Contribution à l’approximation numérique des systèmes hyperboliques

Vivien Desveaux

sous la direction de Christophe Berthon et Yves Coudière

Laboratoire de mathématiques Jean Leray - Université de Nantes

26 novembre 2013
1. Second-order schemes based on dual mesh gradient reconstruction

2. A high-order entropy preserving scheme with \textit{a posteriori} limitation

3. Well-balanced schemes for systems with nonlinear source terms
1. Second-order schemes based on dual mesh gradient reconstruction
   - MUSCL scheme and CFL condition
   - The DMGR scheme
   - Numerical results

2. A high-order entropy preserving scheme with a posteriori limitation
   - Motivations
   - From one to all discrete entropy inequalities
   - The e-MOOD scheme for the Euler equations

3. Well-balanced schemes for systems with nonlinear source terms
   - The Ripa model
   - Relaxation models
   - The relaxation scheme
   - Numerical results
Introduction

- Hyperbolic system of conservation laws in 2D

\[ \partial_t w + \partial_x f(w) + \partial_y g(w) = 0 \]

- \( w : \mathbb{R}^+ \times \mathbb{R}^2 \rightarrow \Omega \subset \mathbb{R}^d \): unknown state vector
- \( f, g : \Omega \rightarrow \mathbb{R}^d \): flux functions
- \( \Omega \) convex set of physical states
- **Objective**: derive a numerical scheme
  - second order accurate
  - \( \Omega \)-preserving
  - works on unstructured meshes
  - with an optimized CFL condition
Mesh notations

- polygonal cells $K_i$ (perimeter $\mathcal{P}_i$, area $|K_i|$)
- $\gamma(i)$: index set of the cells neighbouring $K_i$
- $e_{ij}$: common edge between $K_i$ and $K_j$ (length $|e_{ij}|$)
- $\nu_{ij}$: unit outward normal to $e_{ij}$
First-order scheme

\[ w_i^{n+1} = w_i^n - \frac{\Delta t}{|K_i|} \sum_{j \in \gamma(i)} |e_{ij}| \varphi \left( w_i^n, w_j^n, \nu_{ij} \right) \]

- **2D numerical flux**: \( \varphi \) Godunov-type in each direction \( \nu \)
  [Harten, Lax & van Leer ’83]:

\[
\varphi(w_L, w_R, \nu) = h_\nu(w_L) + \frac{\delta}{2\Delta t} w_L - \frac{1}{\Delta t} \int_{-\frac{\delta}{2}}^{\frac{\delta}{2}} \tilde{w}_\nu \left( \frac{x}{\Delta t}, w_L, w_R \right) \, dx
\]

with respect to the CFL condition \( \frac{\Delta t}{\delta} \max |\lambda^\pm(w_L, w_R, \nu)| \leq \frac{1}{2} \)

- \( h_\nu(w) = \nu_x f(w) + \nu_y g(w) \): flux in the \( \nu \)-direction, with \( \nu = (\nu_x, \nu_y)^T \)
- \( \tilde{w}_\nu \) approximate Riemann solver valued in \( \Omega \)

- **Consistency**: \( \varphi(w, w, \nu) = h_\nu(w) \)
- **Conservation**: \( \varphi(w_L, w_R, \nu) = -\varphi(w_R, w_L, -\nu) \)
Theorem (Robustness of the first-order scheme)

Assume the following CFL condition is satisfied:

\[ \Delta t \frac{P_i}{|K_i|} \max_{j \in \gamma(i)} \left| \lambda^\pm(w^n_i, w^n_j, \nu_{ij}) \right| \leq 1, \quad \forall i \in \mathbb{Z}. \]

Then the first-order scheme preserves \( \Omega \):

\[ \forall i \in \mathbb{Z}, \quad w^n_i \in \Omega \quad \Rightarrow \quad \forall i \in \mathbb{Z}, \quad w^{n+1}_i \in \Omega. \]

Remark: in [Perthame & Shu '96], they obtain the CFL restriction

\[ \Delta t \frac{P_i}{|K_i|} \max_{j \in \gamma(i)} \left| \lambda^\pm(w^n_i, w^n_j, \nu_{ij}) \right| \leq \frac{1}{2}, \quad \forall i \in \mathbb{Z}. \]

However this can easily be improved in the Godunov-type framework.
MUSCL scheme ([van Leer ’79], [Perthame & Shu ’96]...)

First-order scheme on the cell $K_i$

$$w_i^{n+1} = w_i^n - \frac{\Delta t}{|K_i|} \sum_{j \in \gamma(i)} |e_{ij}| \varphi \left( w_i^n, w_j^n, \nu_{ij} \right)$$

Second-order MUSCL scheme on the cell $K_i$

$$w_i^{n+1} = w_i^n - \frac{\Delta t}{|K_i|} \sum_{j \in \gamma(i)} |e_{ij}| \varphi \left( w_{ij}, w_{ji}, \nu_{ij} \right)$$

$w_{ij}$ and $w_{ji}$ are second-order approximations at the interface between $K_i$ and $K_j$.

