# SUPPLEMENTARY MATERIALS: ACTIVE FLUX METHODS FOR HYPERBOLIC CONSERVATION LAWS—FLUX VECTOR SPLITTING AND BOUND-PRESERVATION*

JUNMING DUAN†, WASILIJ BARSUKOW‡, AND CHRISTIAN KLINGENBERG§

# A 2D flux vector splitting

## A.1 Local Lax-Friedrichs flux vector splitting

This flux vector splitting can be written as

$$\boldsymbol{F}_\ell^\pm = \frac{1}{2}(\boldsymbol{F}_\ell(\boldsymbol{U}) \pm \alpha_\ell \boldsymbol{U}),$$

where $\alpha_\ell$ is determined by

$$(\alpha_1)_{i+\frac{1}{2},q} = \max_s \left\{|\varrho_1(\boldsymbol{U}_{s,q})|\right\},\ s \in \left\{i-\frac{1}{2}, i, i+\frac{1}{2}, i+1, i+\frac{3}{2}\right\}, q = j, j+\frac{1}{2},$$
$$(\alpha_2)_{q,j+\frac{1}{2}} = \max_s \left\{|\varrho_2(\boldsymbol{U}_{q,s})|\right\},\ s \in \left\{j-\frac{1}{2}, j, j+\frac{1}{2}, j+1, j+\frac{3}{2}\right\}, q = i, i+\frac{1}{2},$$

and $\varrho_\ell$ is the spectral radius of the Jacobian matrix $\partial \boldsymbol{F}_\ell / \partial \boldsymbol{U}$.

## A.2 Upwind flux vector splitting

The flux can also be split based on each characteristic field as follows

$$\boldsymbol{F}_\ell^\pm = \frac{1}{2}(\boldsymbol{F}_\ell(\boldsymbol{U}) \pm |\boldsymbol{J}_\ell|\boldsymbol{U}), \quad |\boldsymbol{J}_\ell| = \boldsymbol{R}_\ell(\boldsymbol{\Lambda}_\ell^+ - \boldsymbol{\Lambda}_\ell^-)\boldsymbol{R}_\ell^{-1}, \tag{45}$$

with $\boldsymbol{J}_\ell = \partial \boldsymbol{F}_\ell / \partial \boldsymbol{U} = \boldsymbol{R}_\ell \boldsymbol{\Lambda}_\ell \boldsymbol{R}_\ell^{-1}$ the eigen-decomposition of the Jacobian matrix.

For the Euler equations, the explicit expressions in the $x$-direction are

$$\boldsymbol{F}_1^\pm = \begin{bmatrix} \frac{\rho}{2\gamma}\alpha^\pm \\ \frac{\rho}{2\gamma}\left(\alpha^\pm v_1 + a(\lambda_2^\pm - \lambda_3^\pm)\right) \\ \frac{\rho}{2\gamma}\alpha^\pm v_2 \\ \frac{\rho}{2\gamma}\left(\frac{1}{2}\alpha^\pm \|\boldsymbol{v}\|_2^2 + a v_1(\lambda_2^\pm - \lambda_3^\pm) + \frac{a^2}{\gamma-1}(\lambda_2^\pm + \lambda_3^\pm)\right) \end{bmatrix},$$

where $\lambda_1 = v_\ell,\ \lambda_2 = v_\ell + a,\ \lambda_3 = v_\ell - a,\ \alpha^\pm = 2(\gamma-1)\lambda_1^\pm + \lambda_2^\pm + \lambda_3^\pm$, and $a = \sqrt{\gamma p/\rho}$ is the sound speed. The expressions in the $y$-direction can be obtained using the rotational invariance.

## A.3 Van Leer-Hänel flux vector splitting for the Euler equations

For the $x$-direction, the flux is split according to the Mach number $M = v_1/a$ as

$$\boldsymbol{F}_1 = \begin{bmatrix} \rho a M \\ \rho a^2(M^2 + \frac{1}{\gamma}) \\ \rho a M v_2 \\ \rho a^3 M(\frac{1}{2}M^2 + \frac{1}{\gamma-1}) + \frac{\rho a M v_2^2}{2} \end{bmatrix} = \boldsymbol{F}_1^+ + \boldsymbol{F}_1^-,\ \boldsymbol{F}_1^\pm = \begin{bmatrix} \pm\frac{1}{4}\rho a(M \pm 1)^2 \\ \pm\frac{1}{4}\rho a(M \pm 1)^2 v_1 + p^\pm \\ \pm\frac{1}{4}\rho a(M \pm 1)^2 v_2 \\ \pm\frac{1}{4}\rho a(M \pm 1)^2 H \end{bmatrix}$$

with the enthalpy $H = (E+p)/\rho$, and the pressure-splitting $p^\pm = \frac{1}{2}(1 \pm \gamma M)p$.

# B Bound-preserving property of intermediate states

Similar to the proofs in [48, 49], the following lemmas hold.

**Lemma B.1.** *For the scalar conservation laws (17), the intermediate state* $\widetilde{u} = \frac{1}{2}(u_L + u_R) + \frac{1}{2\alpha}(f_\ell(u_L) - f_\ell(u_R))$ *stays in* $\mathcal{G}$ *(25) if* $\alpha \geqslant \max\{\varrho_\ell(u_L), \varrho_\ell(u_R)\}$.

*Proof.* The partial derivatives of the intermediate state satisfy

$$\frac{\partial \widetilde{u}(u_L, u_R)}{\partial u_L} = \frac{1}{2}\left(1 + \frac{f_\ell'(u_L)}{\alpha}\right) \geqslant 0, \quad \frac{\partial \widetilde{u}(u_L, u_R)}{\partial u_R} = \frac{1}{2}\left(1 - \frac{f_\ell'(u_R)}{\alpha}\right) \geqslant 0.$$

As $\widetilde{u}(m_0, m_0) = m_0$, $\widetilde{u}(M_0, M_0) = M_0$, it holds $m_0 \leqslant \widetilde{u} \leqslant M_0$. □

**Lemma B.2.** *For the Euler equations, the intermediate state* $\widetilde{\boldsymbol{U}} = \frac{1}{2}(\boldsymbol{U}_L + \boldsymbol{U}_R) + \frac{1}{2\alpha}(\boldsymbol{F}_\ell(\boldsymbol{U}_L) - \boldsymbol{F}_\ell(\boldsymbol{U}_R))$ *stays in* $\mathcal{G}$ *(26) if* $\alpha \geqslant \max\{\varrho_\ell(\boldsymbol{U}_L), \varrho_\ell(\boldsymbol{U}_R)\}$.

*Proof.* For the Euler equations, as the intermediate state is a convex combination of $\boldsymbol{U}_L - \frac{1}{\alpha}\boldsymbol{F}_\ell(\boldsymbol{U}_L)$ and $\boldsymbol{U}_R + \frac{1}{\alpha}\boldsymbol{F}_\ell(\boldsymbol{U}_R)$, we only need to show that the $\boldsymbol{U} \pm \frac{1}{\alpha}\boldsymbol{F}_\ell(\boldsymbol{U})$ belongs to $\mathcal{G}$. The density component $(\rho \pm (\rho v_\ell)/\alpha)$ is positive since $\alpha > |v_\ell|$. The recovered internal energy is

$$\begin{aligned}
\rho e\left(\boldsymbol{U} \pm \frac{1}{\alpha}\boldsymbol{F}_\ell(\boldsymbol{U})\right) &= E\left(\boldsymbol{U} \pm \frac{1}{\alpha}\boldsymbol{F}_\ell(\boldsymbol{U})\right) - \frac{\left\|\rho \boldsymbol{v}\left(\boldsymbol{U} \pm \frac{1}{\alpha}\boldsymbol{F}_\ell(\boldsymbol{U})\right)\right\|_2^2}{2\rho\left(\boldsymbol{U} \pm \frac{1}{\alpha}\boldsymbol{F}_\ell(\boldsymbol{U})\right)} \\
&= \left(1 - \frac{p^2}{2(\alpha \pm v_\ell)^2 \rho^2 e}\right)\left(1 \pm \frac{v_\ell}{\alpha}\right)\rho e,
\end{aligned}$$

so that one has $\rho e\left(\boldsymbol{U} \pm \frac{1}{\alpha}\boldsymbol{F}_\ell(\boldsymbol{U})\right) > 0 \iff \frac{p^2}{2\rho^2 e} < (\alpha \pm v_\ell)^2 \iff \frac{\gamma-1}{2\gamma}a^2 < (\alpha \pm v_\ell)^2$ for the perfect gas EOS, which holds as $\alpha \geqslant |v_\ell| + a$. □

# C 1D bound-preserving active flux methods

For the scalar conservation law (2), its solutions satisfy a strict maximum principle (MP) [14], i.e.,

$$\mathcal{G} = \{u \mid m_0 \leqslant u \leqslant M_0\}, \quad m_0 = \min_x u_0(x), \ M_0 = \max_x u_0(x). \tag{46}$$

For the compressible Euler equations, the admissible state set is

$$\mathcal{G} = \left\{\boldsymbol{U} = (\rho, \rho v, E) \;\middle|\; \rho > 0, \ p = (\gamma - 1)\left(E - (\rho v)^2/(2\rho)\right) > 0\right\}. \tag{47}$$

which is convex, see e.g. [51].

