Low Froude regime and implicit kinetic schemes for the Saint-Venant system

Mathieu Rigal

Supervisors: Nina Aguillon, Nathalie Ayi, Jacques Sainte-Marie

Amiens – October 10, 2022

SORBONNE UNIVERSITÉ

Inria inventeurs du monde numérique

MATH INNOV

Région îledeFrance

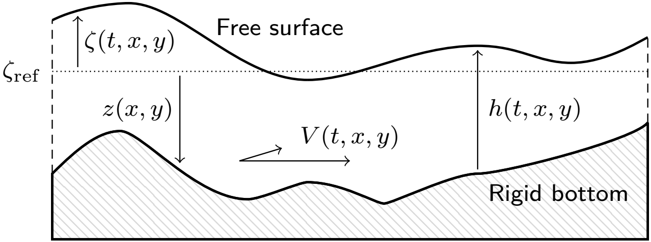
Introduction

Why are we interested in geophysical flows?

- water management (quality, availability);
- forecast natural disasters, mitigate their consequences;
- understand interplay between ocean dynamics and
  - → the weather;
  - → climate change;
  - → coastline erosion;
  - → natural resources (marine energy, seafood);

The 2D Saint-Venant system reads:

$$\begin{cases} \dfrac{\partial h}{\partial t} + \nabla \cdot (hV) = 0 \\ \dfrac{\partial}{\partial t}(hV) + \nabla \cdot (hV \otimes V) + \nabla\left(\dfrac{g}{2}h^2\right) = -gh\nabla z \end{cases} \tag{SV}$$



Important properties

- hyperbolicity;
- positivity of the water height;
- conservation of the water height (and discharge if $\nabla z = 0$);
- existence of non trivial steady states;
- entropy inequality $\partial_t \eta(U, z) + \nabla \cdot G(U, z) \leq 0 \rightarrow$ energy dissipation;
- as surface waves travel faster, solutions become incompressible;

Interest of implicit methods $\rightarrow$ better stability?

"ability to preserve approximation in some domain of physical validity"

IMEX methods for the low Froude regime

Reference Froude number: $\mathrm{Fr} \approx \dfrac{|V|}{\sqrt{gh}} = \dfrac{\text{particles velocity}}{\text{surface waves velocity}}$

Rescaling the quantities of interest, a dimensionless writting for (SV) is

$$\begin{cases} \dfrac{\partial h}{\partial t} + \nabla \cdot (hV) = 0 \\ \dfrac{\partial}{\partial t}(hV) + \nabla \cdot (hV \otimes V) + \dfrac{1}{\mathrm{Fr}^2}\nabla\left(\dfrac{h^2}{2}\right) = -\dfrac{1}{\mathrm{Fr}^2} h \nabla z \end{cases} \qquad (\mathcal{P}_{\mathrm{Fr}})$$

- Projecting along $n \in \mathbb{S}^2$, eigenvalues are $\lambda(U; n) = V \cdot n \pm \sqrt{h}/\mathrm{Fr}$.
- Usual explicit CFL reads $\Delta t \approx \delta/(2\lambda_{\max}) = \mathcal{O}(\mathrm{Fr}\delta)$, with δ the mesh res.
- In coastal flows, we consistently have $\mathrm{Fr} \approx 10^{-2}$.

$\longrightarrow$ We need **asymptotic stability** ($\Delta t = \mathcal{O}(\delta)$).

Split $(\mathcal{P}_{\mathrm{Fr}})$ into slow and fast dynamics: $\partial_t U + \nabla \cdot H(U,z) + L(U,z) = 0$

Ideally we would like

- $U = (h, hV)^T$ is the vector of conserved variables;
- both operators $\nabla \cdot H$ and L should be hyperbolic;
- $\nabla \cdot H$ will be treated explicitly and should have bounded eigenvalues;
- L will be treated implicitly and should be linear;
- Hydrostatic equilibrium should be in the kernel of $\nabla \cdot H$ and L;
- Positivity;

→ difficult to satisfy everything at once

Use same splitting as

Bispen et al. 2014 “IMEX Large Time Step Finite Volume Methods for Low Froude Number Shallow Water Flows.”

Set L such that $\partial_t U + L(U,z) = 0$ is the linearized of $(\mathcal{P}_{\mathrm{Fr}})$ around $(-z,0)^T$, then define H such that $\nabla \cdot H(U,z) = \nabla \cdot F(U) - S(U,z) - L(U,z)$

$$L(U,z) = \begin{pmatrix} \nabla \cdot (hV) \\ \frac{-z}{\mathrm{Fr}^2} \nabla(h+z) \end{pmatrix}, \quad H(U,z) = \begin{pmatrix} 0 \\ hV \otimes V + \frac{1}{2\mathrm{Fr}^2}(h+z)^2 \mathbf{I}_2 \end{pmatrix}$$

Eigenvalues are given by

$$\forall n \in \mathbb{S}^2, \quad \begin{cases} \lambda_H^j(U;n) = j(V \cdot n), & j \in \{0,1,2\} \\ \lambda_L^k(U;n) = k\sqrt{-z}/\mathrm{Fr}, & k \in \{-1,0,1\} \end{cases}$$