→ How to compute $w_{ij}$?
Subcells decomposition

- $T_{ij}$: triangle formed by the mass center $G_i$ and the edge $e_{ij}$ (perimeter $P_{ij}$, area $|T_{ij}|$)
- $\gamma(i, j)$: index set of the two subcells neighbouring $T_{ij}$ in $K_i$
- $e^i_{jk}$: common edge between $T_{ij}$ and $T_{ik}$ (length $|e^i_{jk}|$)
- $\nu^i_{jk}$: unit outward normal to $e^i_{jk}$
Theorem (Robustness of the MUSCL scheme)

Assume the following hypotheses:

(i) The reconstruction preserves $\Omega$:
\[ \forall i \in \mathbb{Z}, \ w^n_i \in \Omega \Rightarrow \forall i \in \mathbb{Z}, \ \forall j \in \gamma(i), \ w_{ij} \in \Omega. \]

(ii) The reconstruction satisfies the conservation property
\[ \sum_{j \in \gamma(i)} \frac{|T_{ij}|}{|K_i|} w_{ij} = w^n_i. \]

(iii) The following CFL condition is satisfied for all $i \in \mathbb{Z}$:
\[ \Delta t \max_{j \in \gamma(i)} \frac{P_{ij}}{|T_{ij}|} \max_{k \in \gamma(i,j)} \left| \lambda^\pm(w_{ij}, w_{ji}, \nu_{ij}), \lambda^\pm(w_{ij}, w_{ik}, \nu_{jk}) \right| \leq 1 \]

Then the MUSCL scheme preserves $\Omega$:
\[ \forall i \in \mathbb{Z}, \ w^n_i \in \Omega \Rightarrow \forall i \in \mathbb{Z}, \ w^{n+1}_i \in \Omega. \]
The DMGR scheme

- Vertex of the primal mesh (unknown state)
- Center of a primal cell
  = Vertex of the dual mesh (known state)
- Center of a dual cell (known state)

Figure: Primal mesh (blue) and dual mesh (red)

We write a MUSCL scheme on both primal and dual meshes
⇒ At time $t^n$, we know a state at the center of each primal and dual cell
The DMGR scheme

The DMGR scheme

Figure: Primal mesh (blue) and dual mesh (red)

Assume we have a reconstruction procedure on a generic cell $K$:

\[
\begin{cases}
\text{state at the center} \\
\text{states at the vertices}
\end{cases} \implies \text{linear reconstruction } \tilde{w}
\]

This procedure will be detailed after.
The DMGR scheme

- Vertex of the primal mesh (unknown state)
- Center of a primal cell = Vertex of the dual mesh (known state)
- Center of a dual cell (known state)

Figure: Primal mesh (blue) and dual mesh (red)

1. We can apply the reconstruction procedure on the dual cells
   ⇒ We get a linear function $\tilde{w}_i^d$ on each dual cell

2. To get the state at the primal vertex $S_i^p$, we take $\tilde{w}_i^d(S_i^p)$
   ⇒ We can apply the reconstruction procedure on the primal cells to get a linear function $\tilde{w}_i^p$
Reconstruction procedure

Geometry of the cell $K$

The states $\hat{w}_{j-1/2}$ have to satisfy:

- $\hat{w}_{j-1/2} \in \Omega$
- $\sum_j \frac{|T_{j-1/2}|}{|K|} \hat{w}_{j-1/2} = w_0$

If we take $\hat{w}_{j-1/2} = \tilde{w}(Q_{j-1/2})$ with $\tilde{w}$ a linear function on $K$, we have

$$\sum_j \frac{|T_{j-1/2}|}{|K|} \hat{w}_{j-1/2} = w_0 \iff \tilde{w}(G) = w_0$$
Reconstruction procedure

Geometry of the cell $K$

Known states and reconstructed states

Gradient reconstruction

We define a continuous function $\overline{w} : K \to \mathbb{R}^d$ piecewise linear on each triangle $T_{j-1/2}$ and such that $\overline{w}(S_j) = w_j$ and $\overline{w}(G) = w_0$. 
Reconstruction procedure

Geometry of the cell $K$

Known states and reconstructed states

**Projection**

For a slope $\alpha \in M_{d,2}(\mathbb{R})$, we define $\tilde{w}_\alpha(X) = w_0 + \alpha \cdot (X - G)$, the linear function whose gradient is $\alpha$.

Let $\mu$ be the slope resulting from the $L^2$–projection of $\overline{w}$:

$$\int_K \|\overline{w}(X) - \tilde{w}_\mu(X)\|^2 dX = \min_{\alpha \in M_{d,2}(\mathbb{R})} \int_K \|\overline{w}(X) - \tilde{w}_\alpha(X)\|^2 dX.$$
Reconstruction procedure

Geometry of the cell $K$

Known states and reconstructed states

### Limitation of the slope $\mu$

We define the optimal slope limiter by:

$$\alpha_{j-1/2} = \sup \left\{ \theta \in [0, 1], \tilde{w}_s\mu(Q_{j-1/2}) \in \Omega, \forall s \in [0, \theta] \right\},$$

$$\beta = \min_j \alpha_{j-1/2} - \epsilon,$$

where $\epsilon > 0$ is a small parameter s.t. $\tilde{w}_{\beta\mu}(Q_{j-1/2}) \in \Omega, \forall j$. 

