

Implicit kinetic schemes for the shallow water system

Chourouk El Hassanieh¹, Mathieu Rigal², Jacques Sainte-Marie¹

¹Laboratoire Jacques-Louis Lions, Sorbonne Université

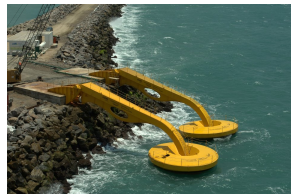
²Institut de Mathématiques de Bordeaux

Bordeaux, November 14, 2023

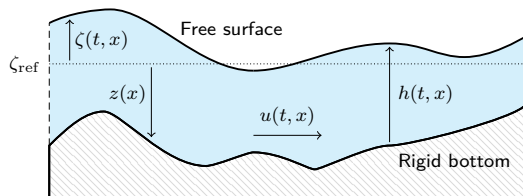


Why are we interested in geophysical flows?

- water management
- natural disasters
- energy production



→ numerical schemes, simulations



Quantities of interest:

h	→	water height
u	→	horizontal velocity
hu	→	horizontal discharge
z	→	bathymetry

Free surface flows \Rightarrow evolving fluid geometry

Shallow water equations: vertically averaged model (reduced complexity)

Simplifying assumptions

- shallow flow
- velocity has small variations along the vertical
- no plunging wave

1D shallow water system:

$$\begin{cases} \partial_t h + \partial_x hu = 0 \\ \partial_t hu + \partial_x (hu^2 + \frac{g}{2}h^2) = -gh\partial_x z \end{cases} \quad \text{in } \mathbb{R} \quad (\text{SV})$$

Convenient vector notation $\partial_t U + \partial_x F(U) = S(U, z)$ with $U = (h, hu)^T$.

Important properties at the continuous level:

- Positivity ($h \geq 0 \forall t$)
- Stationary state $h + z \equiv \text{Cst}$, $u \equiv 0$
- Entropy inequality $\partial_t \eta(U, z) + \partial_x G(U, z) \leq 0$

$$\eta(U, z) = \frac{hu^2}{2} + \frac{gh^2}{2} + ghz, \quad G(U, z) = \left(\eta(U, z) + \frac{gh^2}{2} \right) u$$

1D shallow water system:

$$\begin{cases} \partial_t h + \partial_x hu = 0 \\ \partial_t hu + \partial_x (hu^2 + \frac{g}{2}h^2) = -gh\partial_x z \end{cases} \quad \text{in } \mathbb{R} \quad (\text{SV})$$

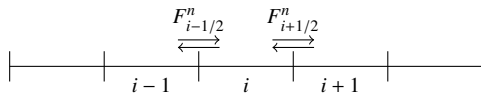
Convenient vector notation $\partial_t U + \partial_x F(U) = S(U, z)$ with $U = (h, hu)^T$.

Important properties at the continuous level:

- Positivity ($h \geq 0 \forall t$)
- Stationary state $h + z \equiv \text{Cst}$, $u \equiv 0$
- Entropy inequality $\partial_t \eta(U, z) + \partial_x G(U, z) \leq 0$

$$\eta(U, z) = \frac{hu^2}{2} + \frac{gh^2}{2} + ghz, \quad G(U, z) = \left(\eta(U, z) + \frac{gh^2}{2} \right) u$$

Finite volumes: a simple example



Finite volume scheme of the form

$$\begin{cases} \frac{U_i^{n+1} - U_i^n}{\Delta t} + \frac{1}{\Delta x} (F_{i+1/2}^n - F_{i-1/2}^n) = S_i^n \\ F_{i+1/2}^n = \mathcal{F}(U_i^n, z_i, U_{i+1}^n, z_{i+1}) \end{cases} \quad (1)$$

For instance, Rusanov flux + centered source

$$F_{i+1/2}^n = \frac{1}{2} \left(F(U_i^n, z_i) + F(U_{i+1}^n, z_{i+1}) \right) - \frac{a}{2} (U_{i+1}^n - U_i^n), \quad a > 0$$

$$S_i^n = -gh_i^n \frac{z_{i+1} - z_{i-1}}{2\Delta x} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Problem n°1: if it exists, find $G_{i+1/2}^n$ numerical entropy flux such that

$$\text{Update (1)} \implies \begin{cases} \frac{\eta(U_i^{n+1}, z_i) - \eta(U_i^n, z_i)}{\Delta t} + \frac{1}{\Delta x} (G_{i+1/2}^n - G_{i-1/2}^n) \leq 0 \\ G_{i+1/2}^n = \mathcal{G}(U_i^n, z_i, U_{i+1}^n, z_{i+1}) \end{cases}$$