Time discretization: Implicit-Explicit Runge-Kutta

$$\begin{array}{c|cccc} \tilde{c}_1 & 0 & & & \\ \tilde{c}_2 & \tilde{a}_{21} & 0 & & \\ \vdots & \vdots & \vdots & \ddots & \\ \tilde{c}_s & \tilde{a}_{s1} & \tilde{a}_{s2} & \cdots & 0 \\ \hline & \tilde{b}_1 & \tilde{b}_2 & \cdots & \tilde{b}_s \end{array} \qquad \begin{array}{c|cccc} c_1 & a_{11} & & & \\ c_2 & a_{21} & a_{22} & & \\ \vdots & \vdots & \vdots & \ddots & \\ c_s & a_{s1} & a_{s2} & \cdots & a_{ss} \\ \hline & b_1 & b_2 & \cdots & b_s \end{array}$$

$$\frac{U^{(j)} - U^0}{\Delta t} + \sum_{k=1}^{j-1} \tilde{a}_{jk} \nabla \cdot H(U^{(k)}, z) + \sum_{k=1}^{j} a_{jk} L(U^{(k)}, z) = 0 \qquad \forall 1 \leq j \leq s$$

$$\frac{U^1 - U^0}{\Delta t} + \sum_{k=1}^{s} \tilde{b}_k \nabla \cdot H(U^{(k)}, z) + \sum_{k=1}^{s} b_k L(U^{(k)}, z) = 0$$

Low Froude regime and IMEX methods

L^2-stability

Set $z \equiv \text{Cst}$ and focus on the surface waves system $(\partial_t + L)U = 0$.

Assuming periodic boundary conditions, we conserve the energy

$$E[U](t) = \frac{1}{2}\|h\|^2_{L^2(\mathbb{T}^2)} + \frac{1}{2c^2}\|hV\|^2_{L^2(\mathbb{T}^2)}, \quad c = \sqrt{-z}/\text{Fr}.$$

What about the numerical approximation?

Definition 1 (Modified PDE, Warming and Hyett 1974)

"A modified PDE aims at describing the qualitative behaviour of a scheme by incorporating some of its truncation error into the original PDE."

More precisely, if $\widetilde{U}$ solves the p-th order modified PDE we have

$$\|\widetilde{U}(\Delta t, \cdot) - U^1_{\text{discrete}}\| \leq C\Delta t^{p+2}$$

L^2-stability criteria: $E[\widetilde{U}]' < 0$.

Proposition 1 (Runge-Kutta modified PDE and L^2-stability)

When discretized in time by a p-th order RK scheme $(\mathbf{A}, b)$*, the surface waves system admits the following* $(p+1)$*-th order modified PDE*

$$(\partial_t + L)U = \varphi_p \Delta t^p (-L)^{p+1} U + (\varphi_{p+1} - \varphi_p)\Delta t^{p+1}(-L)^{p+2} U \qquad (1)$$

with $\varphi_p = b^T \mathbf{A}^p \mathbb{1} - 1/(p+1)!$.

The energy E is strictly dissipated by (1) *if either one of the below holds*

- *p is even and* $(-1)^{p/2}[\varphi_{p+1} - \varphi_p] > 0$;
- *p is odd and* $(-1)^{(p+1)/2}\varphi_p < 0$;

Sketch of the proof: perform $U \cdot (1)$ and use integration by parts to find

$$(U, L^{2k+1}U)_{L^2} = 0, \quad (U, L^{2k+2}U)_{L^2} \text{ has the sign of } (-1)^{k+1}$$

Name	*Type*	*Order*	*L^2-stability*
Forward Euler	Explicit	1	No
Heun	Explicit	2	No
Midpoint	Explicit	2	No
Backward Euler	Implicit	1	Yes
Crank-Nicolson	Implicit	2	Inconclusive
Implicit ARS-(2,2,2)	Implicit	2	Yes
Implicit JIN-(2,2,2)	Implicit	2	Yes

Table: Applying various Runge-Kutta schemes to the surface waves equation.

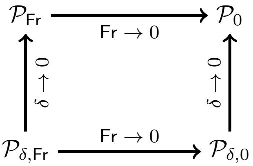
Low Froude regime and IMEX methods

Asymptotic consistency and limiting system

Asymptotic consistency: a scheme $\mathcal{P}_{\delta,\mathrm{Fr}}$ should remain consistent in the vanishing limit.

→ Consistency error needs to stay bounded.

→ Mimic qualitative behavior of solutions.