## C.1 Convex limiting for the cell average

This section presents a convex limiting approach to achieve the BP property of the cell average update. The low-order scheme is chosen as the first-order LLF scheme

$$\begin{aligned}
&\overline{\boldsymbol{U}}_i^{\mathrm{L}} = \overline{\boldsymbol{U}}_i^n - \mu_i\left(\widehat{\boldsymbol{F}}_{i+\frac{1}{2}}^{\mathrm{L}} - \widehat{\boldsymbol{F}}_{i-\frac{1}{2}}^{\mathrm{L}}\right), \quad \mu_i = \Delta t^n/\Delta x_i, \\
&\widehat{\boldsymbol{F}}_{i+\frac{1}{2}}^{\mathrm{L}} = \boldsymbol{F}^{\mathrm{LLF}}(\overline{\boldsymbol{U}}_i^n, \overline{\boldsymbol{U}}_{i+1}^n) = \frac{1}{2}\left(\boldsymbol{F}(\overline{\boldsymbol{U}}_i^n) + \boldsymbol{F}(\overline{\boldsymbol{U}}_{i+1}^n)\right) - \frac{\alpha_{i+\frac{1}{2}}}{2}\left(\overline{\boldsymbol{U}}_{i+1}^n - \overline{\boldsymbol{U}}_i^n\right), \\
&\alpha_{i+\frac{1}{2}} = \max\{\varrho(\overline{\boldsymbol{U}}_i^n), \ \varrho_\ell(\overline{\boldsymbol{U}}_{i+1}^n)\},
\end{aligned}$$

where $\varrho$ is the spectral radius of $\partial \boldsymbol{F}/\partial \boldsymbol{U}$. Note that here $\alpha_{i+\frac{1}{2}}$ is not the same as the one in the LLF FVS (12). Following [22], the first-order LLF scheme can be rewritten as

$$\overline{\boldsymbol{U}}_i^{\mathtt{L}} = \left[1 - \mu_i \left(\alpha_{i-\frac{1}{2}} + \alpha_{i+\frac{1}{2}}\right)\right] \overline{\boldsymbol{U}}_i^n + \mu_i \alpha_{i-\frac{1}{2}} \widetilde{\boldsymbol{U}}_{i-\frac{1}{2}} + \mu_i \alpha_{i+\frac{1}{2}} \widetilde{\boldsymbol{U}}_{i+\frac{1}{2}}, \tag{48}$$

with the first-order LLF intermediate states defined as

$$\widetilde{\boldsymbol{U}}_{i\pm\frac{1}{2}} := \frac{1}{2}\left(\overline{\boldsymbol{U}}_i^n + \overline{\boldsymbol{U}}_{i\pm1}^n\right) \pm \frac{1}{2\alpha_{i\pm\frac{1}{2}}}\left[\boldsymbol{F}(\overline{\boldsymbol{U}}_i^n) - \boldsymbol{F}(\overline{\boldsymbol{U}}_{i\pm1}^n)\right]. \tag{49}$$

The proofs of $\widetilde{\boldsymbol{U}}_{i\pm\frac{1}{2}} \in \mathcal{G}$ are similar to Appendix B, for the scalar case and Euler equations.

**Lemma C.1.** *If the time step size $\Delta t^n$ satisfies*

$$\Delta t^n \leqslant \frac{\Delta x_i}{\alpha_{i-\frac{1}{2}} + \alpha_{i+\frac{1}{2}}}, \tag{50}$$

*then (48) is a convex combination, and the first-order LLF scheme is BP.*

The proof (see e.g. [22, 37]) relies on $\overline{\boldsymbol{U}}_i^n, \widetilde{\boldsymbol{U}}_{i\pm\frac{1}{2}} \in \mathcal{G}$ and the convexity of $\mathcal{G}$.

Upon defining the anti-diffusive flux $\Delta\widehat{\boldsymbol{F}}_{i\pm\frac{1}{2}} := \widehat{\boldsymbol{F}}_{i\pm\frac{1}{2}}^{\mathtt{H}} - \widehat{\boldsymbol{F}}_{i\pm\frac{1}{2}}^{\mathtt{L}}$ with $\widehat{\boldsymbol{F}}_{i\pm\frac{1}{2}}^{\mathtt{H}} := \boldsymbol{F}(\boldsymbol{U}_{i\pm\frac{1}{2}})$, a forward-Euler step applied to the semi-discrete high-order scheme for the cell average (4) can be written as

$$\begin{aligned}
\overline{\boldsymbol{U}}_i^{\mathtt{H}} &= \overline{\boldsymbol{U}}_i^n - \mu_i(\widehat{\boldsymbol{F}}_{i+\frac{1}{2}}^{\mathtt{H}} - \widehat{\boldsymbol{F}}_{i-\frac{1}{2}}^{\mathtt{H}}) = \overline{\boldsymbol{U}}_i^n - \mu_i(\widehat{\boldsymbol{F}}_{i+\frac{1}{2}}^{\mathtt{L}} - \widehat{\boldsymbol{F}}_{i-\frac{1}{2}}^{\mathtt{L}}) - \mu_i(\Delta\widehat{\boldsymbol{F}}_{i+\frac{1}{2}} - \Delta\widehat{\boldsymbol{F}}_{i-\frac{1}{2}}) \\
&= \left[1 - \mu_i \left(\alpha_{i-\frac{1}{2}} + \alpha_{i+\frac{1}{2}}\right)\right] \overline{\boldsymbol{U}}_i^n + \mu_i \alpha_{i-\frac{1}{2}} \widetilde{\boldsymbol{U}}_{i-\frac{1}{2}}^{\mathtt{H},+} + \mu_i \alpha_{i+\frac{1}{2}} \widetilde{\boldsymbol{U}}_{i+\frac{1}{2}}^{\mathtt{H},-}, \qquad (51) \\
\widetilde{\boldsymbol{U}}_{i-\frac{1}{2}}^{\mathtt{H},+} &:= \left(\widetilde{\boldsymbol{U}}_{i-\frac{1}{2}} + \frac{\Delta\widehat{\boldsymbol{F}}_{i-\frac{1}{2}}}{\alpha_{i-\frac{1}{2}}}\right), \quad \widetilde{\boldsymbol{U}}_{i+\frac{1}{2}}^{\mathtt{H},-} := \left(\widetilde{\boldsymbol{U}}_{i+\frac{1}{2}} - \frac{\Delta\widehat{\boldsymbol{F}}_{i+\frac{1}{2}}}{\alpha_{i+\frac{1}{2}}}\right).
\end{aligned}$$

With the low-order scheme (48) and high-order scheme (51) having the same abstract form, one can blend them to define the limited scheme for the cell average as

$$\overline{\boldsymbol{U}}_i^{\mathtt{Lim}} = \left[1 - \mu_i \left(\alpha_{i-\frac{1}{2}} + \alpha_{i+\frac{1}{2}}\right)\right] \overline{\boldsymbol{U}}_i^n + \mu_i \alpha_{i-\frac{1}{2}} \widetilde{\boldsymbol{U}}_{i-\frac{1}{2}}^{\mathtt{Lim},+} + \mu_i \alpha_{i+\frac{1}{2}} \widetilde{\boldsymbol{U}}_{i+\frac{1}{2}}^{\mathtt{Lim},-}, \tag{52}$$

where the limited intermediate states are

$$\widetilde{\boldsymbol{U}}_{i\pm\frac{1}{2}}^{\mathtt{Lim},\mp} = \widetilde{\boldsymbol{U}}_{i\pm\frac{1}{2}} \mp \frac{\Delta\widehat{\boldsymbol{F}}_{i\pm\frac{1}{2}}^{\mathtt{Lim}}}{\alpha_{i\pm\frac{1}{2}}} := \widetilde{\boldsymbol{U}}_{i\pm\frac{1}{2}} \mp \frac{\theta_{i\pm\frac{1}{2}}\Delta\widehat{\boldsymbol{F}}_{i\pm\frac{1}{2}}}{\alpha_{i\pm\frac{1}{2}}}, \tag{53}$$

and $\theta_{i\pm\frac{1}{2}} \in [0,1]$ are the blending coefficients. The limited scheme (52) reduces to the first-order LLF scheme if $\theta_{i\pm\frac{1}{2}} = 0$, and recovers the high-order AF scheme (4) when $\theta_{i\pm\frac{1}{2}} = 1$.

### C.1.1 Application to scalar conservation laws

Similar to the 2D case, the convex limiting is applied to scalar conservation laws (2), such that the limited cell averages (52) satisfy the MP $u_i^{\min} \leqslant \bar{u}_i^{\mathtt{Lim}} \leqslant u_i^{\max}$, where $u_i^{\min} = \min\mathcal{N}$, $u_i^{\max} = \max\mathcal{N}$, and $\mathcal{N}$ will be defined later. The limited anti-diffusive flux is

$$\Delta\hat{f}_{i+\frac{1}{2}}^{\mathtt{Lim}} = \begin{cases} \min\left\{\Delta\hat{f}_{i+\frac{1}{2}},\ \alpha_{i+\frac{1}{2}}(\tilde{u}_{i+\frac{1}{2}} - u_i^{\min}),\ \alpha_{i+\frac{1}{2}}(u_{i+1}^{\max} - \tilde{u}_{i+\frac{1}{2}})\right\}, & \text{if } \Delta\hat{f}_{i+\frac{1}{2}} \geqslant 0, \\ \max\left\{\Delta\hat{f}_{i+\frac{1}{2}},\ \alpha_{i+\frac{1}{2}}(u_{i+1}^{\min} - \tilde{u}_{i+\frac{1}{2}}),\ \alpha_{i+\frac{1}{2}}(\tilde{u}_{i+\frac{1}{2}} - u_i^{\max})\right\}, & \text{otherwise.} \end{cases}$$

Finally, the limited numerical flux is

$$\hat{f}^{\mathtt{Lim}}_{i+\frac{1}{2}} = \hat{f}^{\mathtt{L}}_{i+\frac{1}{2}} + \Delta\hat{f}^{\mathtt{Lim}}_{i+\frac{1}{2}}. \tag{54}$$

If considering the global MP, $\mathcal{N} = \bigcup_i \{\bar{u}^n_i, u^n_{i+\frac{1}{2}}\}$. For the local MP, one can choose $\mathcal{N} = \min\left\{\bar{u}^n_i,\ \tilde{u}_{i-\frac{1}{2}},\ \tilde{u}_{i+\frac{1}{2}}, \bar{u}^n_{i-1}, \bar{u}^n_{i+1}\right\}$, which consists of the neighboring cell averages and intermediate states.