Reconstruction procedure

Geometry of the cell $K$

Finally, the reconstructed states are given by

$$\hat{w}_{j-1/2} = \tilde{w}_\beta \mu (Q_{j-1/2}).$$

Limitation procedure  \Rightarrow  $$\hat{w}_{j-1/2} \in \Omega$$

$$\tilde{w}(G) = w_0$$

\Rightarrow  $$\sum_j \frac{|T_{j-1/2}|}{|K|} \hat{w}_{j-1/2} = w_0$$

\Rightarrow  The DMGR scheme preserves $\Omega$
Numerical results: 2D Euler equations

\[
\partial_t \begin{pmatrix}
\rho \\
\rho u \\
\rho v \\
E
\end{pmatrix} + \partial_x \begin{pmatrix}
\rho u \\
\rho u^2 + p \\
\rho uv \\
u(E + p)
\end{pmatrix} + \partial_y \begin{pmatrix}
\rho v \\
\rho uv \\
\rho v^2 + p \\
v(E + p)
\end{pmatrix} = 0
\]

- \(\rho\): density
- \((u, v)\): velocity
- \(E\): total energy
- \(p\): pressure given by the ideal gas law

\[
p = (\gamma - 1) \left( E - \frac{\rho}{2} \left( u^2 + v^2 \right) \right), \quad \text{with } \gamma \in (1, 3]
\]

- Set of physical states

\[
\Omega = \left\{ (\rho, \rho u, \rho v, E) \in \mathbb{R}^4; \quad \rho > 0, \quad E - \frac{\rho}{2} \left( u^2 + v^2 \right) > 0 \right\}
\]
2D Riemann problems

**Figure:** Four shocks 2D Riemann problem on a Cartesian mesh with $1.5 \times 10^6$ DOF

**Figure:** Four contact discontinuities 2D Riemann problem on a Cartesian mesh with $1.5 \times 10^6$ DOF
Double Mach reflection on a ramp

Figure: Double Mach reflection on a ramp on an unstructured mesh with $3 \times 10^6$ DOF
Mach 3 wind tunnel with a step

Figure: Mach 3 tunnel with a step on an unstructured mesh with $1.5 \times 10^6$ DOF
# A high-order entropy preserving scheme

## Second-order schemes based on dual mesh gradient reconstruction
- MUSCL scheme and CFL condition
- The DMGR scheme
- Numerical results

## A high-order entropy preserving scheme with *a posteriori* limitation
- Motivations
- From one to all discrete entropy inequalities
- The e-MOOD scheme for the Euler equations

## Well-balanced schemes for systems with nonlinear source terms
- The Ripa model
- Relaxation models
- The relaxation scheme
- Numerical results
Introduction

- Hyperbolic system of conservation laws in 1D

\[ \partial_t w + \partial_x f(w) = 0 \]

\( w : \mathbb{R}^+ \times \mathbb{R} \rightarrow \Omega \subset \mathbb{R}^d: \) unknown state vector

\( f : \Omega \rightarrow \mathbb{R}^d: \) flux function

- \( \Omega \) convex set of physical states

- Entropy inequalities:

\[ \partial_t \eta(w) + \partial_x \mathcal{G}(w) \leq 0, \]

where \( w \mapsto \eta(w) \) is convex and \( \nabla_w f \nabla_w \eta = \nabla_w \mathcal{G} \)

- Objectives:
  - Study the entropy stability of high-order schemes
  - Derive a high-order numerical scheme for the Euler equations which is entropy preserving in the sense of Lax-Wendroff
Scheme notations

- Space discretization: cells $K_i = [x_{i-1/2}, x_{i+1/2}]$ with constant size $\Delta x = x_{i+1/2} - x_{i-1/2}$
- $w_i^n$: approximate solution at time $t^n$ on the cell $K_i$
- Update at time $t^{n+1} = t^n + \Delta t$ given by
  \[
  w_i^{n+1} = w_i^n - \frac{\Delta t}{\Delta x} \left( F_{i+1/2}^n - F_{i-1/2}^n \right)
  \]
  where $F_{i+1/2}^n = F(w_{i-s+1}^n, \cdots, w_{i+s}^n)$ and $F$ is a consistent numerical flux ($F(w, \cdots, w) = f(w)$)
- We introduce the piecewise constant function
  \[
  w^\Delta(x, t) = w_i^n, \quad \text{for } (x, t) \in K_i \times [t^n, t^{n+1})
  \]
- The sequence $(\Delta x, \Delta t)$ is devoted to converge to $(0, 0)$, the ratio $\frac{\Delta t}{\Delta x}$ being kept constant.
Lax-Wendroff Theorem

**Theorem**

(i) Assume the following hypotheses:

- There exists a compact $K \subset \Omega$ such that $w^\Delta \in K$;
- $w^\Delta$ converges in $L^1_{loc}(\mathbb{R} \times \mathbb{R}^+; \Omega)$ to a function $w$.

Then $w$ is a weak solution.

(ii) Assume the additional hypothesis:

- For all entropy pair $(\eta, G)$, there exists an entropy numerical flux $G$, consistant with $G(G(w, \cdots, w) = G(w))$, such that we have the discrete entropy inequality

$$\frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} + \frac{G_{i+1/2}^n - G_{i-1/2}^n}{\Delta x} \leq 0,$$

with $G_{i+1/2}^n = G(w_{i-s+1}^n, \cdots, w_{i+s}^n)$.

