



HAL
open science

Revisiting energy estimates of the CPR scheme for the Shallow Water equations

Arnaud Duran

► **To cite this version:**

Arnaud Duran. Revisiting energy estimates of the CPR scheme for the Shallow Water equations. 2023. hal-04220336v1

HAL Id: hal-04220336

<https://hal.science/hal-04220336v1>

Preprint submitted on 27 Sep 2023 (v1), last revised 19 Apr 2024 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Revisiting energy estimates of the CPR scheme for the Shallow Water equations

A. Duran^{a,*}

^a*Université Claude Bernard Lyon 1, Institut Camille Jordan, UMR 5208*

^{*}*Institut Universitaire de France*

The present notes concern some improvements concerning the CPR (Centred Potential Regularization) schemes proposed in [2] and [1] in the particular case of the Shallow Water equations. We propose here a totally explicit variant of the scheme introduced in [1], and clarify the time step condition required to get energy dissipation (in other terms, entropy stability). This work also results in a slight modification of the viscosity terms, that may help improving numerical results far from low-Froude regimes. This document is organized as follows: in Section 1 we introduce the main notations as well as elementary technical lemmas concerning the discrete operators. The numerical scheme is given in Section 2, together with preliminary results regarding the discrete energy. The main stability results are discussed in Section 3, and extended in 2D in Section 4.

1 Introduction and general settings

We start here from the 1d Shallow Water system with topography:

$$\begin{cases} \partial_t h + \partial_x(hu) = 0, \\ \partial_t(hu) + \partial_x\left(hu^2 + \frac{gh^2}{2}\right) = -gh\partial_x z, \end{cases} \quad (1)$$

where $h = h(x, t)$ is the water height, $u = u(x, t)$ is the horizontal average velocity and $z = z(x)$ accounts for the topography variations. This system admits the following energy conservation:

$$\partial_t(he) + \partial_x\left(\left(e + \frac{1}{2}gh\right)hu\right) = 0, \quad (2)$$

where

$$e = \frac{1}{2}u^2 + \frac{1}{2}gh + gz.$$

In view of the design of the numerical scheme, we now introduce the discrete operators used in the next sections, as well as elementary estimates and duality formulas. In the one dimensional case, we consider a uniform grid with space step Δx , and denote $K \in \mathbb{Z}$ the element indices. First, for any sequence of scalar interface quantity $(b_{K+1/2})$, we define the operator:

$$\partial_K b = \frac{1}{\Delta x} (b_{K+1/2} - b_{K-1/2}), \quad (3)$$

with the specific notation $\partial_K^c b$ when the interface quantities correspond to the half sum of the values sharing the interface, that is:

$$\partial_K^c b = \frac{1}{\Delta x} (\bar{b}_{K+1/2} - \bar{b}_{K-1/2}) , \quad (4)$$

where $\bar{b}_{K+1/2} = \frac{1}{2}(b_K + b_{K+1})$. We also introduce the interface jump $[b]_{K+1/2} = \frac{1}{2}(b_{K+1} - b_K)$, so that $b_K = \bar{b}_{K+1/2} - [b]_{K+1/2} = \bar{b}_{K-1/2} + [b]_{K+1/2}$ for all $K \in \mathbb{Z}$. When no confusion is possible, these two definitions will extend to multiple variables by considering the difference of term to term interface products, leading to, in the case of two variables:

$$\partial_K(a, b) = \frac{1}{\Delta x} (a_{K+1/2} b_{K+1/2} - a_{K-1/2} b_{K-1/2}) . \quad (5)$$

We also define the discrete upwind derivative, for any collocated sequence of scalars (a_K)

$$\partial_K^{up}(a, b) = \frac{1}{\Delta x} \left(\mathcal{F}_{K+1/2}^{up}(a, b) - \mathcal{F}_{K-1/2}^{up}(a, b) \right) , \quad (6)$$

where $\mathcal{F}_{K+1/2}^{up}(a, b) = a_K b_{K+1/2}^+ + a_{K+1} b_{K+1/2}^-$, and $w^+ = \frac{1}{2}(w + |w|)$, $w^- = \frac{1}{2}(w - |w|)$ for any scalar quantity $w \in \mathbb{R}$. In what follows we will regularly use the notation:

$$\mathcal{S}_K(a, b) = a_{K+1/2} b_{K+1/2} + a_{K-1/2} b_{K-1/2} , \quad (7)$$

with its natural generalization to the sum of interface products $\mathcal{S}_K(a_1, \dots, a_n)$. When no confusion is possible, we will use the convention $\mathcal{S}_K(a^2) = \mathcal{S}_K(a, a)$. Remark that $a_K \mathcal{S}_K(b) = \bar{a}_{K+1/2} b_{K+1/2} + \bar{a}_{K-1/2} b_{K-1/2} - [a]_{K+1/2} b_{K+1/2} + [a]_{K-1/2} b_{K-1/2}$, which gives the duality formula:

$$a_K \mathcal{S}_K(b) = \mathcal{S}_K(\bar{a}, b) - \Delta x \partial_K([a], b) . \quad (8)$$

