(x)\}_i$  is a basis of  $P^k(I_e)$ . The semi-discrete DG scheme can be formulated as follow,

$$\left( \int_e \phi_{e,i} \phi_{e,i} dx \right) \cdot \frac{d}{dt} u_e^{(i)} = \int_e f(u_h) \frac{d\phi_{e,i}}{dx} dx - \left[ \widehat{f}_{e+\frac{1}{2}}(u) \cdot \phi_{e,i} \left( x_{e+\frac{1}{2}}^- \right) - \widehat{f}_{e-\frac{1}{2}}(u) \cdot \phi_{e,i} \left( x_{e-\frac{1}{2}}^+ \right) \right] \tag{5.1}$$

Here,  $\widehat{f}_{e+\frac{1}{2}}(u) = \frac{1}{2}[f(u_{e+\frac{1}{2}}^-) + f(u_{e+\frac{1}{2}}^+) - \alpha(u_{e+\frac{1}{2}}^+ - u_{e+\frac{1}{2}}^-)]$  is the Local-Lax Friedrichs numerical flux and  $u(x^\pm) = \lim_{\epsilon \rightarrow 0^+} u(x \pm \epsilon)$ . We use SSP-RK3 [10] scheme to discretize time derivative.

To suppress oscillations across the discontinuities, we use TVB limiters to correct the high order dofs of polynomial inside  $I_e$ , TVB-based indicators uses the 3-cell stencil to detect the troubled-cells. The process for TVB limiting technique and the present indicator based limiting technique are given as Algorithms 5.1 and 5.2,

---

**Algorithm 5.1** TVB limiting process

---

- Step-1: Calculate dofs  $u_{e,n}^{(i)}$  and cell-average  $\bar{u}_e$  for each e and i when  $t = t^n$
  - Step-2: Calculate left and right cell-interface values  $u_{e+\frac{1}{2}}^-$  and  $u_{e-\frac{1}{2}}^+$  for each e
  - Step-3: Correct high order dofs in  $I_e$  if  $|\bar{u}_e| > Mh^2$  or  $|\widetilde{u}_e| > Mh^2$ .
  - Step-4: Obtain modified high order dofs  $u_{e,n}^{(i) (mod)}$ .
- 

---

**Algorithm 5.2** present indicator process

---

- Step-1: Calculate dofs  $u_{e,n}^{(i)}$  and cell-average  $\bar{u}_e$  for each e and i when  $t = t^n$
  - Step-2: Calculate output of the present indicator on each cell  $I_e$ , (4.6)  $\widetilde{out}_e$  for scalar and (4.10)  $\widehat{out}_e$  for system conservation laws.
  - Step-3: Flag  $I_e$  if  $\widetilde{out}_e \geq 0.5$  or  $\widehat{out}_e \geq 0.5$ .
  - Step-4: Modify high order dofs in flagged cells, obtain the modified high order dofs  $u_{e,n}^{(i) (mod)}$ .
- 

Next section, we will test numerical cases under the framework of RKDG to compare the behaviors of the present indicators and the TVB-based indicators.

## 6 Numerical Results

We now demonstrate the capability of the present indicators (4.6) and (4.10) when applied to RKDG framework. We compare the performance of the indicators (4.6) and (4.10) with that of the TVB-based indicators. The notation TVB-1, TVB-2, TVB-3 are used to refer to the TVB limiter with the parameter  $M = 10$ ,  $M = 100$ , and  $M = 1000$ , respectively.

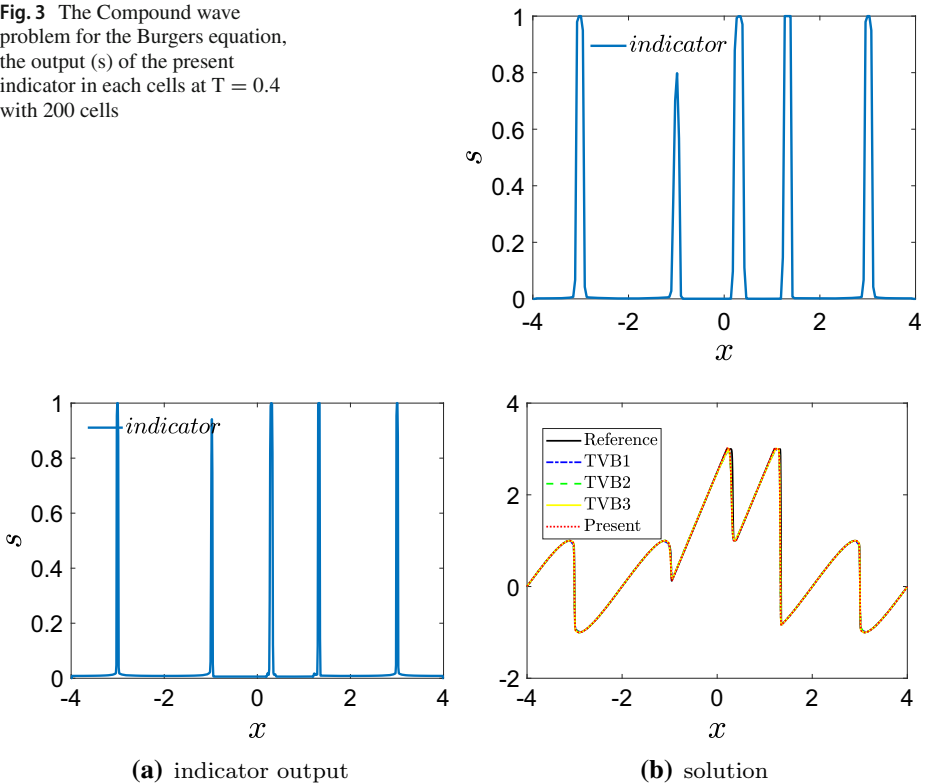
### 6.1 The 1-D Burgers Equation

#### 6.1.1 Compound Wave Case

The initial condition for this test is a composition of smooth and discontinuous data [28], expressed as

$$u_0(x) = \begin{cases} \sin(\pi x) & \text{if } |x| \geq 1 \\ 3 & \text{if } -1 < x \leq -0.5 \\ 1 & \text{if } -0.5 < x \leq 0 \\ 3 & \text{if } 0 < x \leq 0.5 \\ 2 & \text{if } 0.5 < x \leq 1, \end{cases}$$

**Fig. 3** The Compound wave problem for the Burgers equation, the output (s) of the present indicator in each cells at  $T = 0.4$  with 200 cells



**Fig. 4** Solutions of the Compound wave problem for the Burgers equation with various indicators at  $T = 0.4$  under RKDG(p4) framework with 800 cells

on the domain  $[-4, 4]$  with periodic boundary conditions, the solution is simulated until time  $T = 0.4$ . As the solution evolves, alternating shock and rarefaction waves begin to develop. We test this testcase under  $N = 200$  and  $N = 800$  cells to assess the behaviour of the present indicator under different cell-size.

Outputs (s) of the present indicator (4.6) at  $T = 0.4$  are shown in Figs. 3 and 4a. The results with various indicators on a mesh with  $N = 200$  cells and under DG(p4) framework are shown in Fig. 5a, b. The solution with TVB-3 indicator suffers from oscillations near the shock waves because some of the shock waves are not detected. The solution with the present indicator is similar to that with TVB-2 indicator, and is superior to that with TVB-1 indicator. We also show the time-history of the troubled-cell indication with 200 cells as shown in Fig. 6a–d, except that there is no indication for initial inadmissible discontinuities, the present indicator has a similar performance to the TVB-1 indicator. With  $N = 800$  cells under DG(p3) framework, as shown in Fig. 7, the TVB-3 indicator performs best among the TVB indicators, TVB-1 and TVB-2 indicators work worse than that with  $N = 200$  cells, this implies that the performance of TVB-based indicators is much related to the cell size. On the other hand, the present indicator works consistently.

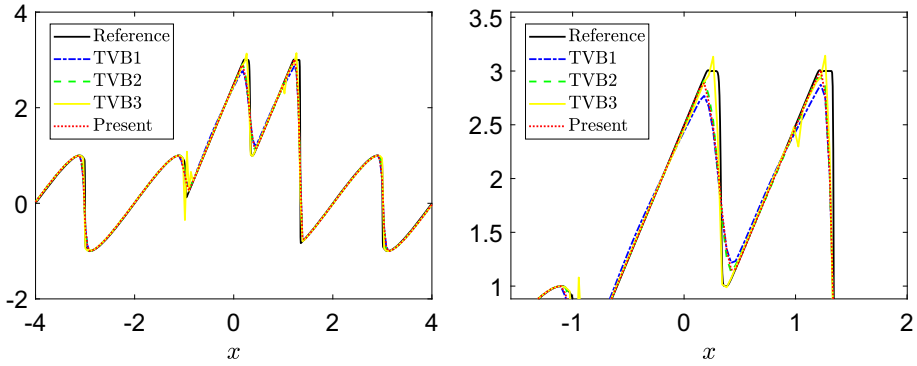


Fig. 5 Solutions of the Compound wave problem for the Burgers equation with various indicators at  $T = 0.4$  under RKDG(p4) framework with 200 cells

