# High Order Well-Balanced Generalized WENO-Interpolation Schemes

The Report in Advanced Research in Applied Mathematics and Scientific Computing 2019, OUC, China

Bao-Shan Wang , School of Mathematical Sciences , Ocean University of China , Qingdao , China

## Shallow Water Equations

The 1D shallow water equation (SWE) is given by

$$\mathbf{Q}_t + \mathbf{F}(\mathbf{Q})_x = \mathbf{G}, \tag{1}$$

where $\mathbf{Q} = (h, hu)^{\mathrm{T}}$, $\mathbf{F}(\mathbf{Q}) = (hu, hu^2 + \frac{1}{2}gh^2)^{\mathrm{T}}$, $\mathbf{G} = (0, -ghb_x)^{\mathrm{T}}$, where $b$ is the vertical height of the bottom topography, from an arbitrary level of reference, $h$ is the water depth above the bottom topography, $u$ is the velocity, and $g = 9.812$ is the gravitational constant.

The SWE (1) has the set of steady states ($a$ is a constant),

$$\mathbf{G}_E = \{\mathbf{Q} = (h, hu)^{\mathrm{T}} |\ h + b = a,\ \ u = 0\}. \tag{2}$$

If $\mathbf{Q} \in \mathbf{G}_E$, then the source term in (1) can be written as

$$\begin{aligned}\mathbf{G} &= \begin{bmatrix} 0 \\ -ghb_x \end{bmatrix} = \sum_{k=1}^{2} \mathbf{A}_k \mathbf{S}_k(\mathbf{B})_x \\ &= \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{2}g \end{bmatrix} \begin{bmatrix} 0 \\ b^2 \end{bmatrix}_x + \begin{bmatrix} 0 & 0 \\ 0 & -ga \end{bmatrix} \begin{bmatrix} 0 \\ b \end{bmatrix}_x, \end{aligned} \tag{3}$$

and the left and right Roe-averaged eigenvectors become

$$\tilde{\mathbf{L}} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2}(\frac{1}{\tilde{c}}) \\ \frac{1}{2} & \frac{1}{2}(\frac{1}{\tilde{c}}) \end{bmatrix}, \quad \tilde{\mathbf{R}} = \begin{bmatrix} 1 & 1 \\ -\tilde{c} & \tilde{c} \end{bmatrix},$$

where $\tilde{h} = \left(\frac{1}{2}(\sqrt{h_i} + \sqrt{h_{i+1}})\right)^2$, $\tilde{u} = 0$, $\tilde{c} = \sqrt{g\tilde{h}}$.

## WENO Interpolation (WI)

The fifth order WENO interpolation of a function $Q(x)$ at $x_{i+\frac{1}{2}}$ is built through

$$Q_{i+\frac{1}{2}} = \sum_{k=0}^{2} \omega_k Q_{i+\frac{1}{2}}^{(k)}, \tag{4}$$

where the local low order interpolations are

$$\begin{aligned} Q_{i+\frac{1}{2}}^{(0)} &= \tfrac{3}{8}Q_{i-2} - \tfrac{5}{4}Q_{i-1} + \tfrac{15}{8}Q_i, \\ Q_{i+\frac{1}{2}}^{(1)} &= -\tfrac{1}{8}Q_{i-1} + \tfrac{3}{4}Q_i + \tfrac{3}{8}Q_{i+1}, \\ Q_{i+\frac{1}{2}}^{(2)} &= \tfrac{3}{8}Q_i + \tfrac{3}{4}Q_{i+1} - \tfrac{1}{8}Q_{i+2}. \end{aligned}$$

And the nonlinear weights $\omega_k$ are defined by

$$\omega_k = \frac{\alpha_k}{\sum_{s=0}^{2}\alpha_s}, \quad \alpha_k = d_k\left(1 + \left(\frac{\tau_5}{\beta_k + \epsilon}\right)^p\right), \tag{5}$$

where $\{d_0 = \frac{1}{16}, d_1 = \frac{5}{8}, d_2 = \frac{5}{16}\}$ are the linear (optimal) weights and the sensitivity parameter $\epsilon = 10^{-12}$ is used to avoid a division of zero. The power parameter $p = 2$ is used to increase the scale separation between the local lower order smoothness indicators

$$\beta_0 = \tfrac{13}{12}(Q_{i-2} - 2Q_{i-1} + Q_i)^2 + \tfrac{1}{4}(Q_{i-2} - 4Q_{i-1} + 3Q_i)^2,$$
$$\beta_1 = \tfrac{13}{12}(Q_{i-1} - 2Q_i + Q_{i+1})^2 + \tfrac{1}{4}(Q_{i-1} - Q_{i+1})^2,$$
$$\beta_2 = \tfrac{13}{12}(Q_i - 2Q_{i+1} + Q_{i+2})^2 + \tfrac{1}{4}(3Q_i - 4Q_{i+1} + 3Q_{i+2})^2,$$

and global smoothness indicator $\tau_5 = |\beta_0 - \beta_2|$.