Lemma 1.1 (*Centred fluxes duality*). *We have the following duality formula:*

$$a_K \partial_K^c b + b_K \partial_K^c a = \widetilde{\partial}_K^c(a, b) ,$$

where

$$\widetilde{\partial}_K^c(a, b) = \partial_K^c(a, b) - \partial_K([a], [b]) . \quad (9)$$

Proof. We have, using $a_K = \bar{a}_{K+1/2} - [a]_{K+1/2} = \bar{a}_{K-1/2} + [a]_{K-1/2}$:

$$\begin{aligned} a_K \partial_K^c b + b_K \partial_K^c a &= \frac{1}{\Delta x} [a_K (\bar{b}_{K+1/2} - \bar{b}_{K-1/2}) + b_K (\bar{a}_{K+1/2} - \bar{a}_{K-1/2})] \\ &= \frac{1}{\Delta x} [a_K (\bar{b}_{K+1/2} - \bar{b}_{K-1/2}) + b_K ([a]_{K+1/2} + [a]_{K-1/2})] \\ &= \frac{1}{\Delta x} \left[\bar{a}_{K+1/2} \bar{b}_{K+1/2} - \bar{a}_{K-1/2} \bar{b}_{K-1/2} \right. \\ &\quad \left. + [a]_{K+1/2} (b_K - \bar{b}_{K+1/2}) + [a]_{K-1/2} (b_K - \bar{b}_{K-1/2}) \right] \\ &= \partial_K^c(a, b) - \partial_K([a], [b]) . \end{aligned}$$

□

Lemma 1.2 (*Upwind fluxes duality*). *We have the following duality formula:*

$$\partial_K^{up}(a, b) = a_K \partial_K b + (b \partial a)_K^{up},$$

where

$$(b \partial a)_K^{up} = b_{K+1/2}^- \left(\frac{a_{K+1} - a_K}{\Delta x} \right) + b_{K-1/2}^+ \left(\frac{a_K - a_{K-1}}{\Delta x} \right). \quad (10)$$

Proof. We have:

$$\begin{aligned} \Delta x [\partial_K^{up}(a, b) - a_K \partial_K b] &= \left[\left(a_K b_{K+1/2}^+ + a_{K+1} b_{K+1/2}^- \right) - \left(a_{K-1} b_{K-1/2}^+ + a_K b_{K-1/2}^- \right) \right] \\ &\quad - \left[a_K \left(b_{K+1/2}^+ + b_{K+1/2}^- \right) - a_K \left(b_{K-1/2}^+ + b_{K-1/2}^- \right) \right] \\ &= b_{K+1/2}^- (a_{K+1} - a_K) + b_{K-1/2}^+ (a_K - a_{K-1}). \end{aligned}$$

□

Lemma 1.3 (*Upwind fluxes duality - Squares*). *We have the following equality:*

$$2a_K (b \partial a)_K^{up} + (a_K)^2 \partial_K b = \partial_K^{up}(a^2, b) + \frac{4}{\Delta x} \mathcal{R}_K^+([a]^2, b),$$

where

$$\mathcal{R}_K^+([a]^2, b) := (-b_{K+1/2})^+ [a]_{K+1/2}^2 + (b_{K-1/2})^+ [a]_{K-1/2}^2. \quad (11)$$

Proof. Using the classical relation

$$(y - x)x = \frac{1}{2}y^2 - \frac{1}{2}x^2 - \frac{1}{2}(x - y)^2, \quad (12)$$

we have:

$$\begin{aligned} \Delta x (2a_K (b \partial a)_K^{up} + (a_K)^2 \partial_K b) &= 2a_K \left(b_{K+1/2}^- (a_{K+1} - a_K) + b_{K-1/2}^+ (a_K - a_{K-1}) \right) + (a_K)^2 (b_{K+1/2} - b_{K-1/2}) \\ &= b_{K+1/2}^- \left((a_{K+1})^2 - (a_K)^2 - (a_{K+1} - a_K)^2 \right) - b_{K-1/2}^+ \left((a_{K-1})^2 - (a_K)^2 - (a_{K-1} - a_K)^2 \right) \\ &\quad + (a_K)^2 (b_{K+1/2}^+ + b_{K+1/2}^-) - (a_K)^2 (b_{K-1/2}^+ + b_{K-1/2}^-) \\ &= \left(b_{K+1/2}^- (a_{K+1})^2 + b_{K+1/2}^+ (a_K)^2 \right) - \left(b_{K-1/2}^- (a_K)^2 + b_{K-1/2}^+ (a_{K-1})^2 \right) \\ &\quad + \left(-b_{K+1/2}^- (a_{K+1} - a_K)^2 + b_{K-1/2}^+ (a_K - a_{K-1})^2 \right) \end{aligned}$$