Limiting system derived formally through the asymptotic expansion:

$$f(t,x,y;\mathrm{Fr}) = f_{(0)}(t,x,y) + \mathrm{Fr}\, f_{(1)}(t,x,y) + \mathrm{Fr}^2\, f_{(2)}(t,x,y) + \mathcal{O}(\mathrm{Fr}^3)\ , \qquad f = h, V$$

Plug into ($\mathcal{P}_{\mathrm{Fr}}$) and isolate terms with same Froude powers

1. Start with the momentum equation to obtain $\nabla(h_{(0)} + z) = \nabla h_{(1)} = 0$
2. For *periodic boundary cond.*, the mass equation yields $\nabla \cdot (hV)_{(0)} = 0$
3. The momentum equation gives $\partial_t V_{(0)} + (V_{(0)} \cdot \nabla) V_{(0)} + \nabla h_{(2)} = 0$

Defining $\mathbb{W} \overset{\text{def}}{=} \Big\{(h, V) : \mathbb{T}^2 \to \mathbb{R}^3,\ \nabla(h+z) = 0,\ \nabla\cdot(hV) = 0\Big\}$, we have

$$\begin{cases} \forall t > 0,\ (h(t,\cdot), V(t,\cdot)) \in \mathbb{W} \\ \partial_t V + (V\cdot\nabla)V + \nabla\Pi = 0 \end{cases} \qquad (\mathcal{P}_0)$$

Remark 1 (Incompressible-like space)

When $\nabla z = 0$, $\mathbb{W}$ becomes the space of incompressible states.

Definition 2 (Space of well prepared data)

$$\mathbb{W}_p \overset{def}{=} \left\{ \sum_{k\in\mathbb{N}} \mathrm{Fr}^k \begin{pmatrix} h_{(k)} \\ V_{(k)} \end{pmatrix} : \mathbb{T}^2 \to \mathbb{R}^3,\ \begin{pmatrix} h_{(0)} \\ V_{(0)} \end{pmatrix} \in \mathbb{W},\ \nabla h_{(1)} = 0 \right\} \qquad (2)$$

The limit $(\mathcal{P}_{\mathrm{Fr}}) \to (\mathcal{P}_0)$ has been proved rigorously for well prepared data

Klainerman and Majda 1982 “Compressible and Incompressible Fluids”

Theorem 1 (Schochet 1994, "Fast singular limits of hyperbolic PDEs.")

Let $z \equiv \mathrm{Cst}$. The distance in $(L^2(\mathbb{T}^2))^3$ between a solution of $(\mathcal{P}_{\mathrm{Fr}})$ and that of $(\mathcal{P}_0)$ remains a $\mathcal{O}(\mathrm{Fr})$ if it is so at initial time.

Similar result obtained when linearizing $(\mathcal{P}_{\mathrm{Fr}})$ around $(-z, V^*)$

$$\partial_t \begin{bmatrix} h \\ V \end{bmatrix} + (V^* \cdot \nabla) \begin{bmatrix} h \\ V \end{bmatrix} + \begin{pmatrix} -z\nabla \cdot V \\ \mathrm{Fr}^{-2}\nabla h \end{pmatrix} = 0 \tag{3}$$

Hodge decomposition: $(L^2(\mathbb{T}^2))^3 = \mathcal{E} \oplus \mathcal{E}^\perp$ with $\mathcal{E} = (L^2(\mathbb{T}^2))^3 \cap \mathbb{W}$

Theorem 2 (Dellacherie 2010)

Let U be a solution of (3) in $(L^2(\mathbb{T}^2))^3$, and denote by $U_\mathcal{E}$ its orthogonal projection onto $\mathcal{E}$. Then $\partial_t U_\mathcal{E} + (V^* \cdot \nabla) U_\mathcal{E} = 0$ and the compressible energy $E_{\mathcal{E}^\perp} = \mathrm{Fr}^{-2}\|h - h_\mathcal{E}\|_{L^2}^2 + \|V - V_\mathcal{E}\|_{L^2}^2$ is constant in time.

Define characteristic $\gamma : (t, x, y) \mapsto (x - tV_x^*, y - tV_y^*)$. Theorem 2 implies

$$\begin{cases} \|h^0 - h^0_{\mathcal{E}}\|_{L^2} = \mathcal{O}(\mathrm{Fr}^2) \\ \|V^0 - V^0_{\mathcal{E}}\|_{L^2} = \mathcal{O}(\mathrm{Fr}) \end{cases} \Rightarrow \begin{cases} \|h - h^0_{\mathcal{E}} \circ \gamma\|_{L^2}(t > 0) = \mathcal{O}(\mathrm{Fr}^2) \\ \|V - V^0_{\mathcal{E}} \circ \gamma\|_{L^2}(t > 0) = \mathcal{O}(\mathrm{Fr}) \end{cases}$$

We want the scheme to mimic this behavior → study its modified PDE.

Theorem 3 (Refinement of Dellacherie's criteria)

Let the linear PDE $\partial_t U + \mathcal{F} U = 0$ be well-posed on $(L^2(\mathbb{T}^2))^3$. To have

$$\forall U^0 \in (L^2(\mathbb{T}^2))^3, \quad \|U^0 - U^0_{\mathcal{E}}\|_{L^2} = \mathcal{O}(\mathrm{Fr}) \Rightarrow \|U - U^0_{\mathcal{E}} \circ \gamma\|_{L^2} = \mathcal{O}(\mathrm{Fr})$$

it is sufficient to check that $U^0 \in \mathcal{E} \Rightarrow \|U - U^0 \circ \gamma\|_{L^2} = \mathcal{O}(\mathrm{Fr})$.