Problem n°2: preserve lakes at rest ($h + z = 0$, $u = 0$)

Steady state $\partial_t U = 0$ implies $\partial_x F(U) = S(U, z)$, whereas at discrete level:

$$\frac{1}{\Delta x} (F_{i+1/2}^n - F_{i-1/2}^n) = \begin{pmatrix} \frac{a}{2} \frac{z_{i+1} - 2z_i + z_{i-1}}{\Delta x} \\ \frac{g}{2\Delta x} ((z_{i+1})^2 - (z_{i-1})^2) \end{pmatrix} \neq \begin{pmatrix} 0 \\ gz_i \frac{z_{i+1} - z_{i-1}}{2\Delta x} \end{pmatrix},$$

and therefore $\frac{U_i^{n+1} - U_i^n}{\Delta t} \neq 0$.

Solution proposed by Audusse et. al (2016)

- Explicit kinetic scheme preserving lakes at rest...
- but satisfying a discrete entropy inequality with error term

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

where in some cases $D_i^n > 0$

Our goal is to implicit this scheme to improve its stability

Outline of the talk:

- 1 Brief recall of the kinetic formalism
- 2 The case of a flat topography
- 3 The case of a varying topography

Kinetic equation with BGK collision operator

$$\partial_t f + \xi \partial_x f - g(\partial_x z) \partial_\xi f = \frac{1}{\epsilon} (M(U, \xi) - f) \quad (\text{BGK})$$

- $f(t, x, \xi) \geq 0$ density of particles with velocity ξ
- Moment relations $\int (1, \xi, \xi^2)^T M(U, \xi) d\xi = (h, hu, hu^2 + gh^2/2)^T$
- In the limit $\epsilon \rightarrow 0$, we formally have $f \rightarrow M$

Proposition 1 (Audusse, Bouchut, Bristeau, and Sainte-Marie 2016)

If the bathymetry $z(x)$ is Lipschitz continuous, then U is solution of the shallow water system iff $M(U, \xi)$ satisfies the kinetic equation

$$\partial_t M + \xi \partial_x M - g(\partial_x z) \partial_\xi M = Q \quad (2)$$

for some collision term $Q(t, x, \xi)$ that satisfies $\int_{\mathbb{R}} (1, \xi)^T Q d\xi = 0$ for a.e. (t, x) .

Definition 1 (Kinetic entropy H)

$H(f, \xi)$ convex in f and satisfying

$$\int_{\mathbb{R}} H(M(U, \xi), \xi) d\xi = \eta(U), \quad \int_{\mathbb{R}} H(M(U_f, \xi), \xi) d\xi \leq \int_{\mathbb{R}} H(f, \xi) d\xi \quad \forall f$$

If flat bottom ($z \equiv \text{Const}$), integrate (BGK) against $\partial_1 H(f, \xi)$ to get

$$\underbrace{\partial_t \int_{\mathbb{R}} H(f, \xi) d\xi}_{\xrightarrow{\varepsilon \rightarrow 0} \partial_t \eta(U_f)} + \underbrace{\partial_x \int_{\mathbb{R}} \xi H(f, \xi) d\xi}_{\xrightarrow{\varepsilon \rightarrow 0} \partial_x G(U_f)} = \frac{1}{\varepsilon} \underbrace{\int_{\mathbb{R}} \partial_1 H(f, \xi) (M(U_f, \xi) - f) d\xi}_{\leq \int_{\mathbb{R}} H(M, \xi) - H(f, \xi) d\xi \leq 0}$$

Extends to varying bottoms if

$$\int_{\mathbb{R}} \partial_3 H(f, z, \xi) d\xi = hu,$$

which implies

$$\lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}} \partial_1 H(f, z, \xi) (\xi \partial_x f - g(\partial_x z) \partial_\xi f) d\xi = \partial_x G(U, z)$$

Kinetic representation of the shallow water system

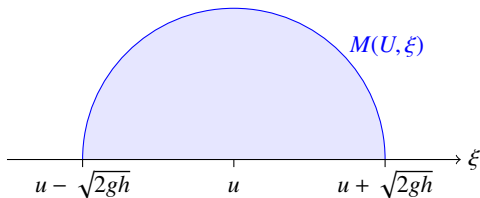
Given a convex H , determine $M(U, \cdot)$ by minimizing

$$f \mapsto \int_{\mathbb{R}} H(f, \xi) d\xi \quad \text{constrained by} \quad \int_{\mathbb{R}} (1, \xi)^T f d\xi = U$$