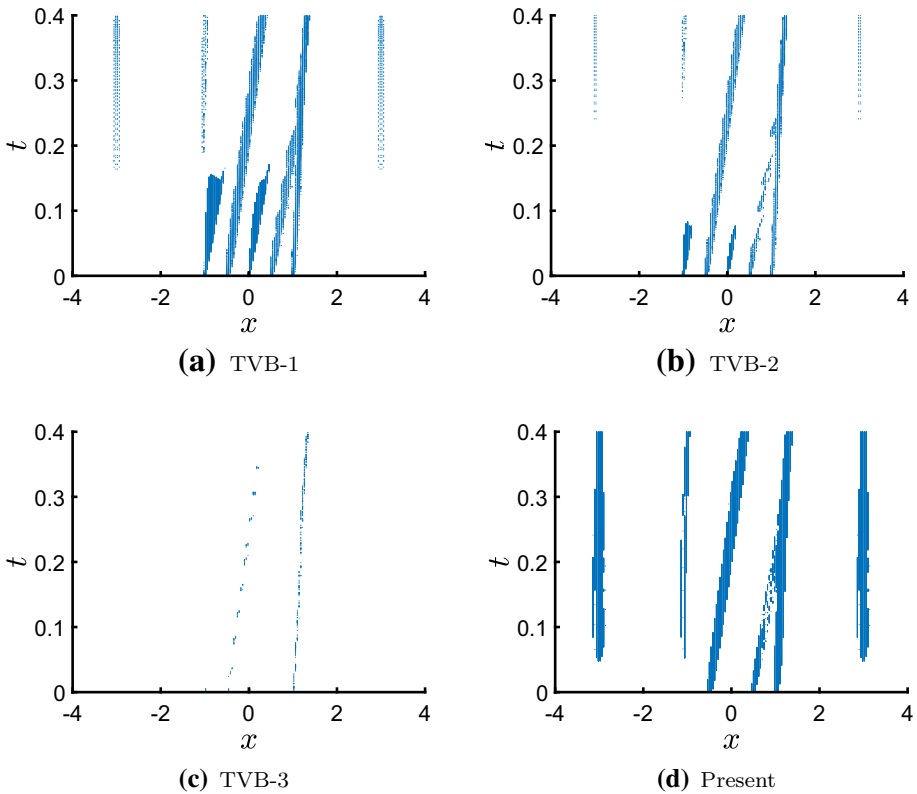
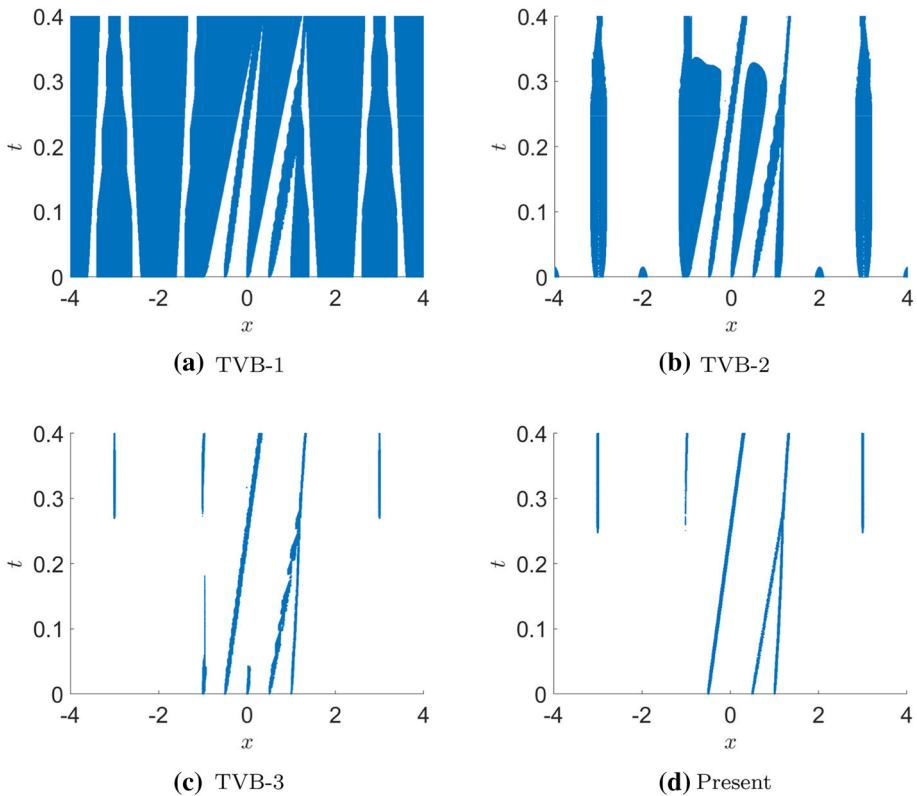


Fig. 6 The time-history of flagged troubled-cells of the compound wave problem for the Burgers equation with 200 cells





**Fig. 7** The time-history of flagged troubled-cells of the compound wave problem for the Burgers equation with 800 cells

### 6.2 Buckley–Leverett Case

In order to demonstrate the performance of various indicators for non-convex flux, we consider the Buckley–Leverett case with the flux function

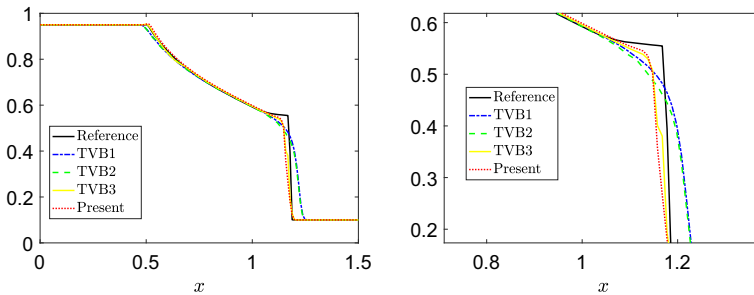
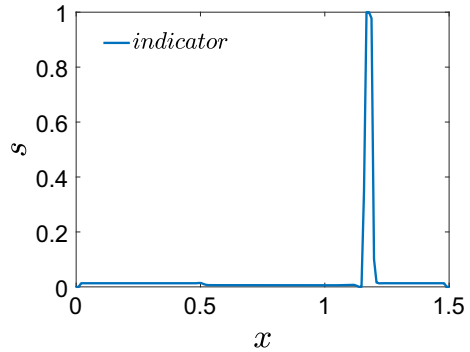
$$f(u) = \frac{u^2}{u^2 + 0.5(1 - u)^2},$$

where  $u$  represents the water saturation in a mixture of oil and water [18]. We consider the initial condition

$$u_0(x) = \begin{cases} 0.95 & \text{if } x \leq 0.5 \\ 0.1 & \text{if } x > 0.5, \end{cases}$$

on the domain  $[0, 1.5]$ , which evolves into a compound wave consisting of a shock wave and a rarefaction wave. The numerical solutions are evaluated at time  $T = 0.4$ , on a mesh with  $N = 150$  and open boundary conditions and under DG(p3) framework. The output ( $s$ ) of the present indicator is shown in Fig. 8. Solutions of Buckley–Leverett case with various indicators are compared in Fig. 9a, b, as one can observe, the solutions with the present indicator are superior to those obtained with TVB-3 indicator which are superior to those obtained by TVB-1 and TVB-2 indicators. The time-history of flagged troubled-cells with various indicators are also compared in Fig. 10a–d, the present indicator detects the zones

**Fig. 8** The Buckley–Leverett case, the output ( $s$ ) of the present indicator in each cell at  $T = 0.4$  with 150 cells



**Fig. 9** Solutions of the Buckley–Leverett case with various indicators at  $T = 0.4$  under RKDG(p3) framework with 150 cells

more regularly, and flags no cells near the rarefaction wave compared to TVB-1 and TVB-2 indicators which flag much more cells near the rarefaction wave. Note that the discontinuity appearing in this case is a contact discontinuity due to non-convex flux, the present indicator works well for this case because its mechanism of detecting troubled cell is via compressing of characteristic curves, although we never train the indicator with data set of non-convex flux.

### 6.3 The 1-D Euler Equations

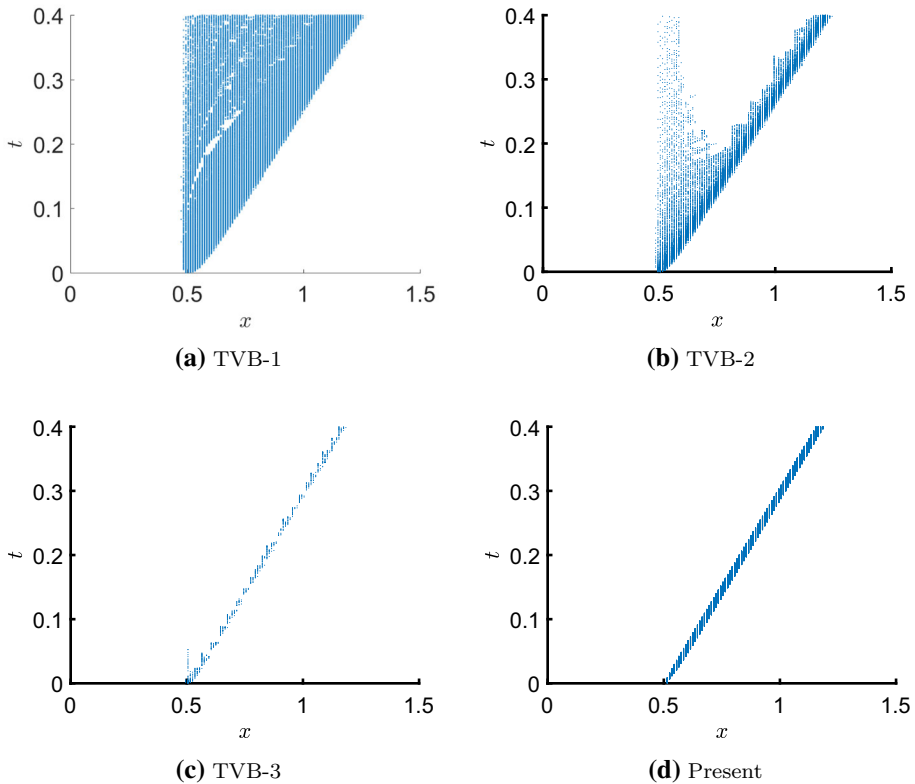
We consider the one-dimensional Euler equations to demonstrate the performance of the indicator (4.10) in system conservation laws.