Simple proof using the triangle inequality

$$\|U - U^0_{\mathcal{E}} \circ \gamma\|_{L^2} \leq \|U - U^*\|_{L^2} + \|U^* - U^0_{\mathcal{E}} \circ \gamma\|_{L^2}, \quad \begin{cases} \partial_t U^* + \mathcal{F} U^* = 0 \\ U^*(t = 0) = U^0_{\mathcal{E}} \end{cases}$$

To summarize, when $V^* \equiv 0$ and $\gamma(t,\cdot) = \mathrm{id}$:

- Near incompressible states (e.g. well prepared data) should behave as near steady states;
- It is sufficient to check that the linear modified PDE keeps any incompressible data constant in time up to a $\mathcal{O}(\mathrm{Fr})$;
- We will apply this criteria in (h, hV)-coordinates instead of (h, V);

In practice we work with Fourier coefficients, remarking that

$$U \in \mathcal{E} \Leftrightarrow \widehat{U} \in \widehat{\mathcal{E}} \overset{\text{def}}{=} \left\{ \begin{pmatrix} \widehat{h} \\ \widehat{hV} \end{pmatrix} \in \ell^2(\mathbb{Z}^2; \mathbb{C}^3),\ \widehat{h}(k \neq 0) = 0 \text{ and } k \cdot \widehat{hV} = 0 \right\}$$

$\longrightarrow$ The modified PDE becomes an ODE $\partial_t \widehat{U}(t,k) = A(k)\widehat{U}(t,k)$.

$\longrightarrow$ $U \in \mathcal{E}$ is a steady state if $\forall k\ \widehat{U}(k) \in \ker A(k)$.

First consider the following first order discretization over a **cartesian mesh**

- In time: IMEX Euler;
- In space: HLL for $\nabla \cdot H$, Rusanov for L;

$$L_{i,j}^{\text{Rusanov}} = \begin{pmatrix} \mu_x \bar{\partial}_x(\mathrm{hV}_x) + \mu_y \bar{\partial}_y(\mathrm{hV}_y) \\ \mathrm{c}^2 \mu_x \bar{\partial}_x(\mathrm{h}+\mathrm{z}) \\ \mathrm{c}^2 \mu_y \bar{\partial}_y(\mathrm{h}+\mathrm{z}) \end{pmatrix}_{i,j} - \frac{\delta \mathrm{c}_{\max}}{2} \begin{pmatrix} [(\bar{\partial}_x)^2 + (\bar{\partial}_y)^2](\mathrm{h}+\mathrm{z}) \\ (\bar{\partial}_x)^2(\mathrm{hV}_x) \\ (\bar{\partial}_y)^2(\mathrm{hV}_y) \end{pmatrix}_{i,j} \quad (4)$$

Modified PDE for the surface waves discretization:

$$\begin{cases} \dfrac{\partial h}{\partial t} + \nabla \cdot Q = c\left[\dfrac{\Delta x}{2}\dfrac{\partial^2}{\partial x^2} + \dfrac{\Delta y}{2}\dfrac{\partial^2}{\partial y^2} + c\Delta t \varphi_1 \Delta\right] h \\ \dfrac{\partial Q}{\partial t} + c^2 \nabla h = c\left[\mathrm{diag}\left(\dfrac{\Delta x}{2}\dfrac{\partial^2}{\partial x^2}, \dfrac{\Delta y}{2}\dfrac{\partial^2}{\partial y^2}\right) + c\Delta t \varphi_1 \nabla \otimes \nabla\right] Q \end{cases} \quad (5)$$

If $(h, hV)^T \in \mathcal{E} + \mathcal{O}(\mathrm{Fr})$, in blue $= \mathcal{O}(\mathrm{Fr}^{-1})$, in red $= \mathcal{O}(\mathrm{Fr})$.

Low Froude regime and IMEX methods

A first order scheme

Proposition 2

Neglecting $\nabla \cdot Q$ in (5), the resulting system admits solutions satisfying

$$U^0 \in \mathcal{E} \Rightarrow \|U - U_{\mathcal{E}}\|_{L^2} = \mathcal{O}(\mathrm{Fr}) \quad \textit{and} \quad \lim_{\mathrm{Fr}\to 0} (\|\partial_x Q_x\|_{L^2} + \|\partial_y Q_y\|_{L^2})(\tau) = 0$$

for $\Delta t, \tau$ scale independent. ***The near steady condition isn't verified.***

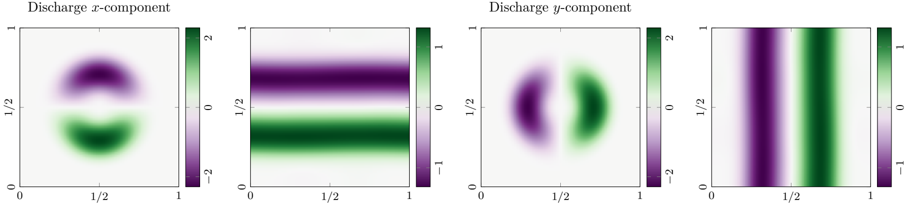


Figure: Steady vortex over non flat bottom, 100×100 mesh, $\mathrm{Fr} = 10^{-2}$.