François Bouchut. "Construction of BGK Models with a Family of Kinetic Entropies for a Given System of Conservation Laws." (1999)

Lemma 1 (Perthame and Simeoni 2001)

$$H(f, z, \xi) = \frac{\xi^2}{2} f + \frac{g^2 \pi^2}{6} f^3 + gz f \text{ is a kinetic entropy for } M(U, \xi) = \frac{1}{g\pi} \sqrt{(2gh - (\xi - u)^2)_+}.$$



Explicit time discretization involving BGK splitting

$$\begin{cases} \frac{f^{n+1/2} - f^n}{\Delta t} = \frac{1}{\varepsilon} (M(U_f^n, \xi) - f^{n+1/2}) & \text{collision step} \\ \frac{f^{n+1} - f^{n+1/2}}{\Delta t} + \xi \partial_x f^{n+1/2} = 0 & \text{transport step} \end{cases}$$

Explicit first order upwind scheme when $\varepsilon \rightarrow 0$

$$\frac{f_i^{n+1} - M_i^n}{\Delta t} + \frac{\xi}{\Delta x} \left(\mathbb{1}_{\xi < 0} (M_{i+1}^n - M_i^n) + \mathbb{1}_{\xi > 0} (M_i^n - M_{i-1}^n) \right) = 0 \quad (3)$$

Macroscopic rewriting by integrating (3) against $(1, \xi)^T$

$$\frac{U_i^{n+1} - U_i^n}{\Delta t} + \frac{1}{\Delta x} \left(F(U_i^n, U_{i+1}^n) - F(U_{i-1}^n, U_i^n) \right) = 0$$

Kinetic numerical flux $F(U_L, U_R) = \int_{\xi < 0} \xi \left(\frac{1}{\xi} \right) M(U_R, \xi) d\xi + \int_{\xi > 0} \xi \left(\frac{1}{\xi} \right) M(U_L, \xi) d\xi$.

Explicit time discretization involving BGK splitting

$$\begin{cases} \frac{f^{n+1/2} - f^n}{\Delta t} = \frac{1}{\varepsilon} (M(U_f^n, \xi) - f^{n+1/2}) & \text{collision step} \\ \frac{f^{n+1} - f^{n+1/2}}{\Delta t} + \xi \partial_x f^{n+1/2} = 0 & \text{transport step} \end{cases}$$

Explicit first order upwind scheme when $\varepsilon \rightarrow 0$

$$\frac{f_i^{n+1} - M_i^n}{\Delta t} + \frac{\xi}{\Delta x} \left(\mathbb{1}_{\xi < 0} (M_{i+1}^n - M_i^n) + \mathbb{1}_{\xi > 0} (M_i^n - M_{i-1}^n) \right) = 0 \quad (3)$$

Macroscopic rewriting by integrating (3) against $(1, \xi)^T$

$$\frac{U_i^{n+1} - U_i^n}{\Delta t} + \frac{1}{\Delta x} \left(F(U_i^n, U_{i+1}^n) - F(U_{i-1}^n, U_i^n) \right) = 0$$

Kinetic numerical flux $F(U_L, U_R) = \int_{\xi < 0} \xi \left(\frac{1}{\xi} \right) M(U_R, \xi) d\xi + \int_{\xi > 0} \xi \left(\frac{1}{\xi} \right) M(U_L, \xi) d\xi$.

The case of a flat topography

Do we satisfy a discrete counterpart to $\partial_t \eta + \partial_x G \leq 0$?

Proposition 2 (Audusse, Bouchut, Bristeau, and Sainte-Marie 2016)

If the CFL $\frac{\Delta t}{\Delta x} |\xi| \leq 1$ holds for any $\xi \in \text{supp } M^n$, then the explicit kinetic scheme (3) satisfies

$$h_i^{n+1} \geq 0 \quad \text{together with} \quad \frac{\eta(U_i^{n+1}) - \eta(U_i^n)}{\Delta t} + \frac{1}{\Delta x} (G_{i+1/2}^n - G_{i-1/2}^n) \leq 0$$

Proof: set $\sigma = \frac{\Delta t}{\Delta x}$ and rewrite (3) as

$$f_i^{n+1} = (1 - \sigma|\xi|)M_i^n + \sigma|\xi|M_{i\pm 1}^n \geq 0$$

$$\text{Also } \eta_i^{n+1} = \int_{\mathbb{R}} H(M_i^{n+1}, \xi) d\xi \leq \int_{\mathbb{R}} H(f_i^{n+1}, \xi) d\xi \leq \underbrace{\int_{\mathbb{R}} (1 - \sigma|\xi|)H_i^n + \sigma|\xi|H_{i\pm 1}^n d\xi}_{\eta_i^n - \sigma(G_{i+1/2}^n - G_{i-1/2}^n)}$$