### C.1.2 Application to the compressible Euler equations

This section aims at enforcing the positivity of density and pressure. To avoid the effect of the round-off error, we need to choose the desired lower bounds. Denote the lowest density and pressure in the domain by

$$\varepsilon^\rho := \min_i\{\overline{\boldsymbol{U}}^{n,\rho}_i, \boldsymbol{U}^{n,\rho}_{i+\frac{1}{2}}\},\ \varepsilon^p := \min_i\{p(\overline{\boldsymbol{U}}^n_i), p(\boldsymbol{U}^n_{i+\frac{1}{2}})\}, \tag{55}$$

where $\boldsymbol{U}^{*,\rho}$ and $p(\boldsymbol{U}^*)$ denote the density component and pressure recovered from $\boldsymbol{U}^*$, respectively. The limiting (53) is feasible if the constraints are satisfied by the first-order LLF intermediate states (49), thus the lower bounds can be defined as

$$\varepsilon^\rho_i := \min\{10^{-13}, \varepsilon^\rho, \widetilde{\boldsymbol{U}}^\rho_{i-\frac{1}{2}}, \widetilde{\boldsymbol{U}}^\rho_{i+\frac{1}{2}}\},\ \varepsilon^p_i := \min\{10^{-13}, \varepsilon^p, p(\widetilde{\boldsymbol{U}}_{i-\frac{1}{2}}), p(\widetilde{\boldsymbol{U}}_{i+\frac{1}{2}})\}.$$

i) **Positivity of density.** The first step is to impose the density positivity $\widetilde{\boldsymbol{U}}^{\mathtt{Lim},\pm,\rho}_{i+\frac{1}{2}} \geqslant \bar{\varepsilon}^\rho_{i+\frac{1}{2}} := \min\{\varepsilon^\rho_i, \varepsilon^\rho_{i+1}\}$. Similarly to the derivation of the scalar case, the corresponding density component of the limited anti-diffusive flux is

$$\Delta\widehat{\boldsymbol{F}}^{\mathtt{Lim},*,\rho}_{i+\frac{1}{2}} = \begin{cases} \min\left\{\Delta\widehat{\boldsymbol{F}}^\rho_{i+\frac{1}{2}},\ \alpha_{i+\frac{1}{2}}\left(\widetilde{\boldsymbol{U}}^\rho_{i+\frac{1}{2}} - \bar{\varepsilon}^\rho_{i+\frac{1}{2}}\right)\right\}, & \text{if } \Delta\widehat{\boldsymbol{F}}^\rho_{i+\frac{1}{2}} \geqslant 0, \\ \max\left\{\Delta\widehat{\boldsymbol{F}}^\rho_{i+\frac{1}{2}},\ \alpha_{i+\frac{1}{2}}\left(\bar{\varepsilon}^\rho_{i+\frac{1}{2}} - \widetilde{\boldsymbol{U}}^\rho_{i+\frac{1}{2}}\right)\right\}, & \text{otherwise.} \end{cases}$$

Then the density component of the limited flux is $\widehat{\boldsymbol{F}}^{\mathtt{Lim},*,\rho}_{i+\frac{1}{2}} = \widehat{\boldsymbol{F}}^{\mathtt{L},\rho}_{i+\frac{1}{2}} + \Delta\widehat{\boldsymbol{F}}^{\mathtt{Lim},*,\rho}_{i+\frac{1}{2}}$, with the other components remaining the same as $\widehat{\boldsymbol{F}}^{\mathtt{H}}_{i+\frac{1}{2}}$.

ii) **Positivity of pressure.** The second step is to enforce pressure positivity $p(\widetilde{\boldsymbol{U}}^{\mathtt{Lim},\pm}_{i+\frac{1}{2}}) \geqslant \bar{\varepsilon}^p_{i+\frac{1}{2}} := \min\{\varepsilon^p_i, \varepsilon^p_{i+1}\}$. Since

$$\widetilde{\boldsymbol{U}}^{\mathtt{Lim},\pm}_{i+\frac{1}{2}} = \widetilde{\boldsymbol{U}}_{i+\frac{1}{2}} \pm \frac{\theta_{i+\frac{1}{2}}\Delta\widehat{\boldsymbol{F}}^{\mathtt{Lim},*}_{i+\frac{1}{2}}}{\alpha_{i+\frac{1}{2}}}, \quad \Delta\widehat{\boldsymbol{F}}^{\mathtt{Lim},*}_{i+\frac{1}{2}} = \widehat{\boldsymbol{F}}^{\mathtt{Lim},*}_{i+\frac{1}{2}} - \widehat{\boldsymbol{F}}^{\mathtt{L}}_{i+\frac{1}{2}},$$

the constraints lead to two inequalities

$$A_{i+\frac{1}{2}}\theta^2_{i+\frac{1}{2}} \pm B_{i+\frac{1}{2}}\theta_{i+\frac{1}{2}} \leqslant C_{i+\frac{1}{2}}, \tag{56}$$

with the coefficients

$$\begin{aligned} A_{i+\frac{1}{2}} &= \frac{1}{2}\left(\Delta\widehat{\boldsymbol{F}}^{\mathtt{Lim},*,\rho v}_{i+\frac{1}{2}}\right)^2 - \Delta\widehat{\boldsymbol{F}}^{\mathtt{Lim},*,\rho}_{i+\frac{1}{2}}\Delta\widehat{\boldsymbol{F}}^{\mathtt{Lim},*,E}_{i+\frac{1}{2}}, \\ B_{i+\frac{1}{2}} &= \alpha_{i+\frac{1}{2}}\left(\Delta\widehat{\boldsymbol{F}}^{\mathtt{Lim},*,\rho}_{i+\frac{1}{2}}\widetilde{\boldsymbol{U}}^E_{i+\frac{1}{2}} + \widetilde{\boldsymbol{U}}^\rho_{i+\frac{1}{2}}\Delta\widehat{\boldsymbol{F}}^{\mathtt{Lim},*,E}_{i+\frac{1}{2}} - \Delta\widehat{\boldsymbol{F}}^{\mathtt{Lim},*,\rho v}_{i+\frac{1}{2}}\widetilde{\boldsymbol{U}}^{\rho v}_{i+\frac{1}{2}} - \tilde{\varepsilon}\Delta\widehat{\boldsymbol{F}}^{\mathtt{Lim},*,\rho}_{i+\frac{1}{2}}\right), \\ C_{i+\frac{1}{2}} &= \alpha^2_{i+\frac{1}{2}}\left(\widetilde{\boldsymbol{U}}^\rho_{i+\frac{1}{2}}\widetilde{\boldsymbol{U}}^E_{i+\frac{1}{2}} - \frac{1}{2}\left(\widetilde{\boldsymbol{U}}^{\rho v}_{i+\frac{1}{2}}\right)^2 - \tilde{\varepsilon}\widetilde{\boldsymbol{U}}^\rho_{i+\frac{1}{2}}\right), \quad \tilde{\varepsilon} = \bar{\varepsilon}^p_{i+\frac{1}{2}}/(\gamma-1). \end{aligned}$$

Following [31], the inequalities (56) hold under the linear sufficient condition

$$\left(\max\left\{0, A_{i+\frac{1}{2}}\right\} + \left|B_{i+\frac{1}{2}}\right|\right)\theta_{i+\frac{1}{2}} \leqslant C_{i+\frac{1}{2}},$$

if making use of $\theta^2_{i+\frac{1}{2}} \leqslant \theta_{i+\frac{1}{2}},\ \theta_{i+\frac{1}{2}} \in [0,1]$. Thus the coefficient can be chosen as

$$\theta_{i+\frac{1}{2}} = \min\left\{1,\ \frac{C_{i+\frac{1}{2}}}{\max\{0, A_{i+\frac{1}{2}}\} + |B_{i+\frac{1}{2}}|}\right\},$$

and the final limited numerical flux is

$$\widehat{\boldsymbol{F}}^{\mathtt{Lim},**}_{i+\frac{1}{2}} = \widehat{\boldsymbol{F}}^{\mathtt{L}}_{i+\frac{1}{2}} + \theta_{i+\frac{1}{2}}\Delta\widehat{\boldsymbol{F}}^{\mathtt{Lim},*}_{i+\frac{1}{2}}. \tag{57}$$

### C.1.3 Shock sensor-based limiting

In 1D, the Jameson's shock sensor [29] is

$$(\varphi_1)_i = \frac{|\bar{p}_{i+1} - 2\bar{p}_i + \bar{p}_{i-1}|}{|\bar{p}_{i+1} + 2\bar{p}_i + \bar{p}_{i-1}|},$$

and the modified Ducros' shock sensor reduced from the 2D case [15] is

$$(\varphi_2)_i = \max\left\{-\frac{\bar{v}_{i+1} - \bar{v}_{i-1}}{|\bar{v}_{i+1} - \bar{v}_{i-1}| + 10^{-40}},\ 0\right\}.$$

Note that $\bar{v}_i$ and $\bar{p}_i$ are the velocity and pressure recovered from the cell average $\overline{\boldsymbol{U}}_i$. The blending coefficient is designed as

$$\begin{aligned}
&\theta^s_{i+\frac{1}{2}} = \exp(-\kappa(\varphi_1)_{i+\frac{1}{2}}(\varphi_2)_{i+\frac{1}{2}}) \in (0,1],\\
&(\varphi_s)_{i+\frac{1}{2}} = \max\left\{(\varphi_s)_i, (\varphi_s)_{i+1}\right\},\ s = 1,2,
\end{aligned}$$

where the problem-dependent parameter $\kappa$ adjusts the strength of the limiting, and its optimal choice needs further investigation. The final limited numerical flux is

$$\widehat{\boldsymbol{F}}^{\mathtt{Lim}}_{i+\frac{1}{2}} = \widehat{\boldsymbol{F}}^{\mathtt{L}}_{i+\frac{1}{2}} + \theta^s_{i+\frac{1}{2}}\Delta\widehat{\boldsymbol{F}}^{\mathtt{Lim},**}_{i+\frac{1}{2}}, \tag{58}$$

with $\Delta\widehat{\boldsymbol{F}}^{\mathtt{Lim},**}_{i+\frac{1}{2}} = \widehat{\boldsymbol{F}}^{\mathtt{Lim},**}_{i+\frac{1}{2}} - \widehat{\boldsymbol{F}}^{\mathtt{L}}_{i+\frac{1}{2}}$, and $\widehat{\boldsymbol{F}}^{\mathtt{Lim},**}_{i+\frac{1}{2}}$ given in (57).