#### 6.3.1 Sod Test

This problem describes a mild shock tube test proposed by Sod [33], whose initial condition is given by

$$(\rho_0, u_0, p_0)(x) = \begin{cases} (1, 0, 1) & \text{if } x < 0 \\ (0.125, 0, 0.1) & \text{if } x \geq 0, \end{cases} \quad x \in [-5, 5].$$

The solution is simulated on a mesh with  $N = 500$  cells and under DG(p3) framework, until the time  $T = 2$ . TVB-3 is unable to produce correct results due to loss of positivity of density and TVB-2 leads to highly oscillatory results. The solutions with TVB-1 and the present indicator have similar simulations and eliminate the spurious oscillations near



**Fig. 10** The time-history of flagged troubled-cells of the Buckley–Leverett case with 150 cells

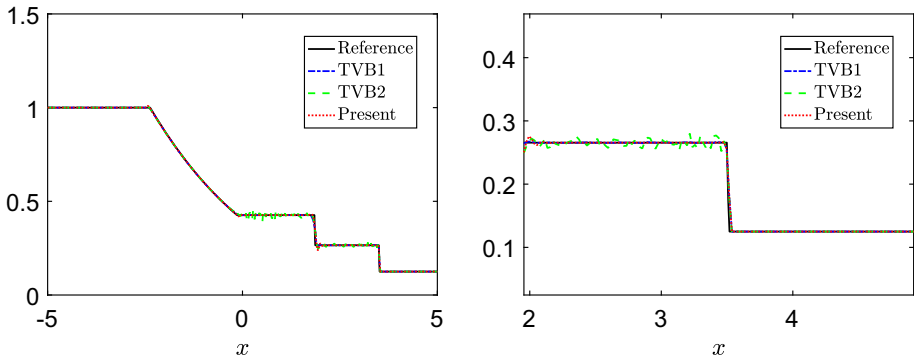
the 3-shock wave at  $x = 3.5$ , as shown in Fig. 11b. Output (s) in each cell of the present indicator (4.10) at  $T = 2$  are shown in Fig. 12. The time-history of the troubled-cells is shown in Fig. 13, we can observe the present indicator marks less troubled-cells near the 3-shock wave than TVB-1 indicator. It should be pointed out that the present indicator with 0.5 as a threshold, passes over the admissible jump of contact discontinuity because its output is less than 0.5, as shown in Fig. 11b.

### 6.3.2 Lax Test

We consider the Lax shock tube problem [17], whose initial condition is given by

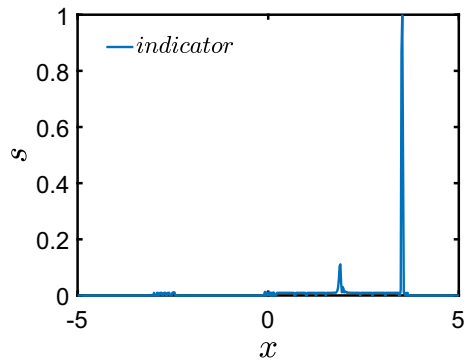
$$(\rho_0, u_0, p_0)(x) = \begin{cases} (0.445, 0.698, 3.528) & \text{if } x < 0 \\ (0.5, 0, 0.571) & \text{if } x \geq 0, \end{cases} \quad x \in [-5, 5].$$

The solution is simulated on a mesh with  $N = 200$  cells and under DG(p3) framework, until the time  $T = 1.3$ . TVB-3 is unable to produce correct results due to loss of positivity of density. The solutions with TVB-1, TVB-2 and present indicators are similar, as shown in Fig. 14. Output (s) of the present indicator (4.10) in each cell at  $T = 1.3$  is shown at Fig. 15. The time-history of the troubled-cells is shown in Fig. 16a–c, we can observe the present indicator marks similar regions as TVB-1 indicator near the 3-shock wave and marks contact discontinuity regions well.



**Fig. 11** Solutions (density) of the Sod test for the Euler equations with various indicators at  $T = 2$  under RKDG(p3) framework with 500 cells

**Fig. 12** The Sod test for the Euler equations, the output ( $s$ ) of the present indicator in each cell at  $T = 2$  with 500 cells



### 6.3.3 Left Half of the Blast-Wave

We consider a severe test case, describing the left of the blast-wave problem [37], the initial condition is given by

$$(\rho_0, u_0, p_0)(x) = \begin{cases} (1, 0, 1000) & \text{if } x < 0.5 \\ (1, 0, 0.01) & \text{if } x \geq 0.5, \end{cases} \quad x \in [0, 1].$$

the solution is simulated on a mesh with  $N = 400$  cells and under DG(p3) framework, until the time  $T = 0.012$ . Output  $s$  of the present indicator at  $T = 0.012$  is shown in (Fig. 17). As shown in Fig. 18, the results obtained by the present indicator (4.10) hold the best simulation of the wave peak compared to that obtained by TVB indicators. The time-history of the troubled-cells is shown in Fig. 19a–d, we can observe the present indicator marks much less cells than TVB-1 and TVB-2 and TVB-3, specially marked thinner zones of wave peak.

### 6.3.4 Shu–Osher Problem

This case proposed in [32], describes the interaction of a right moving shock with an oscillatory smooth wave. The initial condition is given as follows:

$$(\rho_0, u_0, p_0)(x) = \begin{cases} (3.857143, 2.629369, 10.333333) & \text{if } x < -4 \\ (1 + 0.2\sin(5x), 0, 1) & \text{if } x \geq -4, \end{cases} \quad x \in [-5, 5].$$

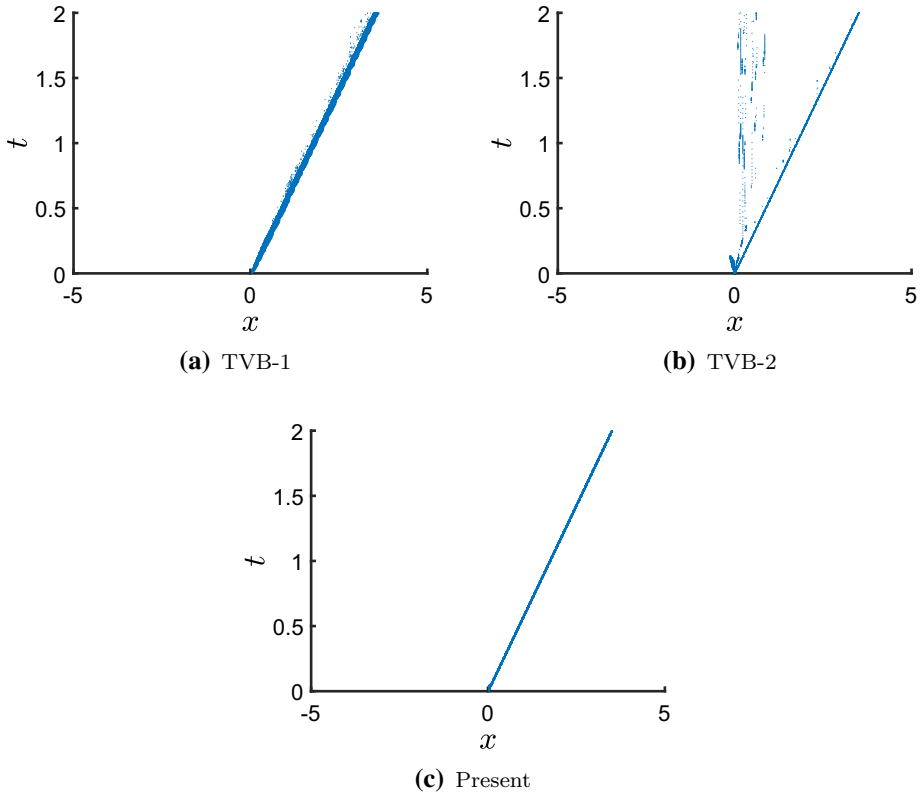


Fig. 13 The time-history of flagged troubled-cells of the Sod test for the Euler equations with 500 cells