The case of a flat topography

We study the implicit version of the previous scheme.

$$\frac{f_i^{n+1} - M_i^n}{\Delta t} + \frac{\xi}{\Delta x} \left(\mathbb{1}_{\xi < 0} (f_{i+1}^{n+1} - f_i^{n+1}) + \mathbb{1}_{\xi > 0} (f_i^{n+1} - f_{i-1}^{n+1}) \right) = 0 \quad (4)$$

Solve the system $(\mathbf{I} + \sigma \mathbf{L}) f^{n+1} = M^n + \sigma B^{n+1}$ with $\sigma = \Delta t / \Delta x$ and

$$\mathbf{L} = |\xi| \begin{pmatrix} 1 & -\mathbb{1}_{\xi < 0} & & & 0 \\ -\mathbb{1}_{\xi > 0} & 1 & -\mathbb{1}_{\xi < 0} & & \\ & \ddots & \ddots & \ddots & \\ & & -\mathbb{1}_{\xi > 0} & 1 & -\mathbb{1}_{\xi < 0} \\ 0 & & & -\mathbb{1}_{\xi > 0} & 1 \end{pmatrix}_{N \times N}, \quad B^{n+1} = |\xi| \begin{pmatrix} M_0^{n+1} \mathbb{1}_{\xi > 0} \\ 0 \\ \vdots \\ 0 \\ M_{N+1}^{n+1} \mathbb{1}_{\xi < 0} \end{pmatrix}_N$$

In practice, ghost cell contribution B^{n+1} unknown \rightarrow substitute it by B^n .

Proposition 3 (El Hassanieh, R., Sainte-Marie)

The implicit kinetic scheme (4) is well defined, its update can be computed analytically and it enjoys the same properties as the explicit scheme $\forall \Delta t > 0$.

The case of a flat topography

We study the implicit version of the previous scheme.

$$\frac{f_i^{n+1} - M_i^n}{\Delta t} + \frac{\xi}{\Delta x} \left(\mathbb{1}_{\xi < 0} (f_{i+1}^{n+1} - f_i^{n+1}) + \mathbb{1}_{\xi > 0} (f_i^{n+1} - f_{i-1}^{n+1}) \right) = 0 \quad (4)$$

Solve the system $(\mathbf{I} + \sigma \mathbf{L}) f^{n+1} = M^n + \sigma B^{n+1}$ with $\sigma = \Delta t / \Delta x$ and

$$\mathbf{L} = |\xi| \begin{pmatrix} 1 & -\mathbb{1}_{\xi < 0} & & & 0 \\ -\mathbb{1}_{\xi > 0} & 1 & -\mathbb{1}_{\xi < 0} & & \\ & \ddots & \ddots & \ddots & \\ & & -\mathbb{1}_{\xi > 0} & 1 & -\mathbb{1}_{\xi < 0} \\ 0 & & & -\mathbb{1}_{\xi > 0} & 1 \end{pmatrix}_{N \times N}, \quad B^{n+1} = |\xi| \begin{pmatrix} M_0^{n+1} \mathbb{1}_{\xi > 0} \\ 0 \\ \vdots \\ 0 \\ M_{N+1}^{n+1} \mathbb{1}_{\xi < 0} \end{pmatrix}_N$$

In practice, ghost cell contribution B^{n+1} unknown \rightarrow substitute it by B^n .

Proposition 3 (El Hassanieh, R., Sainte-Marie)

The implicit kinetic scheme (4) is well defined, its update can be computed analytically and it enjoys the same properties as the explicit scheme $\forall \Delta t > 0$.

The case of a flat topography

$$\frac{f_i^{n+1} - M_i^n}{\Delta t} + \frac{\xi}{\Delta x} (f_{i+1/2}^{n+1} - f_{i-1/2}^{n+1}) = 0 \iff (\mathbf{I} + \sigma \mathbf{L}) f^{n+1} = M^n + \sigma B^n \quad (4)$$

Sketch of the proof:

Well-defined: The mass matrix has a strictly dominant diagonal \Rightarrow invertible

Positivity: The mass matrix is monotone and RHS is positive

Analytic expression: Decompose $\mathbf{L} = |\xi| \mathbf{I} - \mathbf{N}$ so that

$$(\mathbf{I} + \sigma \mathbf{L})^{-1} = \frac{1}{1 + \sigma |\xi|} \left(\mathbf{I} - \frac{\sigma}{1 + \sigma |\xi|} \mathbf{N} \right)^{-1} = \frac{1}{1 + \sigma |\xi|} \sum_{k=0}^N \left(\frac{\sigma}{1 + \sigma |\xi|} \mathbf{N} \right)^k$$