## C.2 Scaling limiter for point value

A first-order LLF scheme for the point value update can be written as

$$\boldsymbol{U}^{\mathtt{L}}_{i+\frac{1}{2}} = \boldsymbol{U}^n_{i+\frac{1}{2}} - \frac{2\Delta t^n}{\Delta x_i + \Delta x_{i+1}}\left(\widehat{\boldsymbol{F}}^{\mathtt{L}}_{i+1}(\boldsymbol{U}^n_{i+\frac{1}{2}}, \boldsymbol{U}^n_{i+\frac{3}{2}}) - \widehat{\boldsymbol{F}}^{\mathtt{L}}_{i}(\boldsymbol{U}^n_{i-\frac{1}{2}}, \boldsymbol{U}^n_{i+\frac{1}{2}})\right), \tag{59}$$

with the numerical flux

$$\begin{aligned}
&\widehat{\boldsymbol{F}}^{\mathtt{L}}_i = \widehat{\boldsymbol{F}}^{\mathtt{LLF}}(\boldsymbol{U}^n_{i-\frac{1}{2}}, \boldsymbol{U}^n_{i+\frac{1}{2}}) = \frac{1}{2}\left(\boldsymbol{F}(\boldsymbol{U}^n_{i-\frac{1}{2}}) + \boldsymbol{F}(\boldsymbol{U}^n_{i+\frac{1}{2}})\right) - \frac{\alpha_i}{2}\left(\boldsymbol{U}^n_{i+\frac{1}{2}} - \boldsymbol{U}^n_{i-\frac{1}{2}}\right),\\
&\alpha_i = \max\{\varrho(\boldsymbol{U}^n_{i-\frac{1}{2}}),\ \varrho(\boldsymbol{U}^n_{i+\frac{1}{2}})\}.
\end{aligned}$$

Similarly to Lemma C.1, it is straightforward to obtain the following Lemma.

**Lemma C.2.** *The LLF scheme for the point value (59) is BP under the CFL condition*

$$\Delta t^n \leqslant \frac{\Delta x_i + \Delta x_{i+1}}{2(\alpha_i + \alpha_{i+1})}. \tag{60}$$

The limited solution is obtained by blending the high-order AF scheme (5) with the forward-Euler scheme and the LLF scheme (59) as $\boldsymbol{U}^{\mathtt{Lim}}_{i+\frac{1}{2}} = \theta_{i+\frac{1}{2}}\boldsymbol{U}^{\mathtt{H}}_{i+\frac{1}{2}} + (1-\theta_{i+\frac{1}{2}})\boldsymbol{U}^{\mathtt{L}}_{i+\frac{1}{2}}$, such that $\boldsymbol{U}^{\mathtt{Lim}}_{i+\frac{1}{2}} \in \mathcal{G}$.

### C.2.1 Application to scalar conservation laws

This section enforces the MP $u^{\min}_{i+\frac{1}{2}} \leqslant u^{\mathtt{Lim}}_{i+\frac{1}{2}} \leqslant u^{\max}_{i+\frac{1}{2}}$ using the scaling limiter [48]. The limited solution is

$$u^{\mathtt{Lim}}_{i+\frac{1}{2}} = \theta_{i+\frac{1}{2}} u^{\mathtt{H}}_{i+\frac{1}{2}} + \left(1-\theta_{i+\frac{1}{2}}\right) u^{\mathtt{L}}_{i+\frac{1}{2}}, \tag{61}$$

with the coefficient

$$\theta_{i+\frac{1}{2}} = \min\left\{1,\ \left|\frac{u^{\mathtt{L}}_{i+\frac{1}{2}} - u^{\min}_{i+\frac{1}{2}}}{u^{\mathtt{L}}_{i+\frac{1}{2}} - u^{\mathtt{H}}_{i+\frac{1}{2}}}\right|,\ \left|\frac{u^{\max}_{i+\frac{1}{2}} - u^{\mathtt{L}}_{i+\frac{1}{2}}}{u^{\mathtt{H}}_{i+\frac{1}{2}} - u^{\mathtt{L}}_{i+\frac{1}{2}}}\right|\right\}.$$

The bounds are determined by $u^{\min}_{i+\frac{1}{2}} = \min\mathcal{N}$, $u^{\max}_{i+\frac{1}{2}} = \max\mathcal{N}$, where the set $\mathcal{N}$ consists of all the DoFs in the domain, i.e., $\mathcal{N} = \bigcup_i \{\bar{u}^n_i, u^n_{i+\frac{1}{2}}\}$ for the global MP. One can also consider the local MP, e.g., $\mathcal{N} = \left\{u^n_{i-\frac{1}{2}}, u^n_{i+\frac{1}{2}}, u^n_{i+\frac{3}{2}}\right\}$, which at least includes all the DoFs appeared in the first-order LLF scheme (59).

### C.2.2 Application to the compressible Euler equations

The limiting consists of two steps.

**i) Positivity of density.** First, the high-order solution $\boldsymbol{U}^{\mathtt{H}}_{i+\frac{1}{2}}$ is modified as $\boldsymbol{U}^{\mathtt{Lim},*}_{i+\frac{1}{2}}$, such that $\boldsymbol{U}^{\mathtt{Lim},*,\rho}_{i+\frac{1}{2}} \geqslant \varepsilon^{\rho}_{i+\frac{1}{2}} := \min\{10^{-13}, \varepsilon^{\rho}, \boldsymbol{U}^{\mathtt{L},\rho}_{i+\frac{1}{2}}\}$ with $\varepsilon^{\rho}$ given in (55). Solving the inequality yields

$$\theta^*_{i+\frac{1}{2}} = \begin{cases} \dfrac{\boldsymbol{U}^{\mathtt{L},\rho}_{i+\frac{1}{2}} - \varepsilon^{\rho}_{i+\frac{1}{2}}}{\boldsymbol{U}^{\mathtt{L},\rho}_{i+\frac{1}{2}} - \boldsymbol{U}^{\mathtt{H},\rho}_{i+\frac{1}{2}}}, & \text{if } \boldsymbol{U}^{\mathtt{H},\rho}_{i+\frac{1}{2}} < \varepsilon^{\rho}_{i+\frac{1}{2}}, \\ 1, & \text{otherwise.} \end{cases}$$

Then the density component of the limited solution is $\boldsymbol{U}^{\mathtt{Lim},*,\rho}_{i+\frac{1}{2}} = \theta^*_{i+\frac{1}{2}}\boldsymbol{U}^{\mathtt{H},\rho}_{i+\frac{1}{2}} + (1-\theta^*_{i+\frac{1}{2}})\boldsymbol{U}^{\mathtt{L},\rho}_{i+\frac{1}{2}}$, with the other components remaining the same as $\boldsymbol{U}^{\mathtt{H}}_{i+\frac{1}{2}}$.

**ii) Positivity of pressure.** Then the limited solution $\boldsymbol{U}^{\mathtt{Lim},*}_{i+\frac{1}{2}}$ is modified as $\boldsymbol{U}^{\mathtt{Lim}}_{i+\frac{1}{2}}$, such that it gives positive pressure, i.e., $p(\boldsymbol{U}^{\mathtt{Lim}}_{i+\frac{1}{2}}) \geqslant \varepsilon^{p}_{i+\frac{1}{2}} := \min\{10^{-13}, \varepsilon^{p}, p(\boldsymbol{U}^{\mathtt{L}}_{i+\frac{1}{2}})\}$, with $\varepsilon^{p}$ given in (55). Let the final limited solution be

$$\boldsymbol{U}^{\mathtt{Lim}}_{i+\frac{1}{2}} = \theta^{**}_{i+\frac{1}{2}}\boldsymbol{U}^{\mathtt{Lim},*}_{i+\frac{1}{2}} + \left(1-\theta^{**}_{i+\frac{1}{2}}\right)\boldsymbol{U}^{\mathtt{L}}_{i+\frac{1}{2}}. \tag{62}$$

The pressure is a concave function of the conservative variables (see e.g. [50]), so that $p(\boldsymbol{U}^{\mathtt{Lim}}_{i+\frac{1}{2}}) \geqslant \theta^{**}_{i+\frac{1}{2}} p(\boldsymbol{U}^{\mathtt{Lim},*}_{i+\frac{1}{2}}) + \left(1-\theta^{**}_{i+\frac{1}{2}}\right) p(\boldsymbol{U}^{\mathtt{L}}_{i+\frac{1}{2}})$ based on Jensen's inequality and $\boldsymbol{U}^{\mathtt{Lim},*,\rho}_{i+\frac{1}{2}} > 0$, $\boldsymbol{U}^{\mathtt{L},\rho}_{i+\frac{1}{2}} > 0$, $\theta^{**}_{i+\frac{1}{2}} \in [0,1]$. Thus the coefficient can be chosen as

$$\theta^{**}_{i+\frac{1}{2}} = \begin{cases} \dfrac{p(\boldsymbol{U}^{\mathtt{L}}_{i+\frac{1}{2}}) - \varepsilon^p_{i+\frac{1}{2}}}{p(\boldsymbol{U}^{\mathtt{L}}_{i+\frac{1}{2}}) - p(\boldsymbol{U}^{\mathtt{Lim},*}_{i+\frac{1}{2}})}, & \text{if } \ p(\boldsymbol{U}^{\mathtt{Lim},*}_{i+\frac{1}{2}}) < \varepsilon^p_{i+\frac{1}{2}}, \\ 1, & \text{otherwise.} \end{cases}$$

**Theorem C.1.** If the initial numerical solution $\overline{\boldsymbol{U}}^0_i, \boldsymbol{U}^0_{i+\frac{1}{2}} \in \mathcal{G}$ for all $i$, and the time step size satisfies (50) and (60), then the AF methods (4)-(5) equipped with the SSP-RK3 (14) and the BP limitings

- (54) and (61) preserve the maximum principle for scalar case;
- (57) and (62) preserve positive density and pressure for the Euler equations.

*Remark* C.1. For uniform meshes, and if taking the maximal spectral radius of $\partial \boldsymbol{F}/\partial \boldsymbol{U}$ in the domain as $\|\varrho\|_\infty$, the following CFL condition

$$\Delta t^n \leqslant \frac{\Delta x}{2\,\|\varrho\|_\infty}$$

fulfills the time step size constraints (50) and (60).

# D Additional numerical results

**Example D.1** (1D accuracy test for the Euler equations)**.** This test is used to examine the accuracy of using different point value updates, following the setup in [1]. The domain is $[-1,1]$ with periodic boundary conditions. The adiabatic index is chosen as $\gamma = 3$ so that the characteristic equations of two Riemann invariants $w = u \pm a$ are $w_t + w w_x = 0$. The initial condition is $\rho_0(x) = 1 + \zeta \sin(\pi x), v_0 = 0, p_0 = \rho_0^\gamma$ and $\zeta \in (0,1)$ controls the range of the density. The exact solution can be obtained by the method of characteristics, given by $\rho(x,t) = \frac{1}{2}\left(\rho_0(x_1) + \rho_0(x_2)\right), v(x,t) = \sqrt{3}\left(\rho(x,t) - \rho_0(x_1)\right)$, where $x_1$ and $x_2$ are solved from the nonlinear equations $x + \sqrt{3}\rho_0(x_1)t - x_1 = 0$, $x - \sqrt{3}\rho_0(x_2)t - x_2 = 0$. The problem is solved until $T = 0.1$ with $\zeta = 1 - 10^{-7}$.