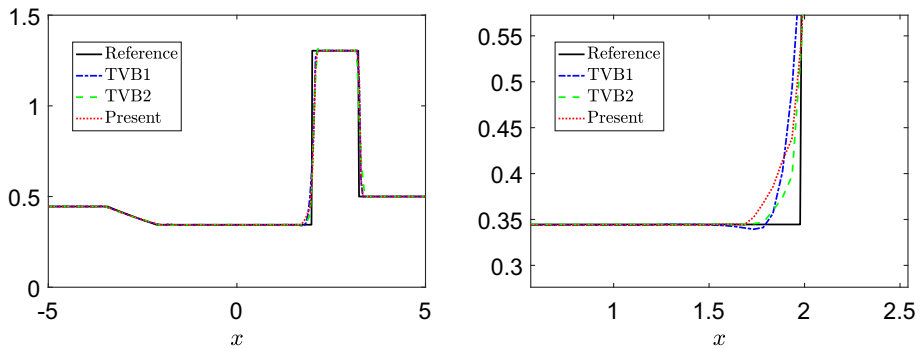
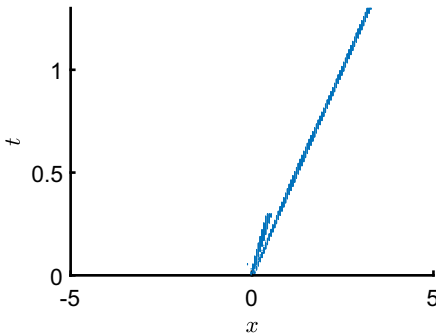
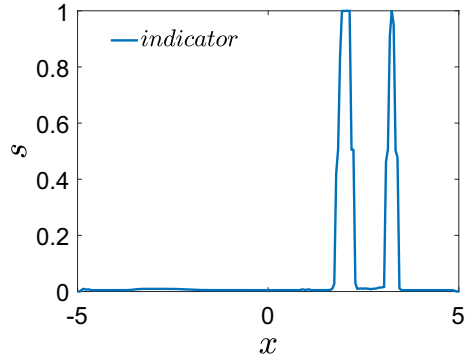
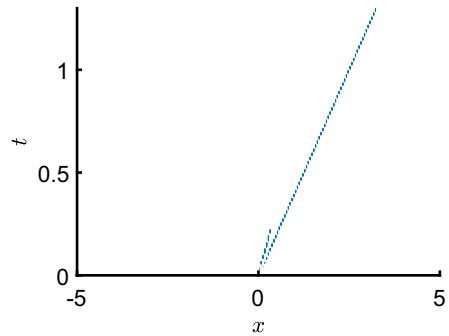


Fig. 14 Solutions (density) of the Lax test for the Euler equations with various indicators at  $T = 1.3$  under RKDG(p3) framework with 200 cells

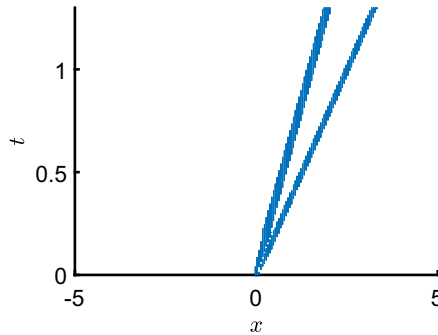
**Fig. 15** The Lax test for the Euler equations, the output ( $s$ ) of the present indicator in each cell at  $T = 1.3$  with 200 cells



**(a)** TVB-1



**(b)** TVB-2

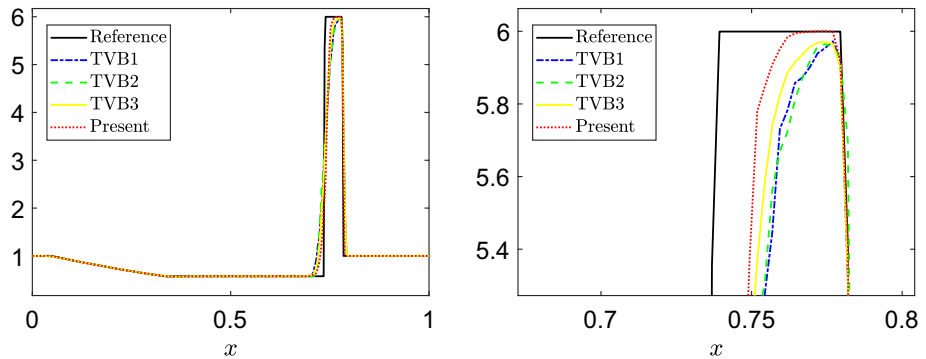
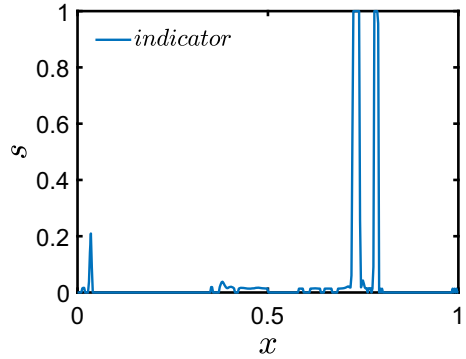


**(c)** Present

**Fig. 16** The time-history of flagged troubled-cells of the Lax test for the Euler equations with 200 cells

The solution is simulated on a mesh with  $N = 400$  cells, until the time  $T = 1.8$ . Various indicators are applied to this case under DG(p3) framework. Output  $s$  of the present indicator at  $T = 1.8$  is shown in (Fig. 20). Numerical results in Fig. 21 show that TVB-3 gives the most accurate simulations in the smooth high-frequency region, with slight oscillations. The simulations obtained by TVB-2 and the present indicators have a similar performance and are superior to that obtained by TVB-1

**Fig. 17** The Left half of the blast-wave test for the Euler equations, the output (s) of the present indicator in each cell at  $T = 0.012$  with 400 cells



**Fig. 18** Solutions (density) of the Left half of the blast-wave test for the Euler equations with various indicators at  $T = 0.012$  under RKDG(p3) framework with 400 cells

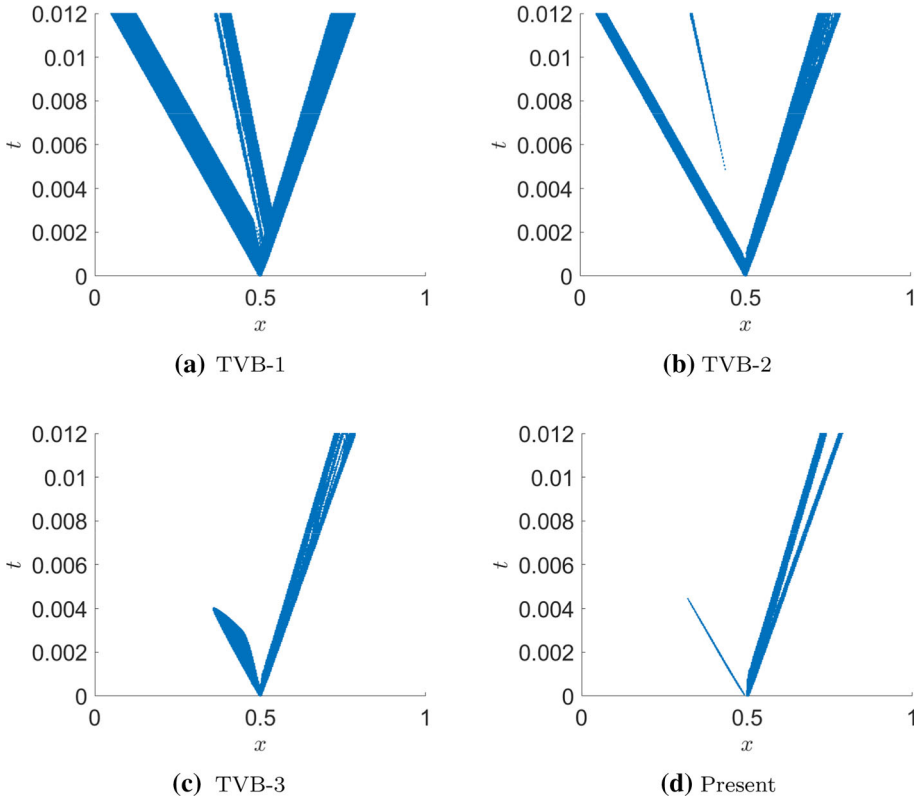
indicator. The Fig. 22 also shows the present indicator flags thinner regions than TVB-2.