Then $w$ is an entropic solution.
Example: the MUSCL scheme

- Limited slope \( \mu_i^n = L (w_i^n - w_{i-1}^n, w_{i+1}^n - w_i^n) \), with \( L \) a limiter
- The MUSCL flux is defined by
  \[
  F_{i+1/2}^n = F \left( w_i^n + \mu_i^n/2, w_{i+1}^n - \mu_{i+1}^n/2 \right)
  \]
- The known discrete entropy inequalities satisfied by the MUSCL scheme all write
  \[
  \frac{\eta(w_{i+1}^{n+1}) - \eta(w_i^n)}{\Delta t} + \frac{G_{i+1/2}^n - G_{i-1/2}^n}{\Delta x} \leq \frac{P_i^n - \eta(w_i^n)}{\Delta t}
  \]

where \( P_i^n = P (w_i^n, \mu_i^n, \Delta x, \eta) \).

- Examples of operator \( P \):
  \[
  P_1(w, \mu, \Delta x, \eta) = \frac{1}{\Delta x} \int_{-\Delta x/2}^{\Delta x/2} \eta \left( w + \frac{x}{\Delta x} \mu \right) dx \quad \text{[Bouchut et al. '96]}
  \]
  \[
  P_2(w, \mu, \Delta x, \eta) = \frac{\eta(w - \mu/2) + \eta(w + \mu/2)}{2} \quad \text{[Berthon '05]}
  \]
Convergence study

- The discrete entropy inequality

\[
\frac{\eta(w_{i}^{n+1}) - \eta(w_{i}^{n})}{\Delta t} + \frac{G_{i+1/2}^{n} - G_{i-1/2}^{n}}{\Delta x} \leq \frac{P_{i}^{n} - \eta(w_{i}^{n})}{\Delta t}
\]

converges weakly to

\[
\partial_{t}\eta(w) + \partial_{x}G(w) \leq \delta,
\]

where \(\delta\) is a positive measure.

Conjecture (Hou-LeFloch ’94)

- \(\delta = 0\) in the areas where \(w\) is smooth
- \(\delta > 0\) on the curves of discontinuity of \(w\)
Numerical study: test cases (Euler equations)

Entropy error (total mass of the right-hand side):

\[ I^\Delta = \Delta x \sum_{i,n} (P^n_i - \eta(w^n_i)) \]

1–rarefaction

\[ I^\Delta \to 0 \]

Double shock

\[ I^\Delta \to c > 0 \]
Numerical results obtained with a second-order time scheme

1-rarefaction

Double shock
Conclusion of motivations

- Numerical results confirm the Hou-le Floch conjecture: when the scheme converges, the measure $\delta$ seems to be concentrated on the curves of discontinuity of $w$.
- This does not imply that the limit is not entropic, but the usual discrete entropy inequalities are not relevant to apply the Lax-Wendroff theorem.
- We have to enforce the stronger discrete entropy inequalities

$$\frac{\eta(w_i^{n+1}) - \eta(w_i^n)}{\Delta t} + \frac{G_{i+1/2}^n - G_{i-1/2}^n}{\Delta x} \leq 0.$$ 

- We suggest to extend the $a posteriori$ methods (MOOD) introduced in [Clain, Diot & Loubère ’11].
The family of entropies for the Euler equations

**Lemma**

The entropy pairs $(\eta, G)$ of the Euler system write

$$
\eta = \rho \psi(r), \quad G = \rho \psi(r) u,
$$

where $r = -\frac{p^{1/\gamma}}{\rho}$ and $\psi$ is a smooth increasing convex function.

We consider the scheme

$$
w_{i+1}^n = w_i^n - \frac{\Delta t}{\Delta x} \left( F_{i+1/2} - F_{i-1/2} \right),
$$

where $w_i^n = (\rho_i^n, \rho_i^n u_i^n, E_i^n)^T$ and $F_{i+1/2} = (F_{i+1/2}^\rho, F_{i+1/2}^{\rho u}, F_{i+1/2}^E)^T$. 
We introduce $r_{i+1/2}^n = \begin{cases} r_{i+1}^n & \text{if } F_{i+1/2}^\rho < 0 \\ r_i^n & \text{if } F_{i+1/2}^\rho > 0 \end{cases}$. 

**Theorem**

Assume the scheme preserves $\Omega$. Assume the scheme satisfies the specific discrete entropy inequality

$$\rho_{i+1}^{n+1} r_{i+1}^{n+1} \leq \rho_i^n r_i^n - \frac{\Delta t}{\Delta x} \left( F_{i+1/2}^\rho r_{i+1/2}^n - F_{i-1/2}^\rho r_{i-1/2}^n \right).$$

Assume the additional CFL like condition

$$\frac{\Delta t}{\Delta x} \left( \max \left( 0, F_{i+1/2}^\rho \right) - \min \left( 0, F_{i-1/2}^\rho \right) \right) \leq \rho_i^n.$$

Then the scheme is entropy preserving: for all smooth increasing convex function $\psi$, we have

$$\rho_{i+1}^{n+1} \psi(r_{i+1}^{n+1}) \leq \rho_i^n \psi(r_i^n) - \frac{\Delta t}{\Delta x} \left( F_{i+1/2}^\rho \psi(r_{i+1/2}^n) - F_{i-1/2}^\rho \psi(r_{i-1/2}^n) \right).$$
First-order scheme