As $\zeta = 1 - 10^{-7}$, the minimum density and pressure are $10^{-7}$ and $10^{-21}$ respectively, so that the BP limitings are necessary to run this test case. The maximal CFL numbers allowing stable simulations are obtained experimentally, which are around $0.47, 0.43, 0.32, 0.18$ for the JS, LLF, SW, and VH FVS, respectively, thus we run the test with the same CFL number as 0.18. Figure 17 shows the errors and corresponding convergence rates for the conservative variables in the $\ell^1$ norm. It is seen that the JS and all the FVS except for the SW FVS achieve the designed third-order accuracy, showing that our BP limitings do not affect the high-order accuracy. To examine the reason why the scheme based on the SW FVS is only second-order accurate, Figure 18 plots the density and velocity profiles obtained using the SW FVS with 80 cells. One can observe some defects in the density when the velocity is zero, similar to the "sonic point glitch" in the literature [41]. One

possible reason is that the SW FVS is based on the absolute value of the eigenvalues, and the corresponding mass flux is not differentiable when the velocity is zero [44]. Such an issue remains to be further explored in the future.

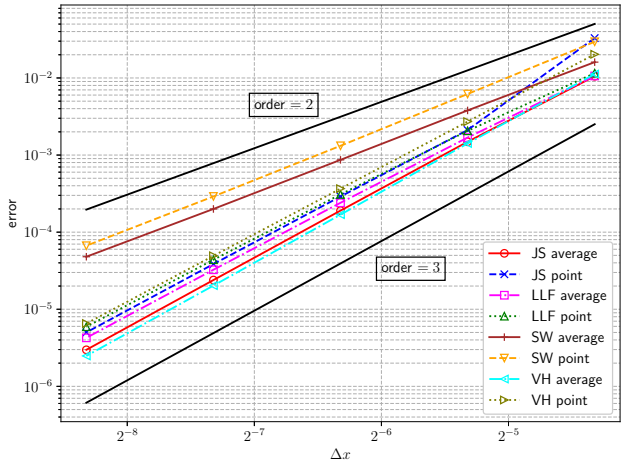


Figure 17: Example D.1, the accuracy test for the 1D Euler equations. The BP limitings are necessary.

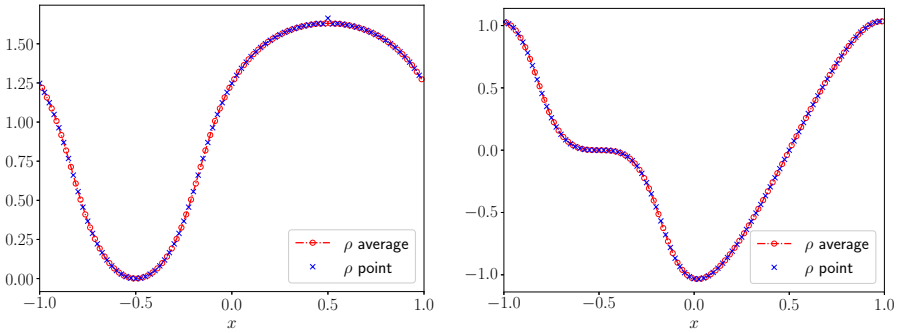


Figure 18: Example D.1, the density (left) and velocity (right) obtained with the SW FVS and 80 cells for the 1D Euler equations.

**Example D.2** (Double rarefaction problem)**.** The exact solution to this problem contains a vacuum, so that it is often used to verify the BP property of numerical methods. The test is solved on a domain $[0, 1]$ until $T = 0.3$ with the initial data

$$(\rho, v, p) = \begin{cases} (7, -1, 0.2), & \text{if } \ x < 0.5, \\ (7, 1, 0.2), & \text{otherwise.} \end{cases}$$

In this test, the AF method based on any kind of point value update mentioned in this paper gives negative density or pressure without the BP limitings. Figure 19 shows the density computed with 400 cells and the BP limitings for the cell average and point value updates. The CFL number is 0.4 for all kinds of point value updates, except for 0.1 for the VH FVS. One observes that the BP AF method gets good performance for this example.

**Example D.3** (Blast wave interaction)**.** The power law reconstruction is useful to reduce oscillations for the fully-discrete AF method [5], thus we would also like to test its ability

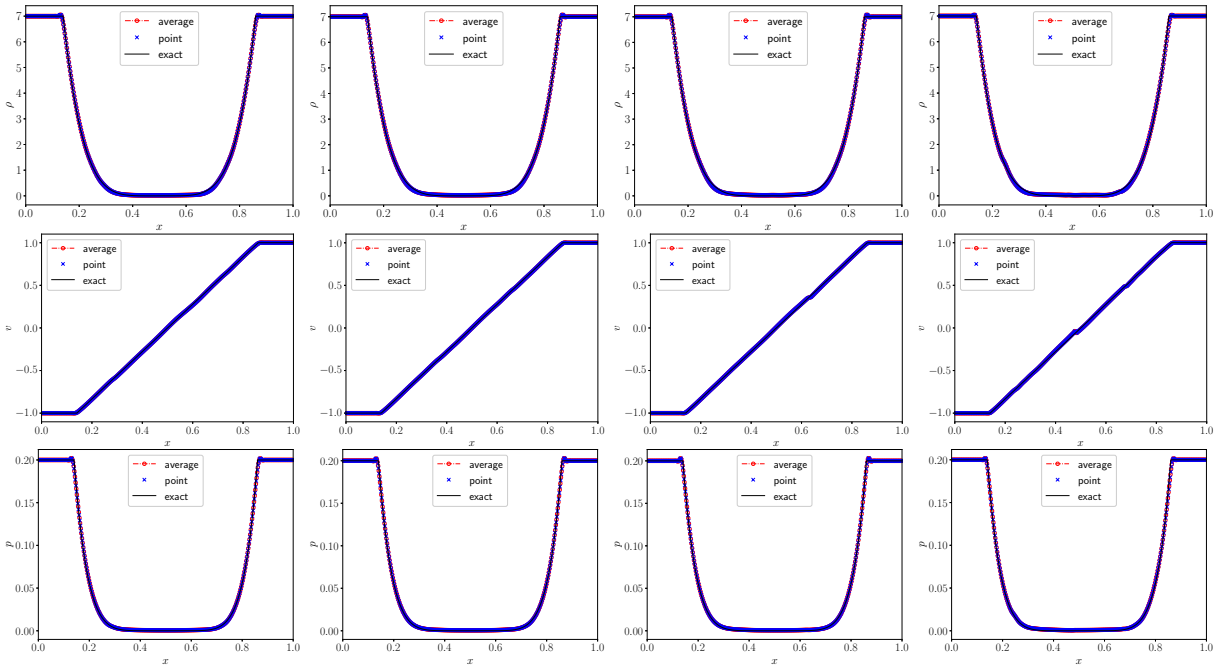


Figure 19: Example D.2, double rarefaction Riemann problem. The density, velocity, and pressure are computed by the BP AF methods on a uniform mesh of 400 cells. From left to right: JS, LLF, SW, and VH FVS.

for the generalized (semi-discrete) AF method. Figure 20 shows the density profiles and corresponding enlarged views obtained by using the BP limitings and power law reconstruction on a uniform mesh of 800 cells. It is seen that the power law reconstruction can suppress oscillations, but the results are still more oscillatory than those using the shock sensor-based limiting. Note that the CFL number reduces to 0.1 when the power law reconstruction is activated. This kind of reduction of the CFL number is also observed in other test cases thus we do not recommend using the power law reconstruction for the generalized AF methods, which also motivates us to develop the shock sensor-based limiting.

**Example D.4** (1D Sedov problem)**.** In this problem, a volume of uniform density and temperature is initialized, and a large quantity of thermal energy is injected at the center, developing into a blast wave that evolves in time in a self-similar fashion [39]. An exact analytical solution based on self-similarity arguments is available [30], which contains very low density with strong shocks. For the background value, the initial density is one, velocity is zero, and total energy is $10^{-12}$ everywhere except that in the centered cell, the total energy of the cell average and point values at two cell interfaces are $3.2 \times 10^6/\Delta x$ with $\Delta x = 4/N$ with $N$ the number of cells, which is used to emulate a $\delta$-function at the center. The test is solved until $T = 10^{-3}$.

This test is run with $N = 801$ cells, and the density plots in the right half domain are shown in Figure 21. The BP limitings are adopted for the cell average and point value updates. The LLF FVS is used and the CFL number is taken as 0.4.

**Example D.5** (Shock reflection problem)**.** The computational domain is $[0, 4] \times [0, 1]$, which is divided into a $120 \times 30$ uniform mesh. The boundary conditions are outflow at the right boundary, reflective at the bottom boundary, and inflow on the other two sides

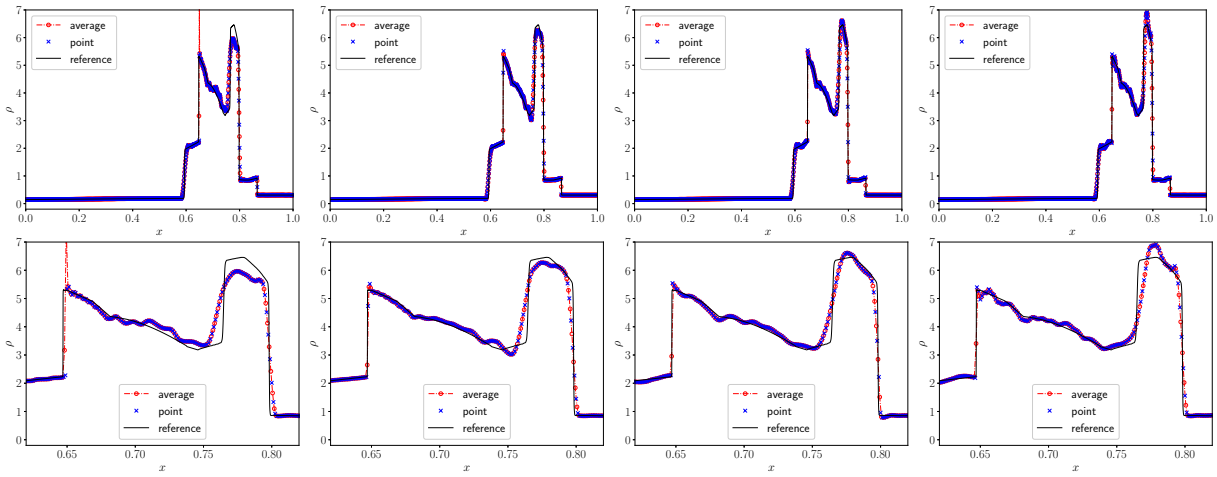


Figure 20: Example 5.3, blast wave interaction. The density computed with the power law reconstruction and BP limitings, and the corresponding enlarged views in $[0.62, 0.82]$ are shown in the bottom row. From left to right: JS, LLF, SW, and VH FVS.