Denote $Q_{i+\frac{1}{2}}$ in (4) by $Q^-_{i+\frac{1}{2}}$ since the stencil is biased to the left. The $Q^+_{i+\frac{1}{2}}$ is mirror-symmetric to $Q^-_{i+\frac{1}{2}}$ with respect to $x_{i+\frac{1}{2}}$.

## Generalized WI Schemes

Generalized WI (GWI) schemes can be expressed as

$$\begin{aligned}\Delta x \sum_{k=-K_1}^{K_1} \alpha_k F'_{i+k} \ &= \ \sum_{k=0}^{K_2} \beta_k \left(\hat{F}_{i+k+\frac{1}{2}} - \hat{F}_{i-k-\frac{1}{2}}\right) \\ &+ \ \sum_{k=1}^{K_3} \gamma_k \left(F_{i+k} - F_{i-k}\right), \end{aligned} \tag{6}$$

where the parameters are given in the table below.

| | WCN[1] | WCN-E[1] | HWCN-E[2] | AWENO[3, 4] |
|---|---|---|---|---|
| $K_1$ | 1 | 0 | 0 | 0 |
| $K_2$ | 1 | 0 | 0 | 0 |
| $K_3$ | – | – | 3 | 3 |
| $\alpha_{-1}$ | $\alpha$ | – | – | – |
| $\alpha_0$ | 1 | 1 | 1 | 1 |
| $\alpha_1$ | $\alpha$ | – | – | – |
| $\beta_0$ | $\frac{225-206\alpha}{192}$ | $\frac{75}{64}$ | $\alpha$ | 1 |
| $\beta_1$ | $\frac{414\alpha-25}{384}$ | $-\frac{25}{384}$ | – | – |
| $\beta_2$ | – | $\frac{3}{640}$ | – | – |
| $\gamma_1$ | – | – | $\frac{192-175\alpha}{256}$ | $\frac{17}{256}$ |
| $\gamma_2$ | – | – | $\frac{35\alpha-48}{320}$ | $-\frac{13}{320}$ |
| $\gamma_3$ | – | – | $\frac{64-45\alpha}{3840}$ | $\frac{19}{3840}$ |

The numerical flux on the interface in (6) is obtained by

$$\hat{F}_{i+\frac{1}{2}} = h(Q^-_{i+\frac{1}{2}}, Q^+_{i+\frac{1}{2}}), \tag{7}$$

where $h(Q^-, Q^+)$ can be used arbitrary monotone flux, such as, the Lax-Friedrichs (LF) flux and HLL/HLLC flux.

## Definitions and Operators

Denote $f$ and $g$ as unknown functions and $A$ as a constant.

**D. 1** *$\lambda = \Delta t/\Delta x$, is the inverse of numerical speed.*

**D. 2** *$\Delta^n[f] = f^{n+1} - f^n$ is the temporal change (from step $n$).*

**D. 3** *$\delta_1[f] = f^+ + f^-$ and $\delta_2[f] = f^+ - f^-$ are the sum and jump of a function at a given point, respectively.*

**D. 4** *$\Delta_i[f] = f_{i+\frac{1}{2}} - f_{i-\frac{1}{2}}$ is the divided difference operator.*

- *If and only if $f$ is a constant, $\Delta_i[f] = 0$.*

**D. 5** *$\mathcal{D}_i[f]$ is the linear high order correction operator (see the last term in (6)).*

- *If $f$ is a constant, $\mathcal{D}_i[f] = 0$.*

**D. 6** *$W[\cdot]$ is the nonlinear WI operator.*

- *$W[f+g] \neq W[f] + W[g]$, unless $f$, $g$ or $f+g$ is constant.*
- *$W[Af] \neq AW[f]$, unless $f = A = 0$ or $f = A = \pm 1$.*
- *In general, $W$ is not commutable. $\Delta_i W \neq W \Delta_i$.*