Proposition 3

Consider a modified PDE $(\partial_t + L)U = (R_{\Delta t} + R_\delta)U$, with $R_{\Delta t}, R_\delta$ the time and spatial errors of the scheme. We always have $\mathcal{E} \subset \ker R_{\Delta t}$.

$\Rightarrow$ Any lack of near steady property is due to the spatial error.

Simple fix: increase order in space so that $R_\delta \ll R_{\Delta t}$.

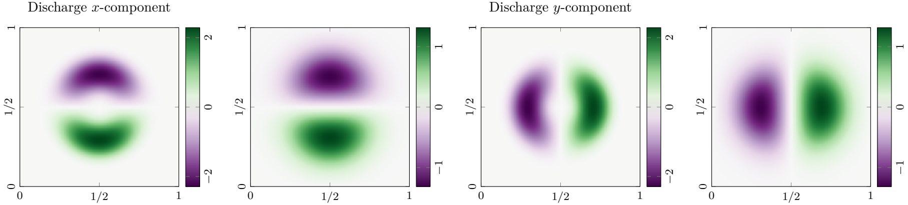


Figure: Improved results when replacing L^{Rusanov} by second order centered differences $L^\star$

Finally we compare second order schemes

- In time: ARS-(2,2,2);
- For $\nabla \cdot H$: HLL + MUSCL;
- For L we compare between:
  centered 2nd order $L^\star$, modified 2nd order $L^\flat$, centered 4th order $L^\sharp$;

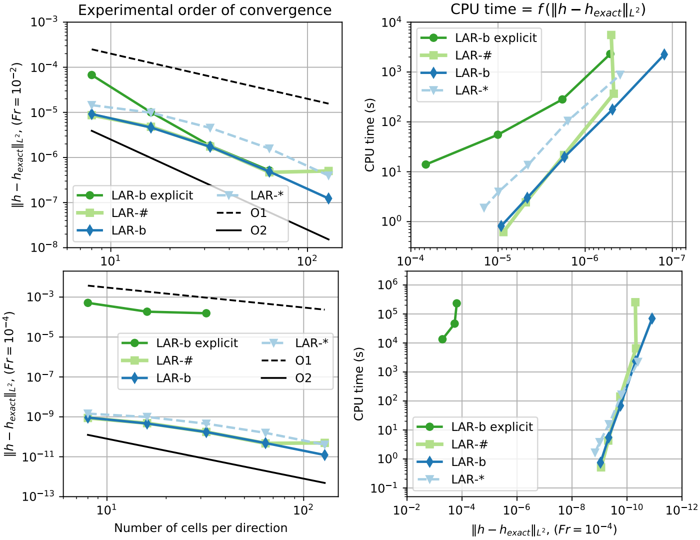
$R_\delta^\star$ not negligible compared to $R_{\Delta t}^{\text{ARS}}$. Nevertheless, the near steady condition is satisfied for Δt scale independent.

$$U^0 \in \mathcal{E} \Rightarrow \|U - U^0\|_{L^2} = \mathcal{O}(\text{Fr})$$

The modified PDEs for $L^\flat, L^\sharp$ satisfy the exact steady condition

$$U^0 \in \mathcal{E} \Rightarrow \|U - U^0\|_{L^2} = 0.$$

Comparing second order schemes



Implicit kinetic schemes

In collaboration with Chourouk El Hassanieh and Jacques Sainte-Marie

Preliminary contribution from Antonin Leprevost

We turn to 1D Saint-Venant system

$$\begin{cases} \partial_t h + \partial_x hu = 0 \\ \partial_t hu + \partial_x (hu^2 + \frac{g}{2} h^2) = -ghz' \end{cases} \tag{SV}$$

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

- Positivity of the water height $h(t, \cdot) \geq 0 \; \forall t$;
- Stationary state $h + z \equiv \text{Cst}$, $u \equiv 0$;
- Entropy-entropy flux pair (η, G) defined by

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

  and satisfying the **entropy inequality** $\partial_t \eta(U, z) + \partial_x G(U, z) \leq 0$;

Goal: satisfy these properties at the discrete level

Idea: work at finer scale by introducing the particle distribution $f(t,x,\xi)$.

$\rightarrow$ Particles velocity $\xi \in \mathbb{R}$ as a new variable

$\rightarrow$ Macroscopic quantities of interest recovered by integration

$$U_f(t,x) = \int_{\mathbb{R}} \begin{pmatrix} 1 \\ \xi \end{pmatrix} f(t,x,\xi)\,\mathrm{d}\xi$$

Boltzmann-like kinetic equation: $\underbrace{\partial_t f + \xi \partial_x f}_{\text{linear transport}} = \underbrace{Q[f](t,x,\xi)/\varepsilon}_{\text{collision operator}}$

Gibbs equilibrium: $f \in \ker Q \Leftrightarrow f = M(U_f,\xi) \overset{\text{def}}{=} \frac{1}{g\pi}\sqrt{\left(2gh - (\xi - u)^2\right)_+}$