We consider a first-order scheme

$$w_{i}^{n+1} = w_{i}^{n} - \frac{\Delta t}{\Delta x} \left( F \left( w_{i}^{n}, w_{i+1}^{n} \right) - F \left( w_{i-1}^{n}, w_{i}^{n} \right) \right).$$

For a time step restricted according to the CFL condition

$$\frac{\Delta t}{\Delta x} \max_{i \in \mathbb{Z}} |\lambda^{\pm} \left( w_{i}^{n}, w_{i+1}^{n} \right)| \leq \frac{1}{2},$$

the first-order scheme is assumed to satisfy:

- **Robustness**: $\forall i \in \mathbb{Z}, \ w_{i}^{n} \in \Omega \Rightarrow \forall i \in \mathbb{Z}, \ w_{i}^{n+1} \in \Omega$
- **Stability**:

$$\rho_{i}^{n+1} r_{i}^{n+1} \leq \rho_{i}^{n} r_{i}^{n} - \frac{\Delta t}{\Delta x} \left( F^{\rho} \left( w_{i}^{n}, w_{i+1}^{n} \right) r_{i+1/2}^{n} \right. \left. - F^{\rho} \left( w_{i-1}^{n}, w_{i}^{n} \right) r_{i-1/2}^{n} \right).$$

Example: the HLLC/Suliciu relaxation scheme
Reconstruction procedure

- We consider high-order reconstructed states $w^n_{i,\pm}$ on the cell $K_i$ at the interfaces $x_{i\pm1/2}$.
- These reconstructed states can be obtained by any reconstruction procedure (MUSCL, ENO/WENO, PPM, DMGR...).
- Assumptions:
  - The reconstruction is $\Omega$-preserving: $w^n_{i,\pm} \in \Omega$;
  - The reconstruction is conservative:
    $$w^n_i = \frac{1}{2} \left( w^n_{i,-} + w^n_{i,+} \right).$$
The e-MOOD algorithm

1. **Reconstruction step:** For all $i \in \mathbb{Z}$, we evaluate high-order reconstructed states $w_i^{n,\pm}$ located at the interfaces $x_{i\pm1/2}$.

2. **Evolution step:** The solution is evolved as follows:

   $$w_i^{n+1,*} = w_i^n - \frac{\Delta t}{\Delta x} \left( F\left(w_i^{n,+}, w_{i+1}^{n,-}\right) - F\left(w_{i-1}^{n,+}, w_i^{n,-}\right) \right).$$

3. **A posteriori limitation step:** We have the following alternative:
   - if for all $i \in \mathbb{Z}$, we have
     $$\rho_i^{n+1,*} r_i^{n+1,*} \leq \rho_i^n r\left(w_i^n\right) - \frac{\Delta t}{\Delta x} \left(F^\rho\left(w_i^{n,+}, w_{i+1}^{n,-}\right) r_i^n r_{i+1/2}\right) - F^\rho\left(w_{i-1}^{n,+}, w_i^{n,-}\right) r_i^{n r_{i-1/2}}, \quad (1)$$
     then the solution is valid and the updated solution at time $t^n + \Delta t$ is defined by $w_i^{n+1} = w_i^{n+1,*}$;
   - otherwise, for all $i \in \mathbb{Z}$ such that (1) is not satisfied, we set $w_i^{n,\pm} = w_i^n$ and we go back to step 2.
Theorem

Assume the time step $\Delta t$ is chosen in order to satisfy the two following CFL like conditions:

$$\frac{\Delta t}{\Delta x} \max_{i \in \mathbb{Z}} \left( |\lambda^+ (w^n_i^+, w^n_{i+1})|, |\lambda^- (w^n_i^-, w^n_{i+1})| \right) \leq \frac{1}{4},$$

$$\frac{\Delta t}{\Delta x} \left( \max \left( 0, F^\rho_{i+1/2} \right) - \min \left( 0, F^\rho_{i-1/2} \right) \right) \leq \rho^n_i.$$

Then the updated states $w^n_{i+1}$, given by the e-MOOD scheme, belong to $\Omega$. Moreover, for all smooth increasing convex function $\psi$, the e-MOOD scheme satisfies

$$\frac{1}{\Delta t} \left( \rho^n_{i+1} \psi(r^n_{i+1}) - \rho^n_i \psi(r^n_i) \right) + \frac{1}{\Delta x} \left( F^\rho \left( w^n_i^+, w^n_{i+1} \right) \psi(r^n_{i+1/2}) - F^\rho \left( w^n_{i-1}, w^n_{i} \right) \psi(r^n_{i-1/2}) \right) \leq 0.$$

The e-MOOD scheme is thus entropy preserving.
Numerical results obtained with a second-order time scheme

$L^1$ error:

$$\sum_i \left| \rho_{i}^{N} - \rho_{ex}(x_i, T) \right|$$

1-rarefaction

Double shock
Well-balanced relaxation schemes

1. Second-order schemes based on dual mesh gradient reconstruction
   - MUSCL scheme and CFL condition
   - The DMGR scheme
   - Numerical results