**D. 7** *Linearized nonlinear WENO operator $W^*$*

*To set the nonlinear weights $\omega^f$ and $\omega^g$ of the $W[f]$ and $W[g]$ to be the same, one can use a common smoothness indicator $\beta^*$,*

$$\beta^* = \frac{(\beta^f)^2 + (\beta^g)^2}{\beta^f + \beta^g + 10^{-16}}, \text{or (if } f = Ag,\ A \neq 0)\ \beta^f = \beta^g,$$

*in both nonlinear weights to form a linear WENO operator $W^*$.*

- *$W^*[f+g] = W^*[f] + W^*[g]$, $W^*[Ag] = AW^*[g]$.*

## Well-Balanced Conditions for GWI Schemes ($\mathbf{Q}^n \in \mathbf{G}_E$)

First, $\mathbf{Q}$ and $\mathbf{B}$ in (1) are projected into characteristic fields,

$$\tilde{\mathbf{Q}}_{i+k} \ = \ \tilde{\mathbf{L}}\mathbf{Q}_{i+k} = \left(\frac{h}{2}\right)_{i+k} = \begin{bmatrix} \tilde{q}_1 \\ \tilde{q}_2 \end{bmatrix}_{i+k}, \tag{8}$$

$$\tilde{\mathbf{B}}_{i+k} \ = \ \tilde{\mathbf{L}}\mathbf{B}_{j+k} = b_{i+k}\begin{bmatrix} -\frac{1}{2\tilde{c}} \\ \frac{1}{2\tilde{c}} \end{bmatrix} = \begin{bmatrix} \tilde{b}_1 \\ \tilde{b}_2 \end{bmatrix}_{i+k}. \tag{9}$$

Next, using the WI procedure, the reconstructed $\tilde{\mathbf{Q}}^{\pm}_{i+\frac{1}{2}}$ and $\tilde{\mathbf{B}}^{\pm}_{i+\frac{1}{2}}$ at the cell boundaries become

$$\tilde{\mathbf{Q}}^{\pm}_{i+\frac{1}{2}} \ = \ W^{\pm}_{i+\frac{1}{2}}[\tilde{\mathbf{Q}}] \triangleq [\tilde{q}_1^{\pm}, \tilde{q}_2^{\pm}]^{\mathrm{T}}_{i+\frac{1}{2}}, \tag{10}$$

$$\tilde{\mathbf{B}}^{\pm}_{i+\frac{1}{2}} \ = \ W^{\pm}_{i+\frac{1}{2}}[\tilde{\mathbf{B}}] \triangleq [\tilde{b}_1^{\pm}, \tilde{b}_2^{\pm}]^{\mathrm{T}}_{i+\frac{1}{2}}, \tag{11}$$

where $W[\cdot]$ is the nonlinear WI operator.

Then, $\tilde{\mathbf{Q}}^{\pm}_{i+\frac{1}{2}}$ and $\tilde{\mathbf{B}}^{\pm}_{i+\frac{1}{2}}$ are projected into the physical space,

$$\mathbf{Q}^{\pm}_{i+\frac{1}{2}} \ = \ \tilde{\mathbf{R}}\tilde{\mathbf{Q}}^{\pm}_{i+\frac{1}{2}} = \begin{bmatrix} \tilde{q}_1^{\pm} + \tilde{q}_2^{\pm} \\ \tilde{c}(-\tilde{q}_1^{\pm} + \tilde{q}_2^{\pm}) \end{bmatrix}_{i+\frac{1}{2}}, \tag{12}$$

$$\mathbf{B}^{\pm}_{i+\frac{1}{2}} \ = \ \tilde{\mathbf{R}}\tilde{\mathbf{B}}^{\pm}_{i+\frac{1}{2}} = \begin{bmatrix} \tilde{b}_1^{\pm} + \tilde{b}_2^{\pm} \\ \tilde{c}(-\tilde{b}_1^{\pm} + \tilde{b}_2^{\pm}) \end{bmatrix}_{i+\frac{1}{2}}, \tag{13}$$