Entropy inequality: Multiply (4) by $\partial_1 H(f_i^{n+1}, \xi)$ and use

$$\partial_1 H(b, \xi)(b - a) = H(b, \xi) - H(a, \xi) + \frac{g^2 \pi^2}{6} (2b + a)(b - a)^2$$

to obtain

$$\frac{H(f_i^{n+1}) - H(M_i^n)}{\Delta t} + \frac{\xi}{\Delta x} (H_{i+1/2}^{n+1} - H_{i-1/2}^{n+1}) = \widetilde{D}_i(\xi) \leq 0$$

The case of a flat topography

In practice, cannot obtain explicit expression for $\int_{\mathbb{R}} \begin{pmatrix} 1 \\ \xi \end{pmatrix} (\mathbf{I} + \sigma \mathbf{L})^{-1} M \, d\xi$ with

$$M(U, \xi) = \frac{1}{g\pi} \sqrt{(2gh - (\xi - u)^2)_+}$$

Substitute M with a simpler Maxwellian satisfying the moment relations

$$\tilde{M}(U, \xi) = \frac{h}{2\sqrt{3}c} \mathbb{1}_{|\xi-u| \leq \sqrt{3}c}, \quad c = \sqrt{\frac{gh}{2}}$$

- nonlinear implicit update **can be rewritten explicitly**
- counterpart: unlike M , \tilde{M} doesn't minimize $\int_{\mathbb{R}} H(\cdot, \xi) \, d\xi$
- as a consequence, no proof of discrete entropy inequality...
- ... but in practice, it seems to dissipate energy (numerical validation)

Neglecting the ghost cells, the implicit kinetic update writes

$$h^{n+1} = K((Ah) + (Bh)) \sqrt{h^n}, \quad hu^{n+1} = K'((Bhu) - (Ahu)) \sqrt{h^n}$$

For instance, matrix (Ah) is given by

$$\begin{pmatrix} [z]_{-\min(0,b_1)\sigma}^{-\min(0,a_1)\sigma} & [z-y]_{-\min(0,b_2)\sigma}^{-\min(0,a_2)\sigma} & \dots & \dots & [z - \sum_{l=1}^{N-1} y^l/l]_{-\min(0,b_N)\sigma}^{-\min(0,a_N)\sigma} \\ 0 & [z]_{-\min(0,b_2)\sigma}^{-\min(0,a_2)\sigma} & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & [z-y]_{-\min(0,b_N)\sigma}^{-\min(0,a_N)\sigma} \\ 0 & \dots & \dots & 0 & [z]_{-\min(0,b_N)\sigma}^{-\min(0,a_N)\sigma} \end{pmatrix}$$

where $y = x/(1+x)$, $z = \ln|1+x|$, $a_j = u_j^n - \sqrt{3}c_j^n$ and $b_j = u_j^n + \sqrt{3}c_j^n$

Neglecting the ghost cells, the implicit kinetic update writes

$$h^{n+1} = K((Ah) + (Bh)) \sqrt{h^n}, \quad hu^{n+1} = K'((Ahu) - (Bhu)) \sqrt{h^n}$$

Matrices (Ah) , (Bh) , (Ahu) , (Bhu) have $N(N + 1)/2$ nonzero coefficients

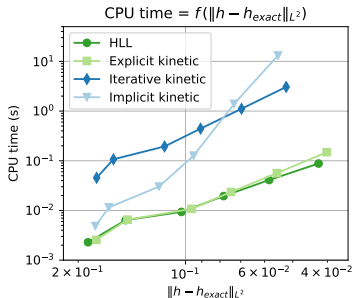
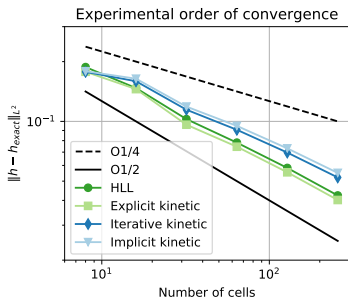
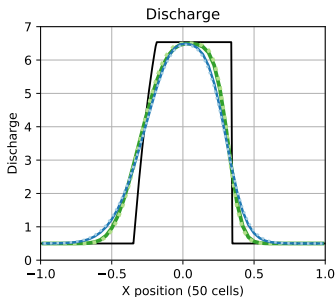
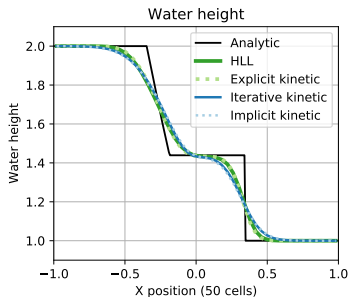
- matrix vector product has complexity $O(N^2)$ (cannot do better)
- up to $O(N)$ steps for each coefficient \Rightarrow matrix assembly in $O(N^3)$