2. A high-order entropy preserving scheme with \textit{a posteriori} limitation
   - Motivations
   - From one to all discrete entropy inequalities
   - The e-MOOD scheme for the Euler equations

3. Well-balanced schemes for systems with nonlinear source terms
   - The Ripa model
   - Relaxation models
   - The relaxation scheme
   - Numerical results
The Ripa model

\[
\begin{align*}
\partial_t h + \partial_x hu &= 0 \\
\partial_t hu + \partial_x (hu^2 + gh^2 \theta/2) &= -gh\theta \partial_x Z \\
\partial_th\theta + \partial_x h\theta u &= 0
\end{align*}
\]

- $h$: water height
- $u$: velocity
- $\theta$: temperature
- $g$: gravity constant
- $Z(x)$: smooth topography function
The Ripa model

\[
\begin{align*}
\partial_t h + \partial_x hu &= 0 \\
\partial_t hu + \partial_x (hu^2 + gh^2\theta/2) &= -gh\theta \partial_x Z \\
\partial_t h\theta + \partial_x h\theta u &= 0
\end{align*}
\]

- We define
  - the vector of conservative variables \( w = (h, hu, h\theta)^T \),
  - the flux function \( f(w) = (hu, hu^2 + gh^2\theta/2, h\theta u)^T \),
  - the source term \( s(w) = (0, -gh\theta, 0)^T \),

to rewrite the system into the compact form

\[
\partial_t w + \partial_x f(w) = s(w) \partial_x Z.
\]

- The set of physical admissible states is

\[
\Omega = \left\{ w \in \mathbb{R}^3, \ h > 0, \ \theta > 0 \right\}.
\]
The Ripa model

\[
\begin{align*}
\partial_t h + \partial_x hu &= 0 \\
\partial_t hu + \partial_x (hu^2 + gh^2 \theta/2) &= -gh\theta \partial_x Z \\
\partial_t h\theta + \partial_x h\theta u &= 0
\end{align*}
\]

**Steady states**

The steady states at rest are governed by the ODE

\[
\begin{align*}
u &\equiv 0, \\
\partial_x (h^2 \theta/2) &= -h\theta \partial_x Z.
\end{align*}
\]

We cannot obtain an explicit expression of all the steady states.

**Lake at rest type solutions**

\[
\begin{align*}
u &= 0, &\quad u &= 0, &\quad u &= 0, \\
\theta &= \text{cst}, &\quad z &= \text{cst}, &\quad h &= \text{cst}, \\
h + Z &= \text{cst}, &\quad h^2 \theta &= \text{cst}, &\quad z + \frac{h}{2} \ln \theta &= \text{cst}.
\end{align*}
\]
The relaxation method without source term

- Initial system:
  \[ \partial_t w + \partial_x f(w) = 0. \] (1)

- Relaxation system:
  \[ \partial_t W + \partial_x F(W) = \frac{1}{\varepsilon} R(W), \] (2)

  - (2) should formally give back (1) when \( \varepsilon \to 0 \).
  - (2) should be “simpler” than (1) (e.g. only linearly degenerate fields)

- The relaxation scheme is based on a splitting strategy:
  
  **Time evolution:** We evolve the initial data by the Godunov scheme for the system \( \partial_t W + \partial_x F(W) = 0 \) (i.e. \( \varepsilon = +\infty \)).

  **Relaxation:** We take into account the relaxation source term by solving \( \partial_t W = \frac{1}{\varepsilon} R(W) \) then taking the limit for \( \varepsilon \to 0 \).
The Suliciu model ([Suliciu ’98], [Bouchut ’04]...)

\[
\begin{align*}
\partial_t h + \partial_x hu &= 0 \\
\partial_t hu + \partial_x (hu^2 + \pi) &= -gh\theta \partial_x Z \\
\partial_t h\theta + \partial_x h\theta u &= 0 \\
\partial_t h\pi + \partial_x (u(h\pi + \nu^2)) &= \frac{h}{\varepsilon} (gh^2\theta/2 - \pi) \\
\partial_t Z &= 0
\end{align*}
\]

<table>
<thead>
<tr>
<th>Eigenvalues</th>
<th>Riemann invariants</th>
</tr>
</thead>
<tbody>
<tr>
<td>(u \pm \frac{\nu}{h})</td>
<td>(u \pm \frac{\nu}{h}, \ \pi \mp \nu u, \ \theta, \ Z)</td>
</tr>
<tr>
<td>(u \times 2)</td>
<td>(u, \ \pi, \ Z)</td>
</tr>
<tr>
<td>0</td>
<td>(hu, \ \pi + \frac{\nu^2}{h}, \ \theta, \ g\theta Z + \frac{u^2}{2} - \frac{\nu^2}{2h^2})</td>
</tr>
</tbody>
</table>

**Difficulties** to compute the solution of the Riemann problem:

- The order of the eigenvalues is not determined \textit{a priori}.
- There are strong nonlinearities in the Riemann invariants for the eigenvalue 0.
Relaxation model with moving topography

\[
\begin{align*}
\partial_t h + \partial_x h u &= 0 \\
\partial_t h u + \partial_x (h u^2 + \pi) &= -gh\theta \partial_x a \\
\partial_t h\theta + \partial_x h\theta u &= 0 \\
\partial_t h\pi + \partial_x (u(h\pi + \nu^2)) &= \frac{h}{\varepsilon} (gh^2\theta/2 - \pi) \\
\partial_t a + u\partial_x a &= \frac{1}{\varepsilon} (Z - a)
\end{align*}
\]