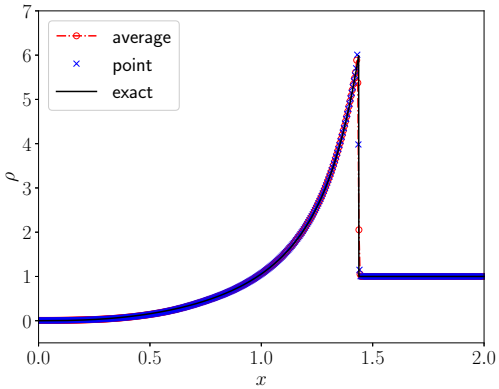


Figure 21: Example D.4, 1D Sedov problem. The numerical solutions are computed with the LLF FVS and the BP limitings on a uniform mesh of 801 cells.

with the data

$$(\rho, v_1, v_2, p) = \begin{cases} (1.0,\ 2.9,\ 0.0,\ 1.0/1.4), & \text{if } x = 0,\ 0 \leqslant y \leqslant 1, \\ (1.69997,\ 2.61934, -0.50632,\ 1.52819), & \text{if } y = 1,\ 0 \leqslant x \leqslant 4. \end{cases}$$

This test is solved until $T = 6$ thus the numerical solution converges.

The density plots obtained without any limiting ($\kappa = 0$) and with the shock sensor-based limiting ($\kappa = 0.5$) are shown in Figure 22, and the blending coefficients based on the shock sensor are plotted in Figure 23. The numerical solutions converge in both cases, and the shock sensor can correctly locate the shock waves. It is also interesting to look at the residual between two successive time steps, presented in Figure 24, with respective to the number of iterations. The limiting based on the shock sensor accelerates the convergence after the reflective shock is fully formed, showing the advantage of using the shock sensor.

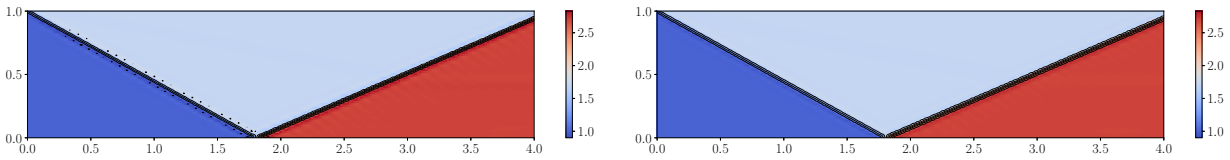

Figure 22: Example D.5, shock reflection problem. The density obtained without ($\kappa = 0$, left) or with the shock sensor ($\kappa = 0.5$, right) on the $120 \times 30$ uniform mesh. 10 equally spaced contour lines from 0.901 to 2.829 are shown.

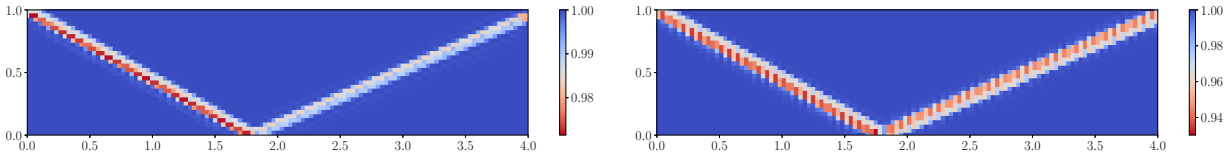

Figure 23: Example D.5, shock reflection problem. The shock sensor-based blending coefficients $\theta^s_{i+\frac{1}{2},j}$ (left) and $\theta^s_{i,j+\frac{1}{2}}$ (right) on the $120 \times 30$ uniform mesh. $\kappa = 0.5$.

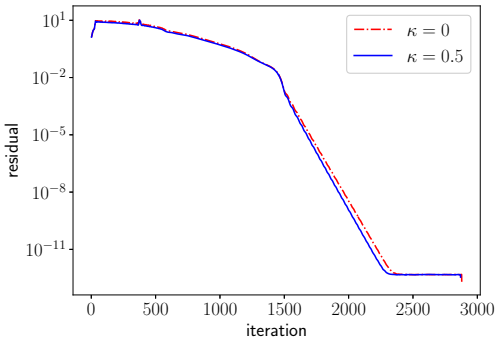


Figure 24: Example D.5, shock reflection problem. The residual decay with respect to the number of iterations.

**Example D.6** (2D Riemann problem)**.** This problem corresponds to the configuration 3 in [33], containing four initial shock waves, with the initial data

$$(\rho, v_1, v_2, p) = \begin{cases} (1.5,\ 0,\ 0,\ 1.5), & x > 0.8,\ y > 0.8, \\ (0.5323,\ 1.206,\ 0,\ 0.3), & x < 0.8,\ y > 0.8, \\ (0.138,\ 1.206,\ 1.206,\ 0.029), & x < 0.8,\ y < 0.8, \\ (0.5323,\ 0,\ 1.206,\ 0.3), & x > 0.8,\ y < 0.8. \end{cases}$$

The test is solved on the domain $[0, 1] \times [0, 1]$ until $T = 0.8$.

Without the BP limitings, the simulation crashes due to negative pressure. The density plots obtained without ($\kappa = 0$) and with the shock sensor ($\kappa = 0.5$) are shown in Figure 25. Without the shock sensor, the numerical solutions contain spurious oscillations, and they are reduced drastically by the shock sensor-based limiting. As mesh refinement, the shock waves are captured sharply, and the small-scale features are preserved well, as evidenced by the roll-ups around the mushroom-shaped jet, which are in good agreement with the results in the literature. The values of the shock sensor-based blending coefficients $\theta_{i+\frac{1}{2},j}, \theta_{i,j+\frac{1}{2}}$ are also plotted in Figure 26, which indicates that the shock sensor can locate the shock waves correctly.

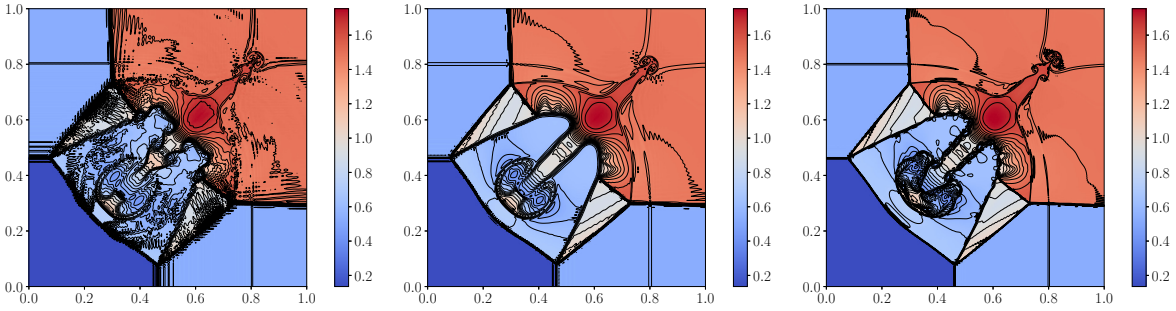

Figure 25: Example D.6, 2D Riemann problem. The density obtained with the BP limitings and without or with the shock sensor. From left to right: $200 \times 200$ mesh with $\kappa = 0$, $200 \times 200$ mesh with $\kappa = 0.5$, $400 \times 400$ mesh with $\kappa = 0.5$. 30 equally spaced contour lines from 0.135 to 1.754.

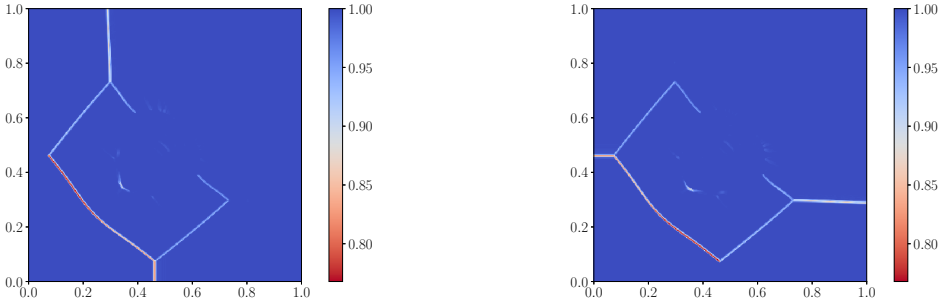

Figure 26: Example D.6, 2D Riemann problem. The shock sensor-based blending coefficients $\theta^s_{i+\frac{1}{2},j}$ (left) and $\theta^s_{i,j+\frac{1}{2}}$ (right) on the $400 \times 400$ uniform mesh.

**Example D.7** (Double Mach reflection)**.** The computational domain is $[0,3]\times[0,1]$ with a reflective wall at the bottom starting from $x=1/6$. A Mach 10 shock is moving towards the bottom wall with an angle of $\pi/6$. The pre- and post-shock states are

$$(\rho, v_1, v_2, p) = \begin{cases} (1.4,\ 0,\ 0,\ 1), & x \geqslant 1/6 + (y+20t)/\sqrt{3}, \\ (8,\ 8.25\cos(\pi/6), -8.25\sin(\pi/6),\ 116.5), & x < 1/6 + (y+20t)/\sqrt{3}. \end{cases}$$

The reflective boundary condition is applied at the wall, while the exact post-shock condition is imposed at the left boundary and for the rest of the bottom boundary (from $x=0$ to $x=1/6$). At the top boundary, the exact motion of the Mach 10 shock is applied and the outflow boundary condition is used at the right boundary. The results are shown at $T=0.2$.