According to $\tilde{q}_1 = \tilde{q}_2$ and $\tilde{b}_1 + \tilde{b}_2 = 0$, one has

$$\mathbf{Q}^{\pm}_{i+\frac{1}{2}} = \begin{bmatrix} 2\tilde{q}_1^{\pm} \\ 0 \end{bmatrix}_{i+\frac{1}{2}}, \quad \mathbf{F}(\mathbf{Q}^{\pm}_{i+\frac{1}{2}}) = \begin{bmatrix} 0 \\ 2g(\tilde{q}_1^{\pm})^2 \end{bmatrix}_{i+\frac{1}{2}}, \tag{14}$$

$$\mathbf{B}^{\pm}_{i+\frac{1}{2}} = \begin{bmatrix} 0 \\ 2\tilde{c}\tilde{b}_2 \end{bmatrix}_{i+\frac{1}{2}}, \quad \mathbf{S}_1(\mathbf{B}^{\pm}_{i+\frac{1}{2}}) = \begin{bmatrix} 0 \\ 4(\tilde{c}\tilde{b}_2^{\pm})^2 \end{bmatrix}_{i+\frac{1}{2}}. \tag{15}$$

A general Euler-forward time integration can be written as

$$\Delta^n[\mathbf{Q}_i] = \lambda\left(-\Delta_i(\widehat{\mathbf{F}}) - \mathcal{D}_i[\mathbf{F}] + \sum_{k=1}^{2}\left(\mathbf{A}_k\Delta_i(\widehat{\mathbf{S}}_k) + \mathcal{D}_i[\mathbf{S}_k]\right)\right).$$

The numerical flux $\widehat{\mathbf{F}}$ and the source term $\widehat{\mathbf{S}}_k$ are given by

$$\widehat{\mathbf{F}} = \tfrac{1}{2}(\delta_1[\mathbf{F}(\mathbf{Q})] - \alpha\delta_2[\mathbf{Q}]), \quad \widehat{\mathbf{S}}_k = \tfrac{1}{2}(\delta_1[\mathbf{S}_k]). \tag{16}$$

Since $\mathcal{D}_i$ is linear, we have $\mathcal{D}_i[\mathbf{F}] = \sum_{k=1}^{2}\mathbf{A}_k\mathcal{D}_i[\mathbf{S}_k]$.

**A well-balanced method is one that balances the flux and source term exactly** [5, 6] for the steady state solution,

$$\Delta_i(\delta_1[\mathbf{F}] - \alpha\delta_2[\mathbf{Q}]) - \sum_{k=1}^{2}\mathbf{A}_k\Delta_i(\delta_1[\mathbf{S}_k]) = \mathbf{0}. \tag{17}$$

The first elements of $\mathbf{F}(\mathbf{Q}^{\pm}_{i+\frac{1}{2}})$ and source term $\mathbf{S}_k$ in (17) are zero, then, we need to satisfy the **first condition**

$$\Delta_i(\delta_2[2\tilde{q}_1]) = 0. \tag{18}$$

In this work, the LF flux (16) is **modified** as

$$\widehat{\mathbf{F}} = \tfrac{1}{2}(\delta_1[\mathbf{F}(\mathbf{Q})] - \alpha\delta_2[\bar{\mathbf{Q}}]), \tag{19}$$

where $\bar{\mathbf{Q}}^{\pm} = \mathbf{Q}^{\pm} + \mathbf{C}\mathbf{B}^{\pm} = (a, 0)^{\mathrm{T}}$ where $\mathbf{C} = [0, 1; 1, 0]$ is the flipped matrix, then $\Delta_i(\delta_2[\bar{\mathbf{Q}}]) = 0$.

The **second condition** becomes

$$\Delta_i(\delta_1[\tilde{q}_1^2 - (\tilde{c}\tilde{b}_2)^2 + a\tilde{c}\tilde{b}_2]) = 0. \tag{20}$$