Optimization: assemble matrices in specific order

- each coefficient computed in $O(1)$ steps from the previous one
- cost of matrix assembly reduced to $O(N^2)$

\rightarrow Fully vectorized implementation in Python

Implicit kinetic scheme: Riemann problem



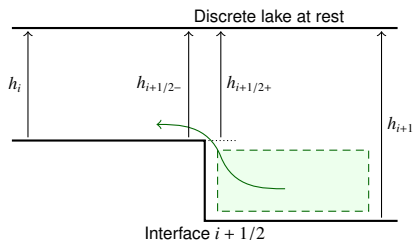
The case of a varying topography: hydrostatic reconstruction

Discretize source term in (SV)
$$\begin{cases} \partial_t h + \partial_x hu = 0 \\ \partial_t hu + \partial_x (hu^2 + \frac{g}{2}h^2) = -gh\partial_x z \end{cases}$$

Problem: how to preserve lakes at rest $h + z \equiv \text{Cst}$, $u \equiv 0$?

- Upwinding introduces diffusion on $h \Rightarrow h^{n+1} \neq h^n$
- Pressure variation should balance with source: $\partial_x \left(\frac{g}{2} h^2 \right) = -gh\partial_x z$

Hydrostatic reconstruction



$$z_{i+1/2} = \max(z_i, z_{i+1})$$

$$h_{i+1/2-} = (h_i + z_i - z_{i+1/2})_+$$

$$h_{i+1/2+} = (h_{i+1} + z_{i+1} - z_{i+1/2})_+$$

Audusse, Bouchut, Bristeau, Klein, et al. 2004 "A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows."

The case of a varying topography: hydrostatic reconstruction

Numerical flux and source term using reconstructed values

$$\begin{cases} \widetilde{F}_{i+1/2} = F(U_{i+1/2-}, U_{i+1/2+}) \\ \widetilde{F}_{i-1/2} = F(U_{i-1/2-}, U_{i-1/2+}) \end{cases}, \quad \widetilde{S}_i = \frac{g}{2\Delta x} (h_{i+1/2-}^2 - h_{i-1/2+}^2) \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

- \widetilde{S}_i is indeed consistent with the source term

$$\frac{1}{2\Delta x} (h_{i+1/2-}^2 - h_{i-1/2+}^2) = \frac{1}{2} \underbrace{(h_{i+1/2-} + h_{i-1/2+})}_{h_i + \mathcal{O}(\Delta z_i)} \times \frac{1}{\Delta x} \underbrace{(h_{i+1/2-} - h_{i-1/2+})}_{-(z_{i+1/2} - z_{i-1/2})/\Delta x} = -h \partial_x z + \mathcal{O}(\Delta x)$$

- If $F(U, U) = F(U)$ (consistency), then over lakes at rest one has

$$U_{i+1/2-} = U_{i+1/2+} \implies \frac{\widetilde{F}_{i+1/2} - \widetilde{F}_{i-1/2}}{\Delta x} = \frac{F(U_{i+1/2-}) - F(U_{i-1/2+})}{\Delta x} = \widetilde{S}_i$$

Explicit kinetic scheme with **hydrostatic reconstruction**:

$$\begin{aligned}\widetilde{F}_{i+1/2} &= \int_{\mathbb{R}} \xi \left(\frac{1}{\xi} \right) \left(\mathbb{1}_{\xi > 0} M(U_{i+1/2-}, \xi) + \mathbb{1}_{\xi < 0} M(U_{i+1/2+}, \xi) \right) d\xi \\ \frac{U_i^{n+1} - U_i^n}{\Delta t} + \frac{1}{\Delta x} (\widetilde{F}_{i+1/2}^n - \widetilde{F}_{i-1/2}^n) &= \widetilde{S}_i^n\end{aligned}\tag{5}$$

Proposition 4 (Audusse, Bouchut, Bristeau, and Sainte-Marie 2016)

Under the CFL condition $\frac{\Delta t}{\Delta x} |\xi| < 1$ the scheme (5) preserves the water height positivity, and admits the discrete entropy inequality

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

where D_i features a quadratic error term, Lipschitz in $\sigma, \Delta x, \Delta z_i$ and vanishing when $u_i^n \rightarrow 0$.