### 6.3.5 1-2-3 Test

We consider the Riemann problem given by the following initial condition

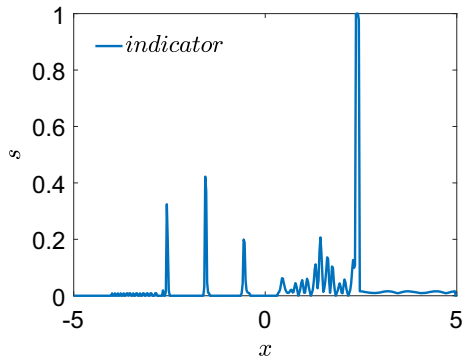
$$(\rho_0, u_0, p_0)(x) = \begin{cases} (7, -1, 0.2) & \text{if } x < 0 \\ (7, 1, 0.2) & \text{if } x \geq 0, \end{cases} \quad x \in [-1, 1].$$

The solution is simulated on a mesh with  $N = 200$  cells and under DG(p3) framework, until the time  $T = 0.6$ . The solution consists of two rarefaction waves pulling away from each other. This test case has an initial inadmissible discontinuity and the present indicator neglects it. As a result, numerical oscillations occur at the beginning, leading to the code broken down due to the loss of positivity of pressure and density. Here we give a simple strategy to solve this trouble. That is to combine a conventional indicator with the present indicator for the first few time steps, and then shift to the present indicator solely. The comparison of the solutions with the combined indicator and TVB-based indicators are shown in Fig. 23 (TVB-3 does not work out for this case). The solution with the combined indicator is slightly superior to solutions with the TVB-based indicators. We can observe the combined indicator marks no cells after the first few time steps, while



**Fig. 19** The time-history of flagged troubled-cells of the Left half of the blast-wave test for the Euler equations with 400 cells

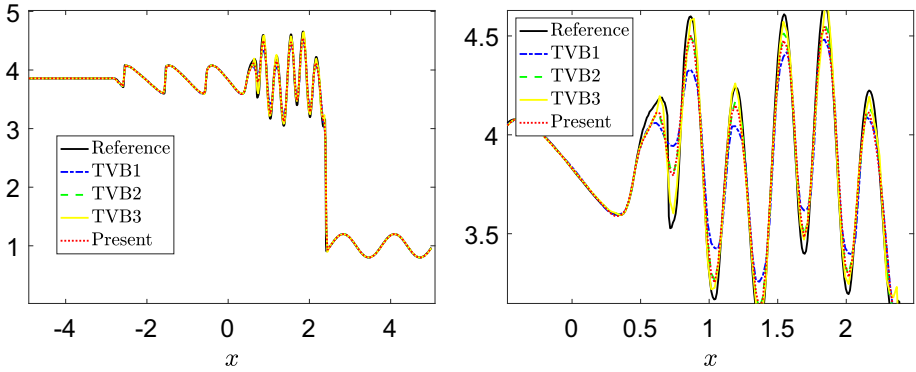
**Fig. 20** The Shu–Osher problem for the Euler equations, the output (s) of the present indicator in each cell at  $T = 1.8$  with 400 cells



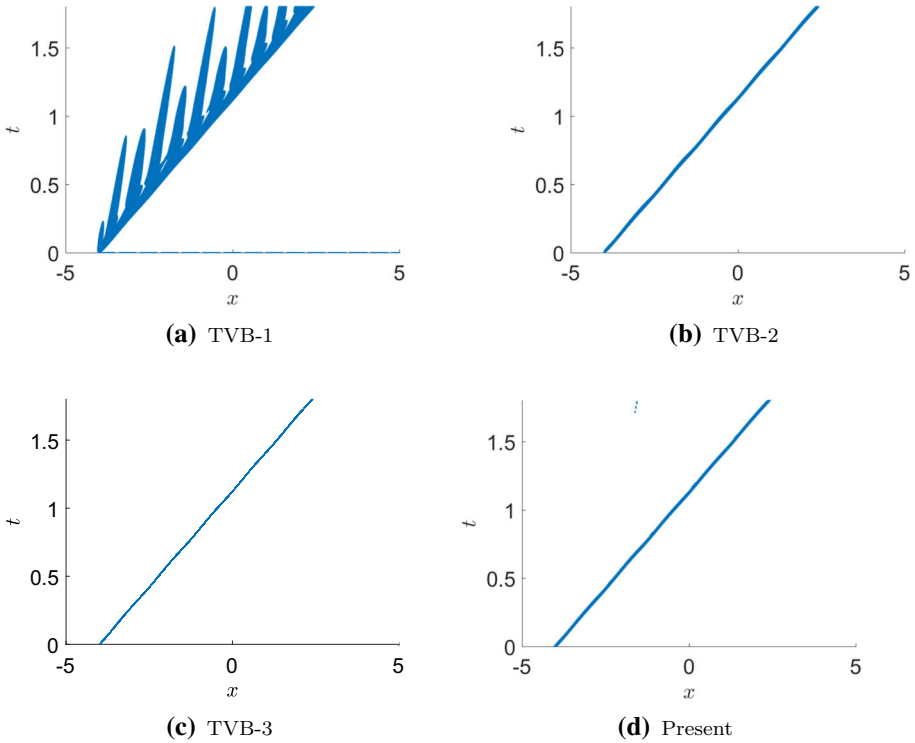
the TVB-based indicators still flag part of rarefaction waves as troubled-cells, as shown in Fig. 24.

The next two cases are to demonstrate the performance of the indicator (4.6) and (4.10) for multi-dimensional conservation laws. The present indicators (4.6) and (4.10) can be extended to multi-dimensional conservation laws by applying them to characteristic fields in different directions.

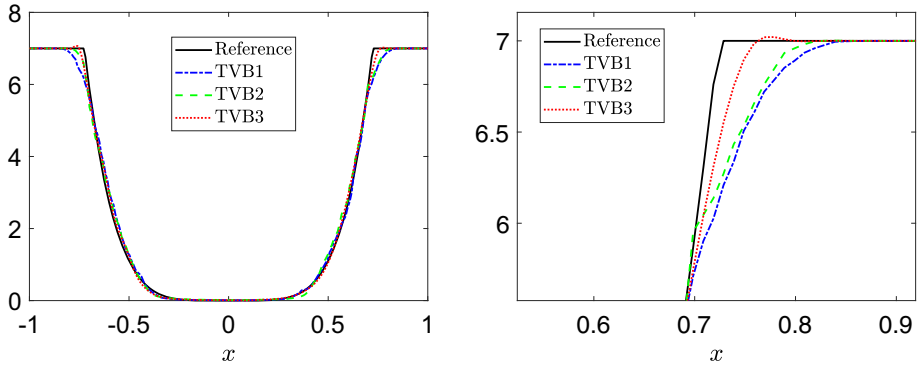




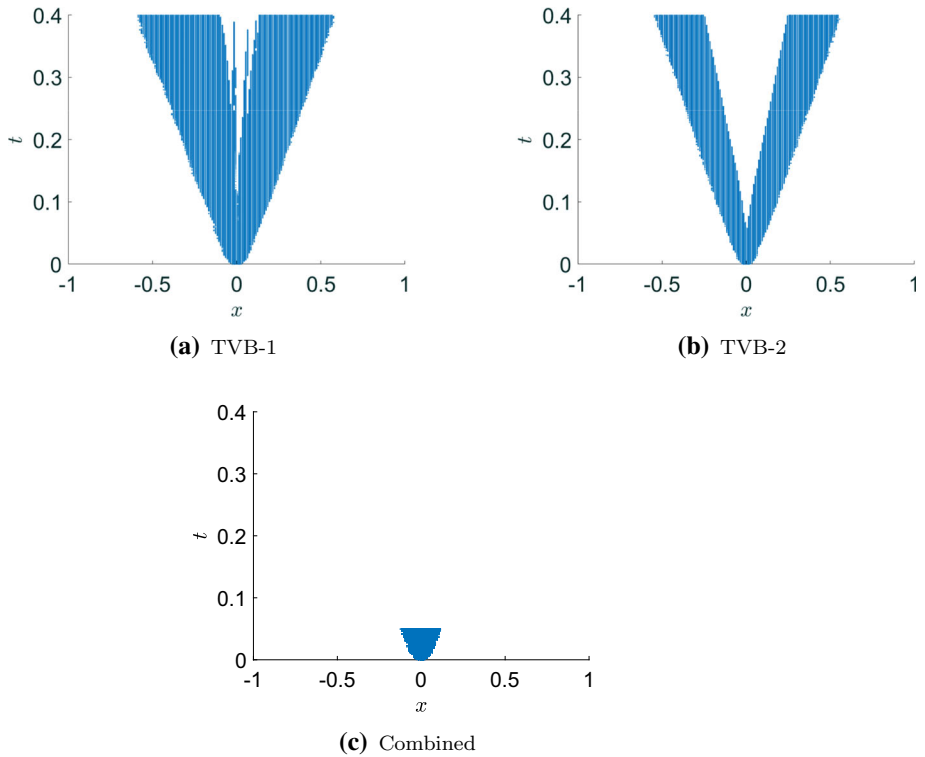
**Fig. 21** Solutions (density) of the Shu-Osher problem for the Euler equations with various indicators at  $T = 1.8$  under RKDG(p3) framework with 400 cells



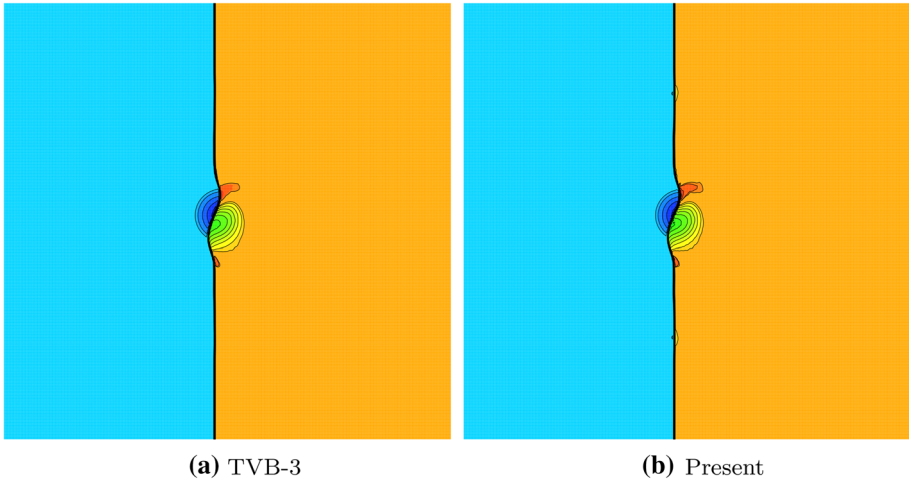
**Fig. 22** The time-history of flagged troubled-cells of the Shu-Osher problem for the Euler equations with 400 cells