The AF method without the BP limitings gives negative density or pressure near the reflective location $(1/6, 0)$, so the BP limitings are necessary for this test. The numerical solutions are computed without or with the shock sensor ($\kappa = 1$) on a series of uniform meshes. The density plots with enlarged views around the double Mach region are shown in Figure 27, and the blending coefficients based on the shock sensor are shown in Figure 28. When the shock sensor is not activated, the noise after the bow shock is obvious, and it is damped with the help of the shock sensor. As mesh refinement, the numerical solutions converge with a good resolution and are comparable to those in the literature. Compared to the third-order $P^2$ DG method using the TVB limiter [12] with the same mesh resolution ($\Delta x = \Delta y = 1/480$), the roll-ups and vortices are comparable while the AF method uses fewer DoFs (4 versus 6 per cell).

# References


[1] R. Abgrall, *A combination of residual distribution and the active flux formulations or a new class of schemes that can combine several writings of the same hyperbolic problem: Application to the 1D Euler equationss*, Commun. Appl. Math. Comput., 5 (2023), pp. 370–402, https://doi.org/10.1007/s42967-021-00175-w.

[2] R. Abgrall and W. Barsukow, *Extensions of active flux to arbitrary order of accuracy*, ESAIM: Math. Model. Numer. Anal., 57 (2023), pp. 991–1027, https://doi.org/10.1051/m2an/2023004, https://www.esaim-m2an.org/articles/m2an/abs/2023/02/m2an220128/m2an220128.html (accessed 2024-03-03).

[3] R. Abgrall, W. Barsukow, and C. Klingenberg, *The active flux method for the Euler equations on Cartesian grids*, Oct. 2023, https://doi.org/10.48550/arXiv.2310.00683, http://arxiv.org/abs/2310.00683. arXiv:2310.00683.

[4] R. Abgrall, J. Lin, and Y. Liu, *Active flux for triangular meshes for compressible flows problems*, Dec. 2023, https://doi.org/10.48550/arxiv.2312.11271, https://arxiv.org/abs/2312.11271.

[5] W. Barsukow, *The active flux scheme for nonlinear problems*, J. Sci. Comput., 86 (2021), p. 3, https://doi.org/10.1007/s10915-020-01381-z.

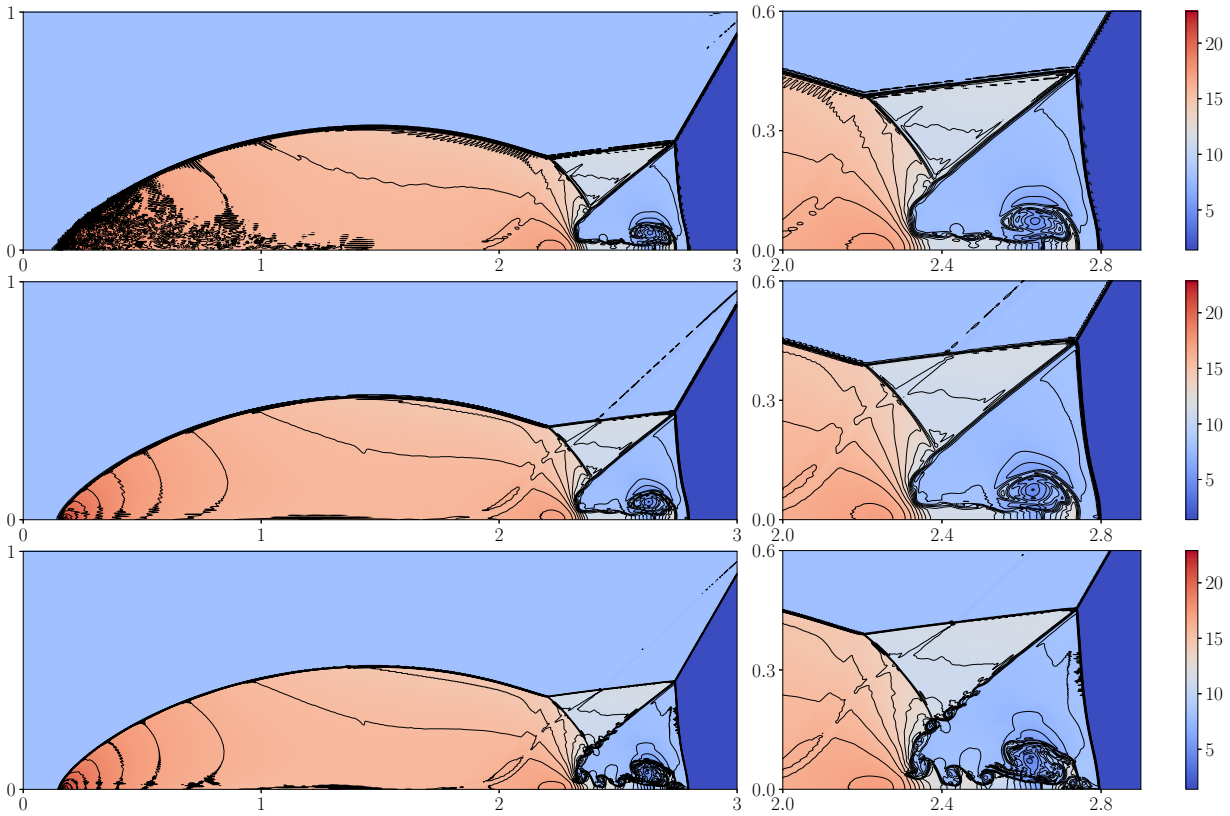

Figure 27: Example D.7, double Mach reflection. The density obtained with the BP limitings and without or with the shock sensor. From top to bottom: $720 \times 240$ mesh without shock sensor, $720 \times 240$ mesh with $\kappa = 1$, $1440 \times 480$ mesh with $\kappa = 1$. 30 equally spaced contour lines from 1.390 to 22.861.

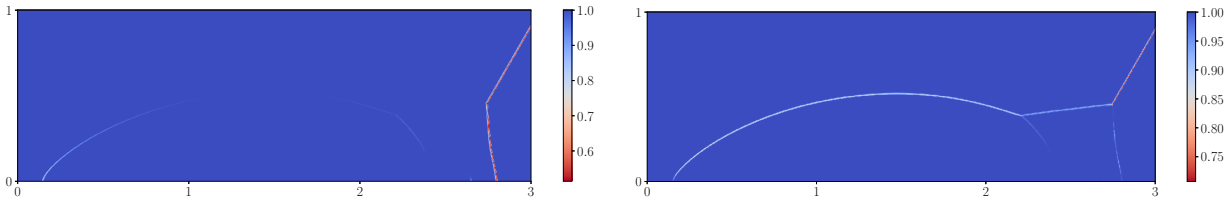

Figure 28: Example D.7, double Mach reflection. The blending coefficients $\theta^s_{i+\frac{1}{2},j}$ (left) and $\theta^s_{i,j+\frac{1}{2}}$ (right) based on the shock sensor with $\kappa = 1$ on the $1440 \times 480$ mesh.

[6] W. Barsukow and J. P. Berberich, *A well-balanced active flux method for the shallow water equations with wetting and drying*, Commun. Appl. Math. Comput., (2023), https://doi.org/10.1007/s42967-022-00241-x.

[7] W. Barsukow, J. P. Berberich, and C. Klingenberg, *On the active flux scheme for hyperbolic PDEs with source terms*, SIAM J. Sci. Comput., 43 (2021), pp. A4015–A4042, https://doi.org/10.1137/20M1346675.

[8] W. Barsukow, J. Hohm, C. Klingenberg, and P. L. Roe, *The active flux scheme on Cartesian grids and its low Mach number limit*, J. Sci. Comput., 81 (2019), pp. 594–622, https://doi.org/10.1007/s10915-019-01031-z.

[9] E. Chudzik and C. Helzel, *A Review of Cartesian Grid Active Flux Methods for Hyperbolic Conservation Laws*, in Finite Volumes for Complex Applications X—Volume 1, Elliptic and Parabolic Problems, E. Franck, J. Fuhrmann, V. Michel-Dansac, and L. Navoret, eds., Cham, 2023, Springer Nature Switzerland, pp. 93–109, https://doi.org/10.1007/978-3-031-40864-9_6.

[10] E. Chudzik, C. Helzel, and D. Kerkmann, *The Cartesian grid active flux method: Linear stability and bound preserving limiting*, Appl. Math. Comput., 393 (2021), p. 125501, https://doi.org/10.1016/j.amc.2020.125501, https://www.sciencedirect.com/science/article/pii/S0096300320304598 (accessed 2024-03-03).

[11] S. Clain, S. Diot, and R. Loubère, *A high-order finite volume method for systems of conservation laws–Multi-dimensional Optimal Order Detection (MOOD)*, J. Comput. Phys., 230 (2011), pp. 4028–4050, https://doi.org/10.1016/j.jcp.2011.02.026, https://www.sciencedirect.com/science/article/pii/S002199911100115X (accessed 2024-03-11).

[12] B. Cockburn and C. W. Shu, *Runge-Kutta discontinuous Galerkin methods for convection-dominated problems*, J. Sci. Comput., 16 (2001), pp. 173–261, https://doi.org/10.1023/a:1012873910884.

[13] C. J. Cotter and D. Kuzmin, *Embedded discontinuous Galerkin transport schemes with localised limiters*, J. Comput. Phys., 311 (2016), pp. 363–373, https://doi.org/10.1016/j.jcp.2016.02.021, https://www.sciencedirect.com/science/article/pii/S0021999116000759 (accessed 2024-03-12).

[14] C. M. Dafermos, *Hyperbolic Conservation Laws in Continuum Physics*, Springer Berlin Heidelberg, 2000, https://doi.org/10.1007/978-3-662-22019-1.