The Maxwellian $M(U,\xi)$ satisfies the **moment relations**

$$\int_{\mathbb{R}} \begin{pmatrix} 1 \\ \xi \end{pmatrix} M(U,\xi)\,\mathrm{d}\xi = U, \quad \int_{\mathbb{R}} \xi \begin{pmatrix} 1 \\ \xi \end{pmatrix} M(U,\xi)\,\mathrm{d}\xi = F(U) \qquad \text{(M)}$$

Lemma 1 (Audusse et al. 2016)

U is a weak solution of (SV) iff $M(U,\xi)$ verifies

$$\partial_t M + \xi \partial_x M - g z' \partial_\xi M = \mu(t,x,\xi) \tag{KR}$$

with $\int_{\mathbb{R}} (1,\xi)^T \mu(t,x,\xi)\,\mathrm{d}\xi = 0$ for a.e. (t,x).

Kinetic representation (KR) obtained from Boltzman eq. in the limit $\varepsilon \to 0$.

Introduce BGK operator $Q[f] = M(U_f,\xi) - f$ and replace (KR) with BGK splitting

$$\begin{cases} \partial_t f = (M(U_f,\xi) - f)/\varepsilon \\ \partial_t f + \xi \partial_x f = 0 \end{cases} \xrightarrow{\varepsilon \to 0} \text{Solve } \begin{cases} \partial_t f + \xi \partial_x f = 0 \\ f^0 = M(U_f^0(x),\xi) \end{cases}$$

First we consider a flat bottom ($z \equiv \mathrm{Cst}$).

Piecewise constant approximation $f_i^n(\xi) \approx \frac{1}{\Delta x} \int_{C_i} f(t^n, x, \xi)\,\mathrm{d}\xi$, $1 \le i \le N$.

Explicit first order upwind scheme

$$\frac{f_i^{n+1} - f_i^n}{\Delta t} + \frac{\xi}{\Delta x}\Big(\mathbb{1}_{\xi<0}\big(f_{i+1}^n - f_i^n\big) + \mathbb{1}_{\xi>0}\big(f_i^n - f_{i-1}^n\big)\Big) = 0 \tag{6}$$

with initialization $f_i^n = M(U_i^n, \xi)$.

Macroscopic rewritting by integrating (6) against $(1,\xi)^T$

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

with kinetic flux $F(U_L, U_R) = \int_{\xi<0} \xi \begin{pmatrix} 1 \\ \xi \end{pmatrix} M(U_R,\xi)\,\mathrm{d}\xi + \int_{\xi>0} \xi \begin{pmatrix} 1 \\ \xi \end{pmatrix} M(U_L,\xi)\,\mathrm{d}\xi$.

We study the implicit version of the previous scheme.

$$\frac{f_i^{n+1} - f_i^n}{\Delta t} + \frac{\xi}{\Delta x}\Big(\mathbb{1}_{\xi<0}\big(f_{i+1}^{n+1} - f_i^{n+1}\big) + \mathbb{1}_{\xi>0}\big(f_i^{n+1} - f_{i-1}^{n+1}\big)\Big) = 0 \qquad (7)$$

Define the vector $f^n = (f_1^n, f_2^n, \dots, f_N^n)^T \in (\mathbb{R}_+)^N$ and $\sigma = \Delta t/\Delta x$.

Then (7) $\Leftrightarrow$ $(\mathbf{I} + \sigma\mathbf{L})f^{n+1} = f^n + \sigma B^{n+1}$ with

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

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

Implicit kinetic schemes

Discrete entropy inequality and positivity

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

Kinetic entropy $H(f, z, \xi) = \frac{\xi^2}{2} f + \frac{g^2 \pi^2}{6} f^3 + gzf$

Lemma 2 (Perthame and Simeoni 2001)

$\forall U, \quad \int_{\mathbb{R}} H(M(U,\xi), z, \xi)\, \mathrm{d}\xi = \eta(U, z), \quad \int_{\mathbb{R}} \xi H(M(U,\xi), z, \xi)\, \mathrm{d}\xi = G(U, z).$
Besides $M(U, \cdot)$ minimizes $f \mapsto \int_{\mathbb{R}} H(f(\xi), z, \xi)\, \mathrm{d}\xi$ under constraint $U_f = U$.

Proposition 1 (Audusse et al. 2016)

Under the CFL $\sigma|\xi| \leq 1$, the explicit kinetic scheme (6) satisfies $h_i^{n+1} \geq 0$ together with an inequality of the form $\eta(U_i^{n+1}) \leq \eta(U_i^n) - \sigma(G_{i+1/2}^n - G_{i-1/2}^n)$.

Proof: rewrite (6) as $f_i^{n+1} = (1 - \sigma|\xi|) M_i^n + \sigma|\xi| M_{i\pm1}^n$ (convex combination)

Also $\eta_i^{n+1} = \int_{\mathbb{R}} H(M_i^{n+1}, \xi)\, \mathrm{d}\xi \leq \int_{\mathbb{R}} H(f_i^{n+1}, \xi)\, \mathrm{d}\xi \leq \int_{\mathbb{R}} (1 - \sigma|\xi|) H_i^n + \sigma|\xi| H_{i\pm1}^n\, \mathrm{d}\xi$

Compare with implicit scheme:

Proposition 2

$\forall \Delta t > 0$, *the implicit kinetic scheme* (7) *satisfies* $h_i^{n+1} \geq 0$ *together with an equality of the form* $\eta(U_i^{n+1}) = \eta(U_i^n) - \sigma(\widetilde{G}_{i+1/2}^{n+1} - \widetilde{G}_{i-1/2}^{n+1}) + D_i$ *with* $D_i \leq 0$.