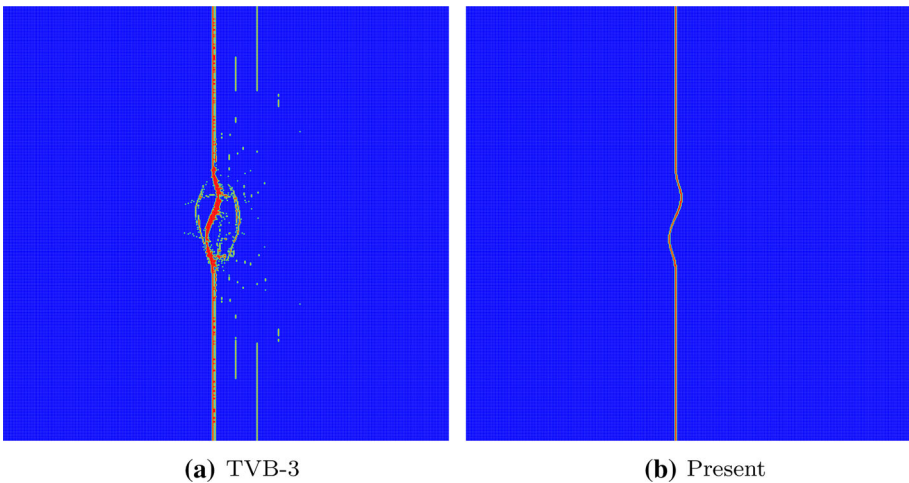
**Fig. 23** Solutions (density) of the 1-2-3 test for the Euler equations with various indicators at  $T = 0.4$  under RKDG(p3) framework with 200 cells



**Fig. 24** The time-history of flagged troubled-cells of the 1-2-3 test for the Euler equations with 200 cells



**Fig. 25** Solutions (density) of the shock–vortex interaction case for the 2-D Euler equations with various indicators with  $480 \times 480$  cells at  $T = 0.2$  under RKDG(p2) framework



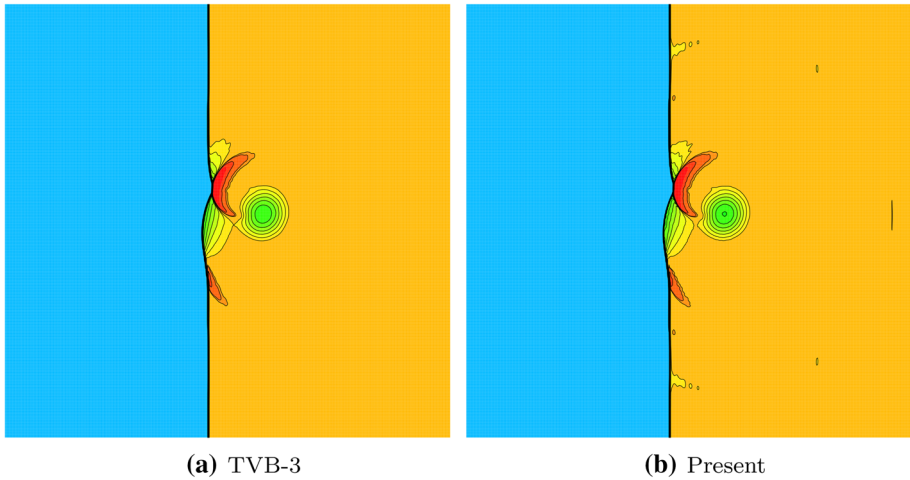
**Fig. 26** troubled-cells flagged by various indicators in the shock–vortex interaction case for 2-D Euler equations with  $480 \times 480$  cells at  $TT = 0.2$  under RKDG(p2) framework

## 6.4 The 2-D Euler Equations

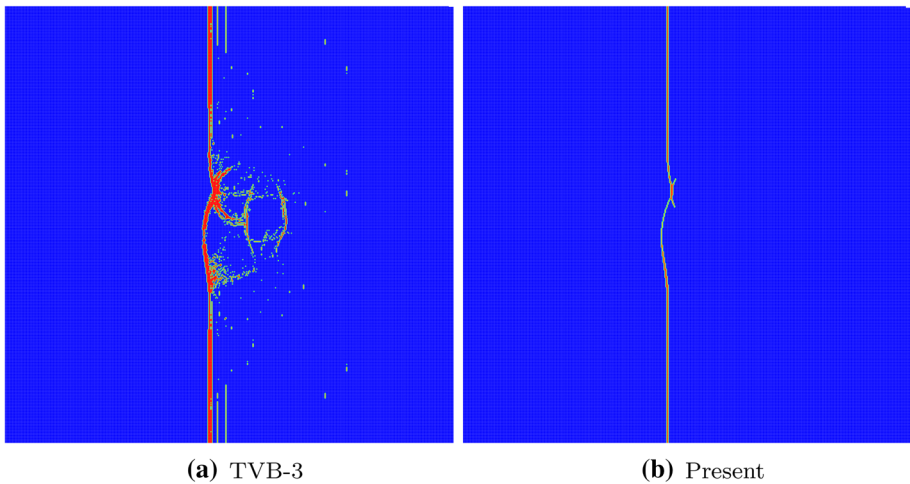
### 6.4.1 Shock–Vortex Interaction

This problem consists of interaction of a left-moving shock wave with a right-moving vortex [9]. The initial shock discontinuity on the domain  $[0, 1] \times [0, 1]$  is given by

$$U_0(x) = \begin{cases} U_L & \text{if } x < 0.5 \\ U_R & \text{if } x \geq 0.5, \end{cases}$$



**Fig. 27** Solutions (density) of the shock–vortex interaction case for the 2-D Euler equations with various indicators with  $480 \times 480$  cells at  $Tt = 0.35$  under RKDG(p2) framework

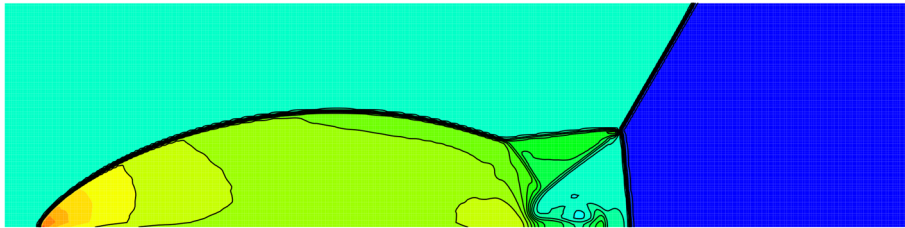


**Fig. 28** troubled-cells flagged by various indicators in the shock–vortex interaction case for 2-D Euler equations with  $480 \times 480$  cells at  $T = 0.35$  under RKDG(p2) framework

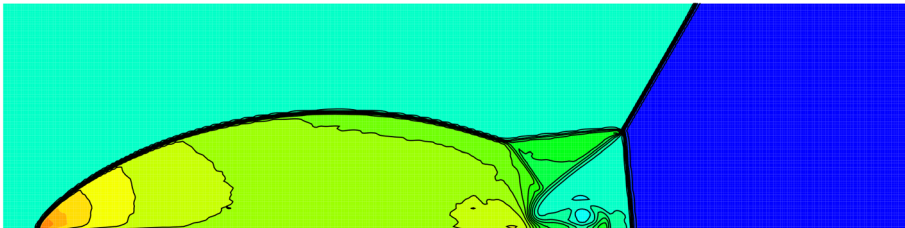
where the left state is given by  $(\rho_L, u_L, v_L, p_L) = (1, \sqrt{\gamma}, 0, 1)$  while the right state is given by

$$\begin{aligned}
 p_R &= 1.3, & \rho_R &= \rho_L \left( \frac{\gamma - 1 + (\gamma + 1)p_R}{\gamma + 1 + (\gamma - 1)p_R} \right), \\
 u_R &= \sqrt{\gamma} + \sqrt{2} \left( \frac{1 - p_R}{\sqrt{\gamma - 1 + (\gamma + 1)p_R}} \right), & v_R &= 0.
 \end{aligned}$$

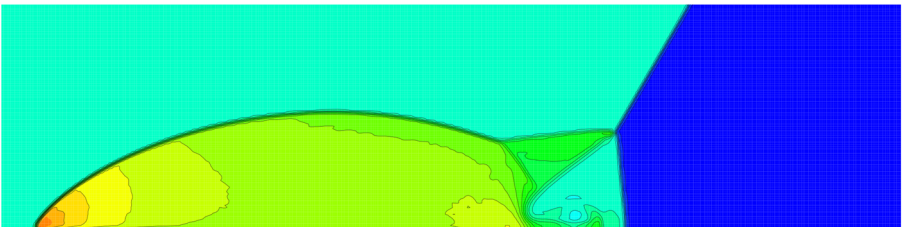
The left state  $\mathbf{U}_L$  is superposed onto an isentropic vortex described by the following perturbations in the velocity, temperature and physical entropy respectively