Using the **linearized WENO operator** by $h$, $W^h$, we have

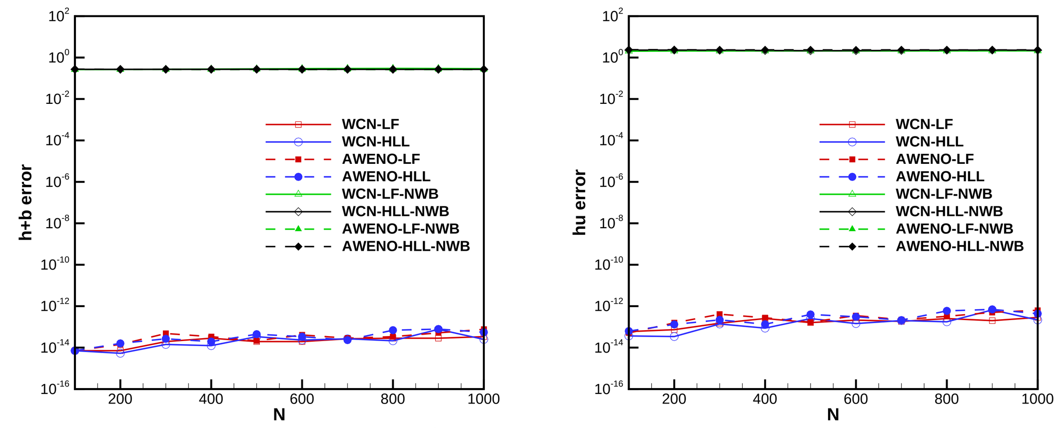
$$\begin{aligned} &\left(W^h\left[\tfrac{h}{2}\right]\right)^2 - \left(\tilde{c}W^h\left[\tfrac{b}{2\tilde{c}}\right]\right)^2 + a\tilde{c}W^h\left[\tfrac{b}{2\tilde{c}}\right] \\ &= \tfrac{1}{4}(W^h[h])^2 - \tfrac{1}{4}(W^h[b])^2 + \tfrac{1}{2}aW^h[b] \\ &= \tfrac{1}{4}(W^h[h] + W^h[b])^2 = \tfrac{1}{4}a^2 \ \ \Rightarrow \ \ (20) \text{ holds.} \end{aligned} \tag{21}$$

Note that the proof for the WCN scheme is similar with an inverse operator of linear constant coefficient matrix. The third order TVD Runge-Kutta method (CFL $= 0.45$) is used for time integration in the numerical tests.

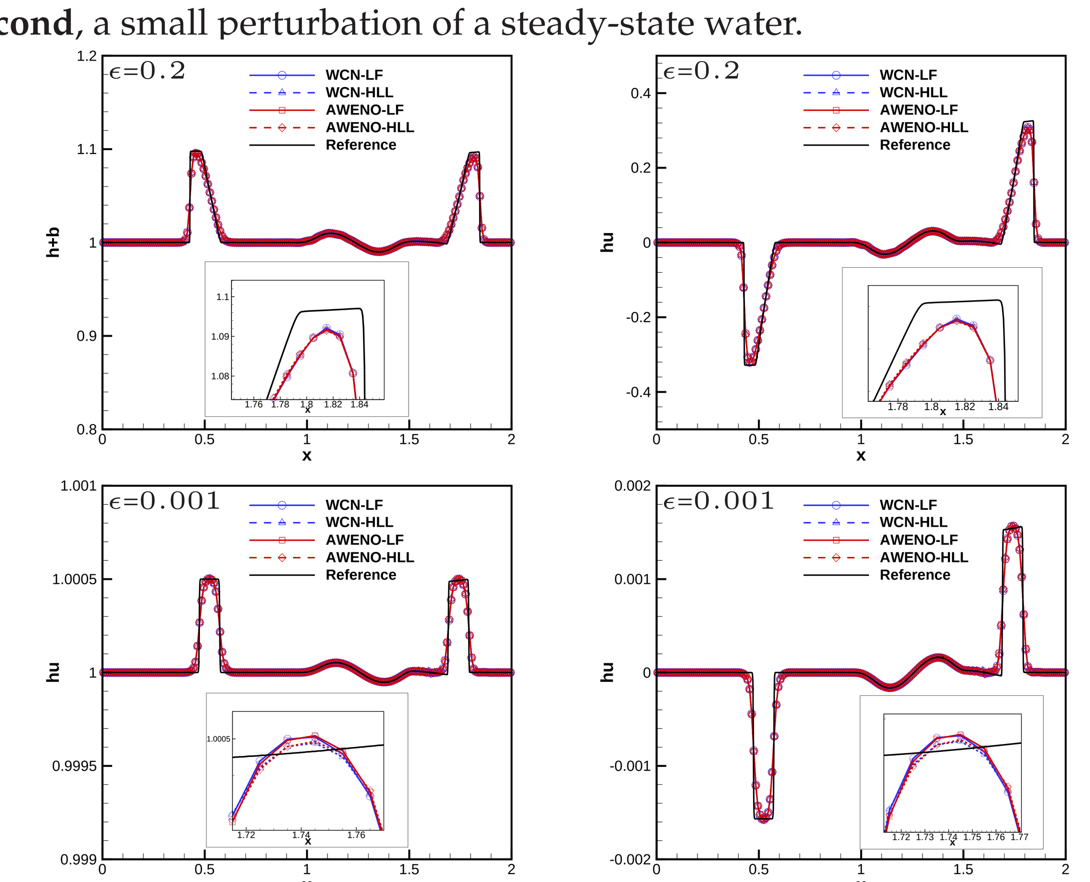
## Numerical Results

The bottom topographies and initial conditions of the examples can be found in [5]. Here, "-NWB" denotes the non-well-balanced schemes and "-C" denotes the schemes without projection.