⇒ We cannot ensure the dissipation of the total energy $\int_{\Omega} \eta(U(t, x)) dx$

The case of a varying topography: iterative kinetic scheme

To solve this issue, implicit the previous scheme

$$\frac{U_i^{n+1} - U_i^n}{\Delta t} + \frac{1}{\Delta x} (\tilde{F}_{i+1/2}^{n+1} - \tilde{F}_{i-1/2}^{n+1}) = \tilde{S}_i^{n+1}$$

Nonlinear system can't be solved directly \rightarrow iterative approximation

$$(1 + \alpha)U_i^{n+1,k+1} = U_i^n + \alpha U_i^{n+1,k} - \frac{\Delta t}{\Delta x} (\tilde{F}_{i+1/2}^{n+1,k} - \tilde{F}_{i-1/2}^{n+1,k}) + \tilde{S}_i^{n+1,k}, \quad \alpha \geq 0 \quad (6)$$

Proposition 5 (El Hassanieh, R., Sainte-Marie)

- We have $h_i^{n+1,k+1} \geq 0$ under the CFL

$$\forall \xi \in \mathbb{R}, \left(\frac{\Delta t}{\Delta x} |\xi| - \alpha \right) M(U_i^{n+1,k}, \xi) \leq M(U_i^n, \xi)$$

- The iterative process (6) satisfies the macroscopic entropy inequality

$$\frac{\eta(U_i^{n+1,k+1}, z_i) - \eta(U_i^n, z_i)}{\Delta t} + \frac{1}{\Delta x} (G_{i+1/2}^{n+1,k} - G_{i-1/2}^{n+1,k}) \leq D_i^{n+1,k},$$

with $D_i^{n+1,k} \leq 0$ from some rank k assuming (6) converges.

Sketch of the proof: relies on the kinetic rewriting

$$(1 + \alpha)f_i^{n+1,k+1} = M_i^n + \alpha M_i^{n+1,k} - \sigma\xi(M_{i+1/2}^{n+1,k} - M_{i-1/2}^{n+1,k}) + \sigma(\xi - u_i^{n+1,k})[M_{i+1/2-}^{n+1,k} - M_{i-1/2+}^{n+1,k}] \quad (7)$$

so that $U^{n+1,k} = \int_{\mathbb{R}} (1, \xi)^T f^{n+1,k} d\xi$ for any $k \in \mathbb{N}$

Positivity: The quantity $(1 + \alpha)h_i^{n+1,k}$ equals

$$\int_{\mathbb{R}} (M_i^n + \alpha M_i^{n+1,k} - \sigma\xi(M_{i+1/2}^{n+1,k} - M_{i-1/2}^{n+1,k})) d\xi \geq \int_{\mathbb{R}} (M_i^n + M_i^{n+1,k}(\alpha - \sigma|\xi|)) d\xi$$

Entropy inequality: Multiply (7) by $\partial_1 H(M_i^{n+1,k}, z_i, \xi)$ and use convexity of H

$$H(M_i^{n+1,k+1}, z_i) \leq H(M_i^n, z_i) - \sigma(\tilde{G}_{i+1/2}^{n+1,k} - \tilde{G}_{i-1/2}^{n+1,k}) + Q(\xi) + \tilde{D}_i,$$

with $\int_{\mathbb{R}} Q(\xi) d\xi = 0$ and

$$\tilde{D}_i = \text{Strictly negative term} + O(M_i^{n+1,k+1} - M_i^{n+1,k})$$

Proposition 6 (El Hassanieh, R., Sainte-Marie)

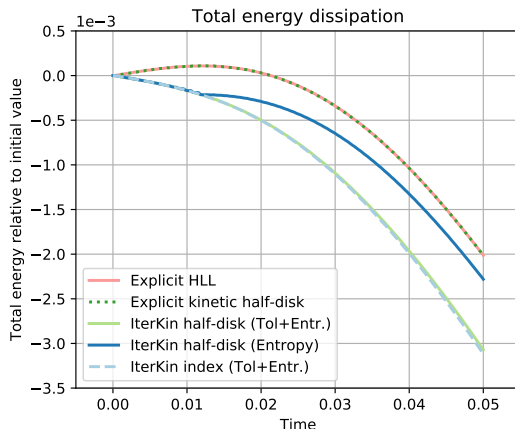
Assume the iterative scheme (7) keeps U_i^{k+1} in $\{(h, hu)^T, \delta \leq h \leq K_1, |u| \leq K_2\}$ for all k .
There exists $C(K_1, K_2, 1/\delta)$ such that $\Delta t \leq C\Delta x$ implies the convergence of $(f_i^{n+1,k})_{k \in \mathbb{N}}$ to f_i^{n+1} solution of the implicit scheme.