(a) TVB-3



(b) Present



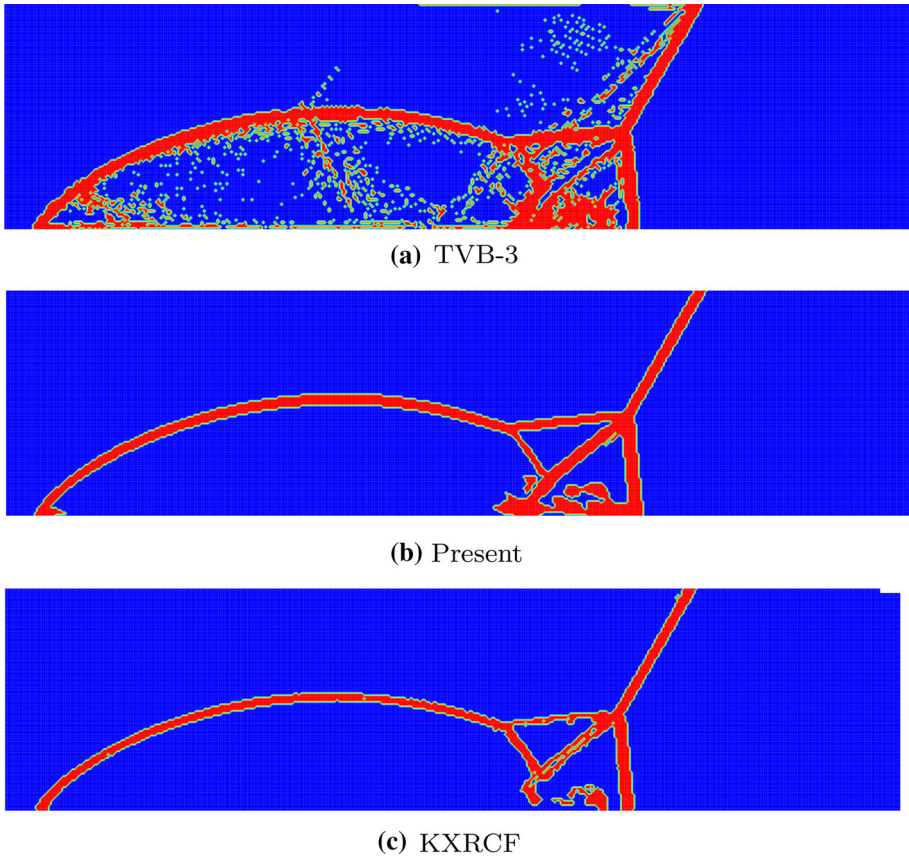
(c) KXRCF

**Fig. 29** Solutions (density) of the Double Mach reflection case for the 2-D Euler equations with various indicators with  $400 \times 100$  cells at  $T = 0.2$  under RKDG(p2) framework

$$\delta u = \theta \frac{(y - y_c)}{r_c} \exp(\beta(1 - r^2)), \quad \delta v = -\theta \frac{(x - x_c)}{r_c} \exp(\beta(1 - r^2))$$

$$\delta T = -\frac{\gamma - 1}{4\beta\gamma} \theta^2 \exp(2\beta(1 - r^2)), \quad \delta s = 0.$$

where  $r^2 = ((x - x_c)^2 + (y - y_c)^2)/r_c^2$ . The various parameters of the perturbation are chosen as  $\theta = 0.3, r_c = 0.05, \beta = 0.204$  and  $(x_c, y_c) = (0.25, 0.5)$ . The domain is discretized by  $480 \times 480$  structured grid with inflow boundary conditions applied on the left/right and reflective boundary conditions on the top/bottom. The simulation is under RKDG(p2) framework and till a final time  $T = 0.35$ . The solution and flagged troubled cells at different time instances are shown in Figs. 25, 26, 27 and 28. Note that the present indicator marked only the moving shock wave and never marked regions near the vortex, which might explain that solution with the present indicator preserves the vortex structure more precisely than that with TVB-3 indicator.

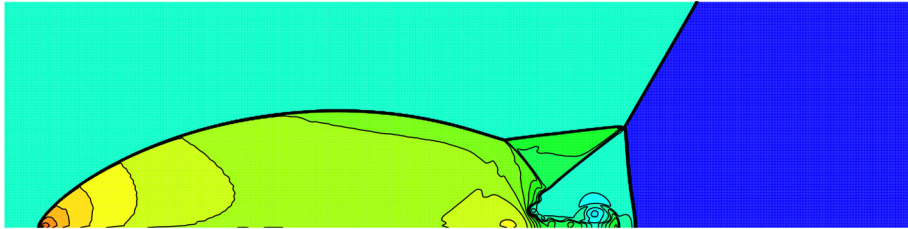


**Fig. 30** troubled-cells flagged by various indicators in the Double Mach reflection problem for 2-D Euler equations with  $400 \times 100$  cells at  $T = 0.2$  under RKDG(p2) framework

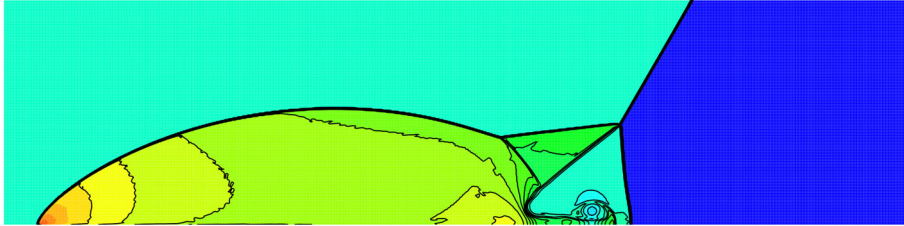
### 6.4.2 Double Mach Reflection Test

This problem is originally from [37], the computational domain for this problem is  $[0, 4] \times [0, 1]$ . The reflecting wall lies at the bottom, starting from  $x = \frac{1}{6}$ . Initially a right-moving Mach 10 shock wave is positioned at  $x = \frac{1}{6}$ ,  $y = 0$  and make a  $60^\circ$  angle with the x-axis. For the bottom boundary, the exact post-shock condition is imposed for the part from  $x = 0$  to  $x = \frac{1}{6}$  and a reflective boundary condition is used for the rest. At the top boundary, the flow values are set to describe the exact motion of a Mach 10 shock. We compute the solution up to  $T = 0.2$  and plot only the simulation results on  $400 \times 100$  cells and  $960 \times 240$  cells under RKDG(p2) framework, all the figures are shown with 23 equally spaced density contours from 1.5 to 22.7.

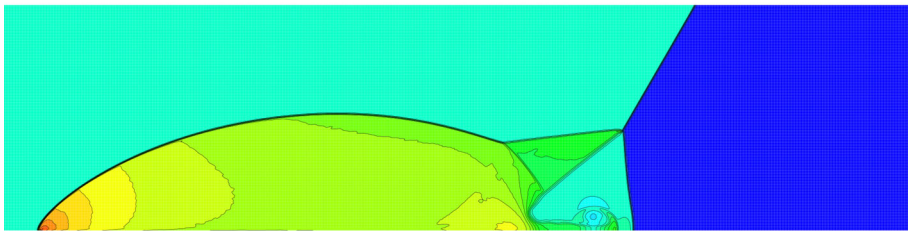
The results for this problem show an excellent performance of the present indicator. Simulations by various TVB indicators have a better performance with the increased parameter  $M$  in the TVB indicator, we only compare the simulations by the present indicator with that by the TVB-3 indicator. Solutions (density) have a significant improvement with the present indicator near bottom boundary as shown in Figs. 29 and 31. The flagged troubled-cells detected by various indicators are compared in Figs. 30 and 32, we can observe the



(a) TVB-3



(b) Present



(c) KXRFC

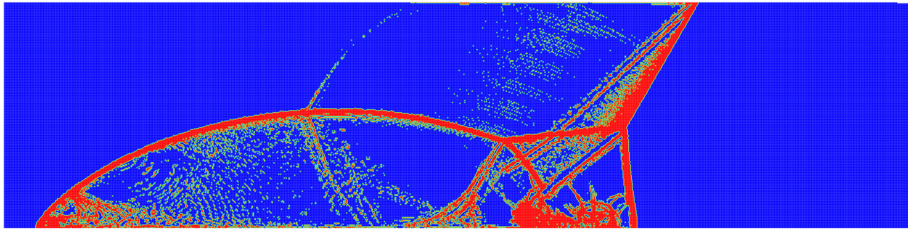
**Fig. 31** Solutions (density) of the Double Mach reflection case for the 2-D Euler equations with various indicators with  $960 \times 240$  cells at  $T = 0.2$  under RKDG(p2) framework

flagged regions detected by the present indicator are almost the discontinuity position of the solution, which are much more accurate than the TVB-3 indicator, and the flagged cells are significantly reduced compared to TVB-3 indicator, which might explain the significant improvement near bottom boundary under  $960 \times 240$  cells. The results with KXRFC shock indicator are presented in Figs. 29, 30, 31 and 32.(c), it shows the KXRFC indicator has similar behaviors with  $400 \times 100$  cells and marked extra cells at the bottom boundary with  $960 \times 240$  cells as compared to the present indicator.