<table>
<thead>
<tr>
<th>Eigenvalues</th>
<th>Riemann invariants</th>
</tr>
</thead>
<tbody>
<tr>
<td>( u \pm \frac{\nu}{h} )</td>
<td>( u \pm \frac{\nu}{h}, \quad \pi \mp \nu u, \quad \theta, \quad a )</td>
</tr>
<tr>
<td>( u \times 3 )</td>
<td>( u )</td>
</tr>
</tbody>
</table>

- The order of the eigenvalues is fixed: \( u - \frac{\nu}{h} < u < u + \frac{\nu}{h} \)
- There is a missing invariant for the eigenvalue \( u \)
  \( \Rightarrow \) we need a closure equation
The Riemann problem

\[ u_L - \frac{\nu}{h_L} \rightarrow u^* \rightarrow u_R + \frac{\nu}{h_R} \]

- 5 unknowns: \( u^*, h^*_L, \pi^*_L, h^*_R, \pi^*_R \)
- 4 equations given by the Riemann invariants
  \[ u \pm \frac{\nu}{h}, \ \pi \mp \nu u \]

Closure equation:
\[ \pi^*_R - \pi^*_L = -g\bar{h}(w_L, w_R)\bar{\theta}(w_L, w_R)(a_R - a_L) \]
- \( \bar{h}(w_L, w_R) \): \( h \)-average s.t. \( \bar{h}(w, w) = h \) and \( \bar{h}(w_L, w_R) = \bar{h}(w_R, w_L) \)
- \( \bar{\theta}(w_L, w_R) \): \( \theta \)-average s.t. \( \bar{\theta}(w, w) = \theta \) and \( \bar{\theta}(w_L, w_R) = \bar{\theta}(w_R, w_L) \)

Solution of the Riemann problem

\[
\begin{align*}
  u^* &= \frac{u_L + u_R}{2} - \frac{\pi_R - \pi_L}{2\nu} - \frac{g}{2\nu} \bar{h}(w_L, w_R)\bar{\theta}(w_L, w_R)(a_R - a_L) \\
  \pi^*_L &= \pi_L + \nu(u_L - u^*) \\
  \frac{1}{h^*_L} &= \frac{1}{h_L} + \frac{u^* - u_L}{\nu} \\
  \pi^*_R &= \pi_R + \nu(u^* - u_R) \\
  \frac{1}{h^*_R} &= \frac{1}{h_R} + \frac{u_R - u^*}{\nu}
\end{align*}
\]
Well-balanced relaxation schemes

Reformulation into a fully determined model

\[
\begin{align*}
\partial_t h + \partial_x hu &= 0 \\
\partial_t hu + \partial_x (hu^2 + \pi) &= -g \bar{h}(X^-, X^+) \bar{\theta}(X^-, X^+) \partial_x a \\
\partial_t h\theta + \partial_x h\theta u &= 0 \\
\partial_t h\pi + \partial_x (u(h\pi + \nu^2)) &= \frac{h}{\varepsilon}(gh^2\theta/2 - \pi) \\
\partial_t a + u\partial_x a &= \frac{1}{\varepsilon}(Z - a) \\
\partial_t X^- + (u - \delta)\partial_x X^- &= \frac{1}{\varepsilon}(W - X^-) \\
\partial_t X^+ + (u + \delta)\partial_x X^+ &= \frac{1}{\varepsilon}(W - X^+) 
\end{align*}
\]

<table>
<thead>
<tr>
<th>Eigenvalues</th>
<th>Riemann invariants</th>
</tr>
</thead>
<tbody>
<tr>
<td>(u \pm \frac{\nu}{h})</td>
<td>(u \pm \frac{\nu}{h}, \pi \mp \nu u, \theta, a, X^-, X^+)</td>
</tr>
<tr>
<td>(u \times 3)</td>
<td>(u, \pi + g \bar{h}(X^-, X^+) \bar{\theta}(X^-, X^+) a, X^-, X^+)</td>
</tr>
<tr>
<td>(u - \delta)</td>
<td>(h, u, \theta, \pi, a, X^+)</td>
</tr>
<tr>
<td>(u + \delta)</td>
<td>(h, u, \theta, \pi, a, X^-)</td>
</tr>
</tbody>
</table>