→ In practice, iterative process seems to converge without restriction

Stopping criteria: tolerance + total energy dissipation

$$\|U^{n+1,k+1} - U^{n+1,k}\| \leq \tau \quad \& \quad \frac{1}{\Delta t} \sum_{1 \leq i \leq N} (\eta(U_i^{n+1,k+1}, z_i) - \eta(U_i^n, z_i)) + \frac{1}{\Delta x} (G_{N+1/2}^{n+1,k} - G_{1/2}^{n+1,k}) \leq 0$$

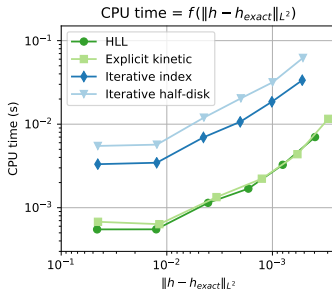
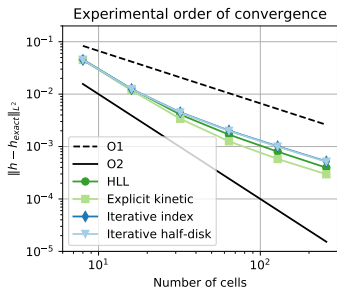
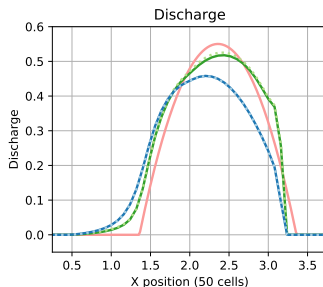
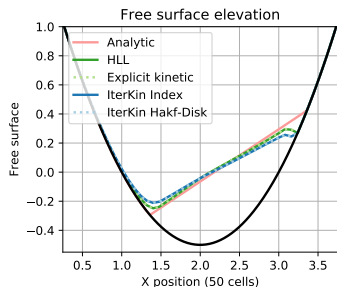
Total energy $\int_{\Omega} \eta \, dx$ should decrease in time due to **entropy inequality**.



Numerical testcase:

$$h + z = \text{Cst}$$

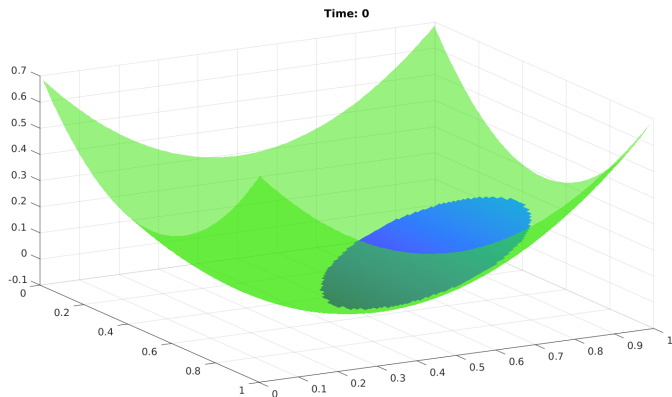
$$u = \text{Cst}' \neq 0$$



Implicit kinetic schemes: extension to 2D

→ Results still valid in 2D

→ Good approximation of the parabolic bowl (difficult numerical testcase)



For a flat topography

- Positivity and entropy inequality obtained unconditionally
- Obtained fully implicit scheme with explicit update for shallow water
- Optimal setting: inversion by hand, no factorization/iterative method
- Computational cost quadratic (cannot be improved further)

With varying bathymetry

- Hydrostatic reconstruction requires iterative strategy
- Positivity and entropy inequality hold under CFL

Advantageous framework for numerical analysis, but costly in practice

Perspectives

- Improve convergence proof
- 2D version of implicit scheme
- Increase order of accuracy (iterative only)

Application in oceanography

- Coarse resolution \Rightarrow dissipation D_i very large
- Improve hydrostatic reconstruction by also reconstructing velocity u
- Make $|D_i|$ smaller near Bernoulli equilibrium

Perspectives

- Improve convergence proof
- 2D version of implicit scheme
- Increase order of accuracy (iterative only)

Application in oceanography

- Coarse resolution \Rightarrow dissipation D_i very large
- Improve hydrostatic reconstruction by also reconstructing velocity u
- Make $|D_i|$ smaller near Bernoulli equilibrium

Thank you for your attention!