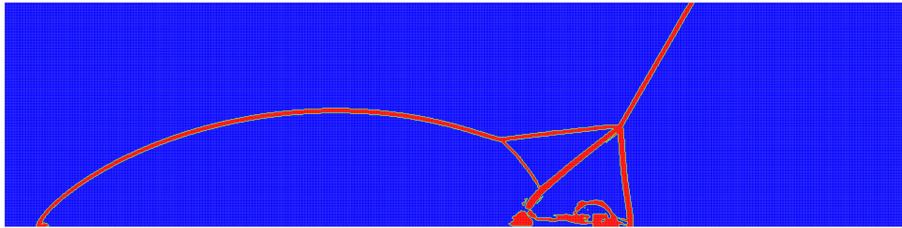
## 7 Conclusion

In this work, we proposed a novel shock wave indicator that can be applied to various numerical schemes such as the RKDG scheme. The present indicator was obtained initially by training an AN, then modified and extended to multi-dimensional conservation laws.

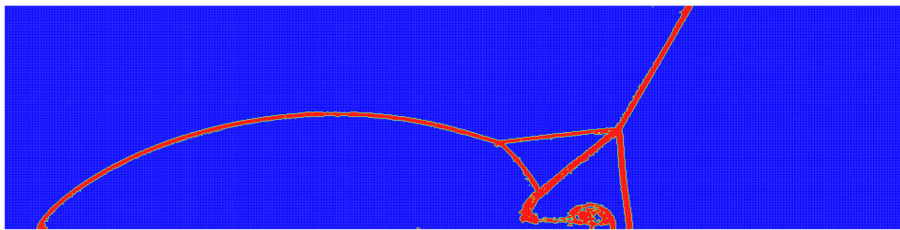
The present indicator is concise, explicable and generalizable. We proved the present indicator can identify and detect the troubled-cells where exist solution discontinuities caused



(a) TVB-3



(b) Present



(c) KXRCF

**Fig. 32** troubled-cells flagged by various indicators in the Double Mach reflection problem for 2-D Euler equations with  $960 \times 240$  cells at  $T = 0.2$  under RKDG(p2) framework

by compressing or intersecting of the characteristic curves, smooth extrema and inadmissible large gradient regions cannot be flagged by the present indicator. Several numerical results in this work demonstrated that the present indicator can improve the indicating efficiency, leading to a reduced computational cost and more accurate numerical results.

Because the present indicator detect troubled-cells by capturing the compressing or intersecting of characteristic curves, theoretically it might not detect the initial inadmissible discontinuities. One way to treat the initial inadmissible discontinuities is to follow the approach done for the 1-2-3 case in this work, the other way is to generate related data into database and deepen the neural network as presented in [28,29], but the resultant indicator might be difficult to be explicable. In addition, we believe that the present indicator can be extended to unstructured mesh.

**Acknowledgements** The authors would like to acknowledge the reviewers for the valuable comments and suggestions. This work is supported by National Numerical Wind Tunnel Project, National Nature Science Foundation of China (Nos. U1730118 and 91530325), Fundamental Research of Civil Aircraft (No. MJ-F-2012-04).



**Funding** This work was funded by National Numerical Wind Tunnel Project, National Nature Science Foundation of China (Nos. U1730118 and 91530325), Fundamental Research of Civil Aircraft (No. MJ-F-2012-04).

## References

1. Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M.: TensorFlow, Large-scale machine learning on heterogeneous systems, (2015) <https://www.tensorflow.org/>
2. Chang, T., Hsiao, L.: The Riemann problem and interaction of waves in gas dynamics. NASA STI/Recon Technical Report A 90 (1989)
3. Cheng, J., Du, Z., Lei, X., Wang, Y., Li, J.: A two-stage fourth-order discontinuous Galerkin method based on the GRP solver for the compressible Euler equations. *Comput. Fluids* **181**, 248–258 (2019)
4. Cockburn, B., Hou, S., Shu, C.-W.: The Runge–Kutta local projection discontinuous Galerkin finite element method for conservation laws. IV. The multidimensional case. *Math. Comput.* **54**, 545–581 (1990)
5. Cockburn, B., Lin, S.-Y., Shu, C.-W.: TVB Runge–Kutta local projection discontinuous Galerkin finite element method for conservation laws III: one-dimensional systems. *J. Comput. Phys.* **84**, 90–113 (1989)
6. Cockburn, B., Shu, C.-W.: TVB Runge–Kutta local projection discontinuous Galerkin finite element method for conservation laws. II. General framework. *Math. Comput.* **52**, 411–435 (1989)
7. Cockburn, B., Shu, C.-W.: The Runge–Kutta discontinuous Galerkin method for conservation laws V: multidimensional systems. *J. Comput. Phys.* **141**, 199–224 (1998)
8. Dafermos, C.M.: *Hyperbolic Conservation Laws in Continuum Physics*, volume 325 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences] (2010)
9. Fjordholm, U.S., Mishra, S., Tadmor, E.: Arbitrarily high-order accurate entropy stable essentially nonoscillatory schemes for systems of conservation laws. *SIAM J. Numer. Anal.* **50**, 544–573 (2012)
10. Gottlieb, S., Shu, C.W., Tadmor, E.: Strong stability-preserving high-order time discretization methods. *SIAM Rev.* **43**, 89–112 (2001)
11. Gurney, K.: *An Introduction to Neural Networks[M]*. UCL Press, Boca Raton (1997)
12. Harten, A.: On a class of high resolution total-variation-stable finite-difference schemes. *SIAM J. Numer. Anal.* **21**, 1–23 (1984)
13. Harten, A., Osher, S.: Uniformly high-order accurate nonoscillatory schemes. I. In: *Upwind and High-Resolution Schemes*, pp. 187–217. Springer (1997)
14. Hesthaven, J.S., Warburton, T.: *Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications*. Springer, Berlin (2007)
15. Jameson, A., Schmidt, W., Turkel, E.: Numerical solution of the Euler equations by finite volume methods using Runge Kutta time stepping schemes, In: *14th Fluid and Plasma Dynamics Conference*, p. 1259 (1981)
16. Krivodonova, L., Xin, J., Remacle, J.-F., Chevaugeon, N., Flaherty, J.E.: Shock detection and limiting with discontinuous Galerkin methods for hyperbolic conservation laws. *Appl. Numer. Math.* **48**, 323–338 (2004)
17. Lax, P.: Weak solutions of nonlinear hyperbolic equations and their numerical computation. *Commun. Pure Appl. Math.* **7**, 159–193 (1954)
18. Leveque, R.J.: Finite volume methods for hyperbolic problems. *Meccanica* **39**, 88–89 (2004)
19. Nair, V., Hinton, G.E.: Rectified linear units improve restricted boltzmann machines. In: *International Conference on International Conference on Machine Learning* (2010)
20. Nasr, G.E., Badr, E.A., Joun, C.: Cross entropy error function in neural networks: forecasting gasoline demand. In: *Fifteenth International Florida Artificial Intelligence Research Society Conference* (2002)
21. Nowlan, S.J., Hinton, G.E.: Simplifying Neural Networks by Soft Weight-Sharing. MIT Press, Cambridge (1992)
22. Oh, S.H.: Error back-propagation algorithm for classification of imbalanced data. *Neurocomputing* **74**, 1058–1061 (2011)
23. Pitts, W.: A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biol.* **52**, 99–115 (1990)
24. Qiu, J., Shu, C.-W.: Hermite weno schemes and their application as limiters for Runge–Kutta discontinuous galerkin method: one-dimensional case. *J. Comput. Phys.* **193**, 115–135 (2004)
25. Qiu, J., Shu, C.-W.: A comparison of troubled-cell indicators for Runge–Kutta discontinuous Galerkin methods using weighted essentially nonoscillatory limiters. *SIAM J. Sci. Comput.* **27**, 995–1013 (2005)
26. Qiu, J., Shu, C.-W.: Hermite weno schemes and their application as limiters for Runge–Kutta discontinuous galerkin method ii: two dimensional case. *Comput. Fluids* **34**, 642–663 (2005)

27. Qiu, J., Shu, C.-W.: Runge–Kutta discontinuous Galerkin method using WENO limiters. *SIAM J. Sci. Comput.* **26**, 907–929 (2005)
28. Ray, D., Hesthaven, J.S.: An artificial neural network as a troubled-cell indicator. *J. Comput. Phys.* **367**, 166–191 (2018)
29. Ray, D., Hesthaven, J.S.: Detecting troubled-cells on two-dimensional unstructured grids using a neural network. *J. Comput. Phys.* **397**, 108845 (2019)
30. Schmidhuber, J.: Deep learning in neural networks: an overview. *Neural Netw.* **61**, 85–117 (2015)
31. Shu, C.-W.: Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws. In: *Advanced Numerical Approximation of Nonlinear Hyperbolic Equations*, pp. 325–432. Springer (1998)
32. Shu, C.-W., Osher, S.: Efficient implementation of essentially non-oscillatory shock-capturing schemes, ii. In: *Upwind and High-Resolution Schemes*, pp. 328–374. Springer (1989)
33. Sod, G.A.: A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws. *J. Comput. Phys.* **27**, 1–31 (1978)
34. Tadmor, E.: Entropy stability theory for difference approximations of nonlinear conservation laws and related time-dependent problems. *Acta Numer.* **12**, 451–512 (2003)
35. Toro, E.F.: *Riemann Solvers and Numerical Methods for Fluid Dynamics*, pp. 87–114. Springer, Berlin (2013)
36. Van Leer, B.: Towards the ultimate conservative difference scheme. V. A second-order sequel to Godunov’s method. *J. Comput. Phys.* **32**, 101–136 (1979)
37. Woodward, C.P.: The numerical simulation of two-dimensional fluid flow with strong shocks. *J. Comput. Phys.* **54**, 115–173 (1984)

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.