Both models lead to the same numerical scheme.
The relaxation scheme

The relaxation scheme associated with both previous models writes:

\[
  w_i^{n+1} = w_i^n - \frac{\Delta t}{\Delta x} \left( f(w_i^n, w_{i+1}^n) - f(w_i^{n-1}, w_i^n) \right) \\
  + \frac{\Delta t}{2} \left( s^+(w_{i-1}^n, w_i^n) \frac{Z_i - Z_{i-1}}{\Delta x} + s^-(w_i^n, w_{i+1}^n) \frac{Z_{i+1} - Z_i}{\Delta x} \right),
\]

where the numerical flux is defined by

\[
  f(w_L, w_R) = \begin{cases} 
    (h_L u_L, \ h_L u_L^2 + gh_L^2 \theta_L/2, \ h_L \theta_L u_L)^T & \text{if } u_L - \frac{\nu}{h_L} > 0, \\
    (h_L^* u^*, \ h_L^*(u^*)^2 + \pi_L^*, \ h_L^* \theta_L u^*)^T & \text{if } u_L - \frac{\nu}{h_L} < 0 < u^*, \\
    (h_R^* u^*, \ h_R^*(u^*)^2 + \pi_R^*, \ h_R^* \theta_R u^*)^T & \text{if } u^* < 0 < u_R + \frac{\nu}{h_R}, \\
    (h_R u_R, \ h_R u_R^2 + gh_R^2 \theta_R/2, \ h_R \theta_R u_R)^T & \text{if } u_R + \frac{\nu}{h_R} < 0,
  \end{cases}
\]

and the numerical source terms are defined by

\[
  s^+(w_L, w_R) = \left( 0, - (\text{sgn}(u^*) + 1) g \bar{h}(w_L, w_R) \bar{\theta}(w_L, w_R), 0 \right)^T, \\
  s^-(w_L, w_R) = \left( 0, - (1 - \text{sgn}(u^*)) g \bar{h}(w_L, w_R) \bar{\theta}(w_L, w_R), 0 \right)^T.
\]
Properties of the relaxation scheme (1)

Theorem (Exact preservation of lake at rest solutions)

Assume the average functions \( \bar{h} \) and \( \bar{\theta} \) are defined by

\[
\bar{h}(w_L, w_R) = \frac{1}{2}(h_L + h_R), \quad \bar{\theta}(w_L, w_R) = \begin{cases} 
\frac{\theta_R - \theta_L}{\ln(\theta_R) - \ln(\theta_L)} & \text{if } \theta_L \neq \theta_R, \\
\theta_L & \text{if } \theta_L = \theta_R.
\end{cases}
\]

Then the relaxation scheme preserves exactly the three lake at rest solutions: if the initial data \( w_i^0 \) is given by

\[
\begin{align*}
&\begin{cases}
  u_i^0 = 0, \\
  \theta_i^0 = \theta, \\
  h_i^0 + z_i = H,
\end{cases} \quad \text{or} \quad
\begin{cases}
  u_i^0 = 0, \\
  z_i = Z, \\
  (h_i^0)^2 \theta_i^0 = P,
\end{cases} \quad \text{or} \quad
\begin{cases}
  u_i^0 = 0, \\
  h_i^0 = H, \\
  z_i + h_i^0 \ln(\theta_i^0)/2 = P,
\end{cases}
\end{align*}
\]
where \( \theta > 0, H > 0 \) and \( P > 0 \) are constants, then the approximate solution \( w_i^n \) stays at rest:

\[
w_i^n = w_i^0, \quad \forall i \in \mathbb{Z}, \quad \forall n \in \mathbb{N}.
\]
Properties of the relaxation scheme (2)

**Theorem (Well-balancedness)**

Assume the initial data $w_i^0$ satisfies for all $i \in \mathbb{Z}$:

\[
\begin{cases}
    u_i^0 = 0, \\
    (h_i^0)^2 \theta_i^0 / 2 - (h_i^0)^2 \theta_i^0 / 2 = -\bar{h}(w_i^0, w_{i+1}^0) \bar{\theta}(w_i^0, w_{i+1}^0)(Z_{i+1} - Z_i).
\end{cases}
\]

Then the approximate solution stays at rest: $w_i^n = w_i^0, \forall i \in \mathbb{Z}, \forall n \in \mathbb{N}$.

**Theorem (Robustness)**

Assume the parameter $\nu$ satisfies the following inequalities:

\[
    u_L - \frac{\nu}{h_L} < u^* < u_R + \frac{\nu}{h_R}.
\]

Assume the following CFL condition is satisfied:

\[
    \frac{\Delta t}{\Delta x} \max_{i \in \mathbb{Z}} |u_i^n \pm \nu / h_i^n| \leq \frac{1}{2}
\]

Then the relaxation scheme preserves the set of physical states:

$\forall i \in \mathbb{Z}, h_i^n > 0$ and $\theta_i^n > 0$ \implies $\forall i \in \mathbb{Z}, h_i^{n+1} > 0$ and $\theta_i^{n+1} > 0$. 
Well-balanced relaxation schemes

Numerical results

Dam break over a non-flat bottom
[Chertock, Kurganov & Liu ’13]

![Graph showing numerical results for dam break over a non-flat bottom](image_url)
Perturbation of a nonlinear steady state

- **Topography:**
  \[ Z(x) = -2e^x \]

- **Steady state solution:**
  \[ (h_s, u_s, \theta_s)(x) = (e^x, 0, e^{2x}) \]

- **Initial perturbation:**
  \[ \delta h(x,0) = 0.1 \chi_{[-0.1,0]}(x) \]
Perspectives

- Extension of the DMGR scheme to higher-order.
- Combination of the DMGR and e-MOOD methods to get high-order entropic schemes in 2D.
- Extension of the e-MOOD scheme to other systems.
- Development of high-order well-balanced schemes for systems with source terms.
- Does the relaxation schemes for Ripa and Euler with gravity satisfy discrete entropy inequalities?