[15] F. Ducros, V. Ferrand, F. Nicoud, C. Weber, D. Darracq, C. Gacherieu, and T. Poinsot, *Large-eddy simulation of the shock/turbulence interaction*, Journal of Computational Physics, 152 (1999), pp. 517–549, https://doi.org/10.1006/jcph.1999.6238, https://www.sciencedirect.com/science/article/pii/S0021999199962381 (accessed 2024-05-29).

[16] T. Eymann and P. Roe, *Active flux schemes*, in 49th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition, Orlando, Florida, Jan. 2011, American Institute of Aeronautics and Astronautics, https://doi.org/10.2514/6.2011-382.

[17] T. Eymann and P. Roe, *Active flux schemes for systems*, in 20th AIAA Computational Fluid Dynamics Conference, Fluid Dynamics and Co-located Conferences, American Institute of Aeronautics and Astronautics, June 2011, https://doi.org/10.2514/6.2011-3840.

[18] T. A. Eymann and P. L. Roe, *Multidimensional active flux schemes*, in 21st AIAA Computational Fluid Dynamics Conference, Fluid Dynamics and Co-located Conferences, American Institute of Aeronautics and Astronautics, June 2013, https://doi.org/10.2514/6.2013-2940.

[19] D. Fan and P. L. Roe, *Investigations of a new scheme for wave propagation*, in 22nd AIAA Computational Fluid Dynamics Conference, American Institute of Aeronautics and Astronautics, 2015, https://doi.org/10.2514/6.2015-2449.

[20] S. Gottlieb, C.-W. Shu, and E. Tadmor, *Strong Stability-Preserving High-Order Time Discretization Methods*, SIAM Rev., 43 (2001), pp. 89–112, https://doi.org/10.1137/S003614450036757X.

[21] J.-L. Guermond, M. Nazarov, B. Popov, and I. Tomas, *Second-order invariant domain preserving approximation of the Euler equations using convex limiting*, SIAM J. Sci. Comput., 40 (2018), pp. A3211–A3239, https://doi.org/10.1137/17M1149961.

[22] J.-L. Guermond and B. Popov, *Invariant domains and first-order continuous finite element approximation for hyperbolic systems*, SIAM J. Numer. Anal., 54 (2016), pp. 2466–2489, https://doi.org/10.1137/16M1074291.

[23] J.-L. Guermond and B. Popov, *Invariant domains and second-order continuous finite element approximation for scalar conservation equations*, SIAM J. Numer. Anal., 55 (2017), pp. 3120–3146, https://doi.org/10.1137/16M1106560.

[24] J.-L. Guermond, B. Popov, and I. Tomas, *Invariant domain preserving discretization-independent schemes and convex limiting for hyperbolic systems*, Comput. Method. Appl. M., 347 (2019), pp. 143–175, https://doi.org/10.1016/j.cma.2018.11.036, https://www.sciencedirect.com/science/article/pii/S0045782518305954 (accessed 2024-03-12).

[25] H. Hajduk, *Monolithic convex limiting in discontinuous Galerkin discretizations of hyperbolic conservation laws*, Comput. Math. Appl., 87 (2021), pp. 120–138, https://doi.org/10.1016/j.camwa.2021.02.012, https://www.sciencedirect.com/science/article/pii/S0898122121000547 (accessed 2024-03-12).

[26] D. Hänel, R. Schwane, and G. Seider, *On the accuracy of upwind schemes for the solution of the Navier-Stokes equations*, Fluid Dynamics and Co-located Conferences, American Institute of Aeronautics and Astronautics, June 1987, https://doi.org/10.2514/6.1987-1105.

[27] C. Helzel, D. Kerkmann, and L. Scandurra, *A new ADER method inspired by the active flux method*, J. Sci. Comput., 80 (2019), pp. 1463–1497, https://doi.org/10.1007/s10915-019-00988-1.

[28] X. Y. Hu, N. A. Adams, and C.-W. Shu, *Positivity-preserving method for high-order conservative schemes solving compressible Euler equations*, J. Comput. Phys., 242 (2013), pp. 169–180, https://doi.org/10.1016/j.jcp.2013.01.024, https://www.sciencedirect.com/science/article/pii/S0021999113000557 (accessed 2024-03-13).

[29] A. Jameson, W. Schmidt, and E. Turkel, *Solutions of the Euler equations by finite volume methods using Runge-Kutta time-stepping schemes*, AIAA J., 1259 (1981).

[30] J. R. Kamm and F. X. Timmes, *On efficient generation of numerically robust Sedov solutions*, Tech. Report LA-UR-07-2849, 2007.

[31] D. Kuzmin, *Monolithic convex limiting for continuous finite element discretizations of hyperbolic conservation laws*, Computer Methods in Applied Mechanics and Engineering, 361 (2020), p. 112804, https://doi.org/10.1016/j.cma.2019.112804, https://www.sciencedirect.com/science/article/pii/S0045782519306966 (accessed 2024-03-12).

[32] D. Kuzmin, R. Löhner, and S. Turek, eds., *Flux-Corrected Transport: Principles, Algorithms, and Applications*, Scientific Computation, Springer Netherlands, Dordrecht, 2012, https://doi.org/10.1007/978-94-007-4038-9.

[33] P. D. Lax and X.-D. Liu, *Solution of two-dimensional Riemann problems of gas dynamics by positive schemes*, SIAM J. Sci. Comput., 19 (1998), pp. 319–340, https://doi.org/10.1137/s1064827595291819.

[34] X.-D. Liu and S. Osher, *Nonoscillatory high order accurate self-similar maximum principle satisfying shock capturing schemes I*, SIAM J. Numer. Anal., 33 (1996), pp. 760–779, https://doi.org/10.1137/0733038.

[35] C. Lohmann, D. Kuzmin, J. N. Shadid, and S. Mabuza, *Flux-corrected transport algorithms for continuous Galerkin methods based on high order Bernstein finite elements*, J. Comput. Phys., 344 (2017), pp. 151–186, https://doi.org/10.1016/j.jcp.2017.04.059, https://www.sciencedirect.com/science/article/pii/S0021999117303388 (accessed 2024-03-12).

[36] J. Maeng, *On the Advective Component of Active Flux Schemes for Nonlinear Hyperbolic Conservation Laws*, PhD thesis, 2017, http://deepblue.lib.umich.edu/handle/2027.42/138695 (accessed 2024-10-19).

[37] B. Perthame and C.-W. Shu, *On positivity preserving finite volume schemes for Euler equations*, Numer. Math., 73 (1996), pp. 119–130, https://doi.org/10.1007/s002110050187.

[38] P. Roe, *Is discontinuous reconstruction really a good idea?*, J. Sci. Comput., 73 (2017), pp. 1094–1114, https://doi.org/10.1007/s10915-017-0555-z.

[39] L. I. Sedov, *Similarity and Dimensional Methods in Mechanics*, Academic Press, New York,, 1959.

[40] J. L. Steger and R. F. Warming, *Flux vector splitting of the inviscid gasdynamic equations with application to finite-difference methods*, J. Comput. Phys., 40 (1981), pp. 263–293, https://doi.org/10.1016/0021-9991(81)90210-2, https://www.sciencedirect.com/science/article/pii/0021999181902102 (accessed 2024-03-05).

[41] H. Z. Tang, *On the sonic point glitch*, J. Comput. Phys., 202 (2005), pp. 507–532, https://doi.org/10.1016/j.jcp.2004.07.013, https://www.sciencedirect.com/science/article/pii/S0021999104002967 (accessed 2024-03-21).

[42] E. F. Toro, *Riemann Solvers and Numerical Methods for Fluid Dynamics*, Springer Berlin Heidelberg, 2009, https://doi.org/10.1007/b79761.

[43] B. Van Leer, *Towards the ultimate conservative difference scheme. IV. A new approach to numerical convection*, J. Comput. Phys., 23 (1977), pp. 276–299, https://doi.org/10.1016/0021-9991(77)90095-X, https://www.sciencedirect.com/science/article/pii/002199917790095X (accessed 2024-03-09).

[44] B. van Leer, *Flux-vector splitting for the Euler equations*, in Eighth International Conference on Numerical Methods in Fluid Dynamics, E. Krause, ed., Lecture Notes in Physics, Berlin, Heidelberg, 1982, Springer, pp. 507–512, https://doi.org/10.1007/3-540-11948-5_66.

[45] P. Woodward and P. Colella, *The numerical simulation of two-dimensional fluid flow with strong shocks*, J. Comput. Phys., 54 (1984), pp. 115–173, https://doi.org/10.1016/0021-9991(84)90142-6, https://www.sciencedirect.com/science/article/pii/0021999184901426 (accessed 2024-03-08).

[46] K. Wu and C.-W. Shu, *Geometric quasilinearization framework for analysis and design of bound-preserving schemes*, SIAM Rev., 65 (2023), pp. 1031–1073, https://doi.org/10.1137/21M1458247.

[47] Z. Xu, *Parametrized maximum principle preserving flux limiters for high order schemes solving hyperbolic conservation laws: one-dimensional scalar problem*, Math. Comput., 83 (2014), pp. 2213–2238, https://doi.org/10.1090/S0025-5718-2013-02788-3, https://www.ams.org/mcom/2014-83-289/S0025-5718-2013-02788-3/ (accessed 2024-03-13).

[48] X. Zhang and C.-W. Shu, *On maximum-principle-satisfying high order schemes for scalar conservation laws*, J. Comput. Phys., 229 (2010), pp. 3091–3120, https://doi.org/10.1016/j.jcp.2009.12.030, https://www.sciencedirect.com/science/article/pii/S0021999109007165 (accessed 2024-06-14).

[49] X. Zhang and C.-W. Shu, *On positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations on rectangular meshes*, J. Comput. Phys., 229 (2010), pp. 8918–8934, https://doi.org/10.1016/j.jcp.2010.08.016,

https://www.sciencedirect.com/science/article/pii/S0021999110004535 (accessed 2024-03-26).

[50] X. Zhang and C.-W. Shu, *Maximum-principle-satisfying and positivity-preserving high-order schemes for conservation laws: survey and new developments*, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 467 (2011), pp. 2752–2776, https://doi.org/10.1098/rspa.2011.0153.

[51] X. Zhang and C.-W. Shu, *Positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations with source terms*, J. Comput. Phys., 230 (2011), pp. 1238–1248, https://doi.org/10.1016/j.jcp.2010.10.036, https://www.sciencedirect.com/science/article/pii/S0021999110006017 (accessed 2024-03-13).