Proof:

- Matrix $(\mathbf{I} + \sigma\mathbf{L})$ is monotone and RHS has positive components.
- Regarding the entropy dissipation, multiply (7) 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

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

We know $(\mathbf{I}+\sigma\mathbf{L})^{-1}$, but cannot compute $\int_{\mathbb{R}} \begin{pmatrix} 1 \\ \xi \end{pmatrix} (\mathbf{I}+\sigma\mathbf{L})^{-1} M \,\mathrm{d}\xi$ where

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

Substitute M with a simpler Maxwellian satisfying the moment relations (M)

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

→ Unlike M, $\widetilde{M}$ doesn't minimize $\int_{\mathbb{R}} H(\cdot,\xi)\,\mathrm{d}\xi$;

→ as a consequence, we loose the discrete entropy inequality...

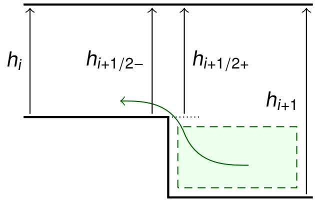
→ ... but the nonlinear implicit update can be rewritten explicitly;

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

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

- Upwinding introduces diffusion on h: $F_h(U_j, U_{j+1}) - F_h(U_{j-1}, U_j) \neq 0$
- Pressure variation should balance with source term: $\partial_x\left(\frac{g}{2} h^2\right) = -ghz'$

Hydrostatic reconstruction



Interface $i + 1/2$

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

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

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

$$\widetilde{U}_{i+1/2-} = h_{i+1/2-} \begin{pmatrix} 1 \\ u_i \end{pmatrix}, \quad \widetilde{U}_{i+1/2+} = h_{i+1/2+} \begin{pmatrix} 1 \\ u_{i+1} \end{pmatrix}$$

The numerical flux is modified as follows

$$\mathcal{F}_{i+1/2-} = F(\widetilde{U}_{i+1/2-}, \widetilde{U}_{i+1/2+}) + \frac{g}{2}(h_i^2 - h_{i+1/2-}^2)\begin{pmatrix}0\\1\end{pmatrix}$$

$$\mathcal{F}_{i-1/2+} = F(\widetilde{U}_{i-1/2-}, \widetilde{U}_{i-1/2+}) + \frac{g}{2}(h_i^2 - h_{i-1/2+}^2)\begin{pmatrix}0\\1\end{pmatrix}$$

Over lakes at rest, $\widetilde{U}_{i+1/2-} = \widetilde{U}_{i+1/2+} \rightarrow$ no more diffusion on h.

Kinetic interpretation:

$$\int_{\mathbb{R}} \begin{pmatrix}1\\ \xi\end{pmatrix}(\xi - u_i)(M_i - M_{i+1/2-})\,\mathrm{d}\xi = \begin{pmatrix}0\\ \frac{g}{2}(h_i^2 - h_{i+1/2-}^2)\end{pmatrix}, \quad M_{i+1/2-} = M(\widetilde{U}_{i+1/2-}, \xi)$$

$$\int_{\mathbb{R}} \begin{pmatrix}1\\ \xi\end{pmatrix}(\xi - u_i)(M_i - M_{i-1/2+})\,\mathrm{d}\xi = \begin{pmatrix}0\\ \frac{g}{2}(h_i^2 - h_{i-1/2+}^2)\end{pmatrix}, \quad M_{i-1/2+} = M(\widetilde{U}_{i-1/2+}, \xi)$$

Interfacial Maxwellian $M_{i+1/2} = M(\widetilde{U}_{i+1/2+}, \xi)\mathbb{1}_{\xi<0} + M(\widetilde{U}_{i+1/2-}, \xi)\mathbb{1}_{\xi>0}$.

Explicit kinetic scheme with hydrostatic reconstruction reads

$$\frac{f_i^{n+1} - f_i^n}{\Delta t} + \frac{\xi}{\Delta x}(M_{i+1/2}^n - M_{i-1/2}^n) + \frac{1}{\Delta x}(\xi - u_i^n)[M_{i-1/2+}^n - M_{i+1/2-}^n] = 0 \quad (8)$$

Proposition 3 (Audusse et al. 2016)

Scheme (8) *admits the discrete entropy inequality*

$$\eta(U_i^{n+1}) \leq \eta(U_i^n) - \sigma(\widetilde{G}_{i+1/2}^n - \widetilde{G}_{i-1/2}^n) + D_i, \quad D_i \geq 0$$

where the error term D_i can be positive.