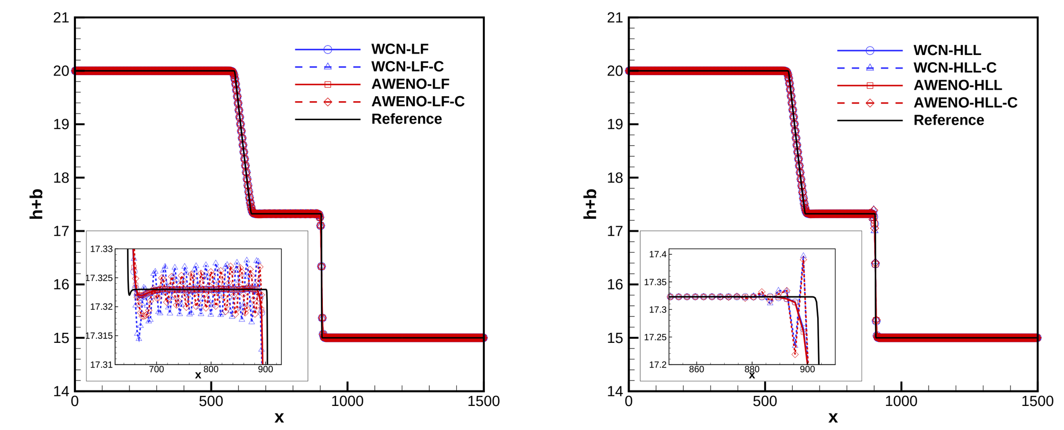
**First**, the exact C-property test.



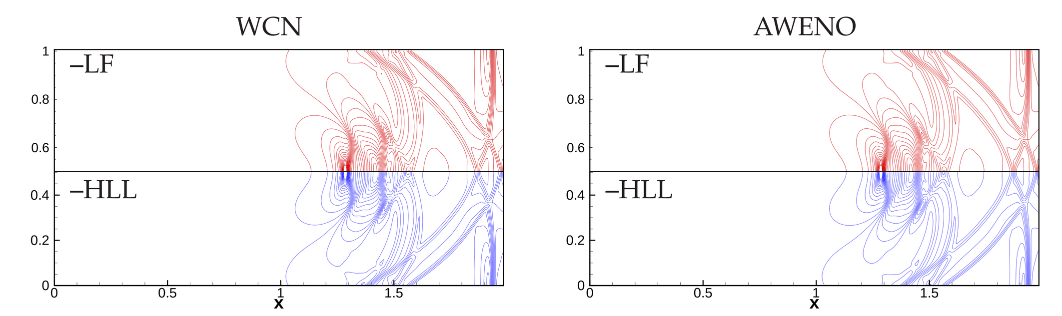

**Second**, a small perturbation of a steady-state water.




**Third**, the dam-breaking problem over a rectangular bump ($t = 15$).




**Fourth**, a small perturbation of 2D steady-state water ($t = 0.6$).




## Acknowledgement

Gratefully acknowledge to

- Prof. Wai Sun Don, Prof. Zhen Gao at OUC and Dr. Peng Li at STDU.
- NSFC (11871443), SDNSF (ZR2017MA016), OUC (201712011).

## References

[1] X. Deng, H. Zhang, J. Comput. Phys. **165**, 22–44 (2000).
[2] X. Deng, M, Mao, *et al.*, AIAA 2011-3847.
[3] Y. Jiang, C.-W. Shu, *et al.*, SIAM J. Sci. Comput. **35**, A1137–A1160 (2013).
[4] B.-S. Wang, P. Li, *et al.*, J. Comput. Phys. **374**, 469–477 (2018).
[5] Y. Xing, C.-W. Shu, J. Comput. Phys. **208**, 206–227 (2005).
[6] Y. Xing, C.-W. Shu, J. Comput. Phys. **214**, 567–598 (2006).