Improve the stability by considering implicit version

$$\frac{f_i^{n+1} - f_i^n}{\Delta t} + \frac{\xi}{\Delta x}(M_{i+1/2}^{n+1} - M_{i-1/2}^{n+1}) + \frac{1}{\Delta x}(\xi - u_i^{n+1})[M_{i-1/2+}^{n+1} - M_{i+1/2-}^{n+1}] = 0 \quad (9)$$

Issue: nonlinear implicit update which cannot be computed exactly.
Instead we approximate f_i^{n+1} by the iterative process

$$(1+\alpha)f_i^{k+1} = f_i^n + \alpha f_i^k - \sigma\xi(M_{i+1/2}^k - M_{i-1/2}^k) + \sigma(\xi - u_i^k)[M_{i+1/2-}^k - M_{i-1/2+}^k] \quad (10)$$

Proposition 4

We have $h_i^{k+1} \geq 0$ *assuming* $\sigma|\xi| \leq \alpha + M_i^0/M_i^k$ *holds for any* $\xi \in \operatorname{supp} M^k$.

Proposition 5

The iterative process (10) *satisfies the macroscopic entropy inequality*

$$\eta(U_i^{k+1}, z_i) \le \eta(U_i^n, z_i) - \sigma(\widetilde{G}_{i+1/2}^k - \widetilde{G}_{i-1/2}^k) + D_i^k,$$

where D_i^k *becomes negative from some rank assuming* (10) *converges.*

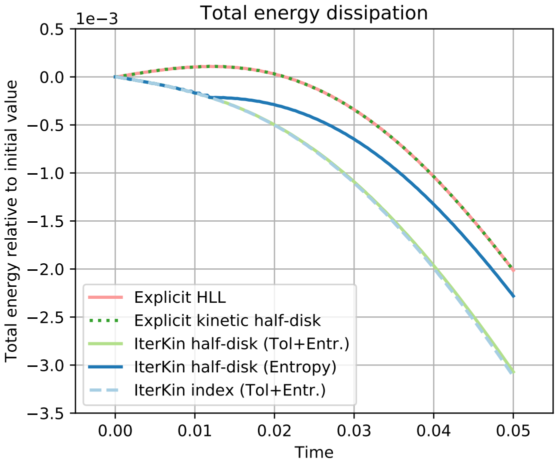
Proposition 6

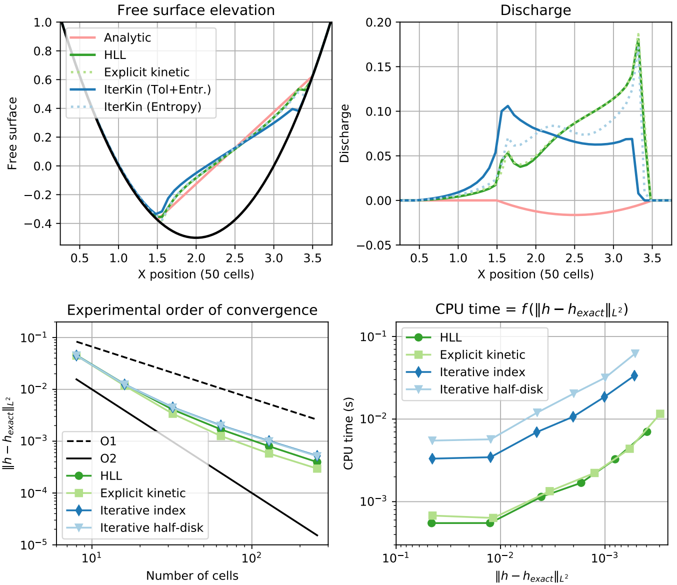
Assume (10) *keeps* U_i^{k+1} *in* $\left\{(h, hu)^T,\ \delta \le h \le K_1,\ |u| \le K_2\right\}$ *for all* k.
There exists $C(K_1, K_2, 1/\delta)$ *such that* $\sigma \le C$ *implies the convergence of* $(f_i^k)_{k\in\mathbb{N}}$ *to* f_i^{n+1} *solution of* (9).

Implicit kinetic schemes

Numerical tests

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




Free surface elevation
Analytic
HLL
Explicit kinetic
IterKin (Tol+Entr.)
IterKin (Entropy)
Free surface
X position (50 cells)
Discharge
Discharge
X position (50 cells)
Experimental order of convergence
O1
O2
HLL
Explicit kinetic
Iterative index
Iterative half-disk
Number of cells
CPU time = $f(\|h - h_{exact}\|_{L^2})$
HLL
Explicit kinetic
Iterative index
Iterative half-disk
CPU time (s)
$\|h - h_{exact}\|_{L^2}$

Conclusion and perspectives

IMEX Runge-Kutta schemes for the low Froude regime

- Stability analysis;
- Refined criteria to ensure near steady updates for surface waves eq.;
- Standard 2^{nd} order discretization leads to good results;

Perspectives: go beyond 2^{nd} order, study dispersion error...

Implicit kinetic schemes

- Without bathymetry, positivity and entropy inequality obtained unconditionally;
- Kinetic interpretation of the hydrostatic reconstruction requires iterative strategy;
- Positivity and entropy inequality hold under CFL;

Perspectives: improve convergence proof? 2D version of implicit scheme?