Going back to the definition (6) and using $-(w^-) = -\frac{1}{2}(w - |w|) = (-w)^+$, we get the announced result. □

Lemma 1.4 (*Square estimate for upwind fluxes*). *With the notations of the previous Lemma, we have the following inequality:*

$$[(b \partial a)_K^{up}]^2 \leq \frac{4}{\Delta x^2} \left((-b_{K+1/2})^+ + b_{K-1/2}^+ \right) \mathcal{R}_K^+([a]^2, b).$$

Proof. We write:

$$[(b\partial a)_K^{up}]^2 = \frac{4}{\Delta x^2} \left[(-b_{K+1/2})^+ \times -[a]_{K+1/2} + b_{K-1/2}^+[a]_{K-1/2} \right]^2,$$

and use Cauchy-Schwarz inequality with the vectors:

$$\mathbf{u} = \begin{pmatrix} \sqrt{(-b_{K+1/2})^+} \\ \sqrt{b_{K-1/2}^+} \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} \sqrt{(-b_{K+1/2})^+} \times -[a]_{K+1/2} \\ \sqrt{b_{K-1/2}^+} [a]_{K-1/2} \end{pmatrix}$$

to obtain

$$|\mathbf{u} \cdot \mathbf{v}|^2 \leq \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 = \left((-b_{K+1/2})^+ + b_{K-1/2}^+ \right) \cdot \left((-b_{K+1/2})^+ [a]_{K+1/2}^2 + b_{K-1/2}^+ [a]_{K-1/2}^2 \right).$$

□

2 Numerical scheme and preliminary energy estimates

Introducing the scalar potential $\phi = g(h + z)$, we consider the following numerical scheme for the model (1):

$$\begin{cases} h_K^{n+1} = h_K^n - \Delta t \partial_K (hu^*), \\ (hu)_K^{n+1} = (hu)_K^n - \Delta t \partial_K^{up}(u, hu^*) - \Delta t h_K^n \partial_K \phi^*. \end{cases} \quad (13)$$

Referring to (3) and (6), we only need to define $(hu)_{K+1/2}^*$ and $\phi_{K+1/2}^*$ to characterize the discrete operators. Following [1], we set:

$$(hu)_{K+1/2}^* = \overline{hu}_{K+1/2} - \Pi_{K+1/2}, \quad (14)$$

and

$$\phi_{K+1/2}^* = \overline{\phi}_{K+1/2} - \Lambda_{K+1/2}, \quad (15)$$

where we recall that the superscript “ $-$ ” refers to the mean interface value taken at time n (the scheme being totally explicit, the time indice will be omitted when possible in the following). The quantities $\Pi_{K+1/2}$ and $\Lambda_{K+1/2}$ govern the numerical viscosity of the scheme, and will be defined in Proposition 3.1 (formulas (37) and (38)) in order to guarantee energy dissipation. Using the mass equation in (13), we easily get:

$$(hu)_K^{n+1} - (hu)_K^n = h_K^{n+1} (u_K^{n+1} - u_K^n) - \Delta t u_K^n \partial_K (hu^*).$$

Invoking the momentum equation:

$$h_K^{n+1} (u_K^{n+1} - u_K^n) = -\Delta t (\partial_K^{up}(u, hu^*) - u_K^n \partial_K (hu^*)) - \Delta t h_K^n \partial_K \phi^*,$$

and using Lemma 1.2, we obtain the velocity equation:

$$u_K^{n+1} = u_K^n - \Delta t \frac{1}{h_K^{n+1}} (hu^* \partial u)_K^{up} - \Delta t \frac{h_K^n}{h_K^{n+1}} \partial_K \phi^*, \quad (16)$$

the operator $(hu^* \partial u)_K^{up}$ being based on (10). In what follows, we will consider an advective CFL condition of the form:

$$\frac{\Delta t}{\Delta x} (|hu_{K+1/2}^*| + |hu_{K-1/2}^*|) \leq \frac{\rho}{1 + \rho} h_K^n, \quad (17)$$

where ρ is a strictly positive constant. Let's remark that the time constraint (17) implies:

$$h_K^n - h_K^{n+1} = \frac{\Delta t}{\Delta x} (hu_{K+1/2}^* - hu_{K-1/2}^*) \leq \frac{\Delta t}{\Delta x} (|hu_{K+1/2}^*| + |hu_{K-1/2}^*|) \leq \frac{\rho}{1+\rho} h_K^n,$$

which gives

$$h_K^n \leq (1+\rho)h_K^{n+1}. \quad (18)$$

In particular, we have, under (17):

$$\frac{\Delta t}{\Delta x} \left((-hu_{K+1/2}^*)^+ + (hu_{K+1/2}^*)^+ \right) \leq \rho h_K^{n+1}. \quad (19)$$

The next proposition gives local estimates for the potential energy:

Proposition 2.1. *Let $\mathcal{E}_K^n = \frac{1}{2}g(h_K^n)^2 + gh_K^n z_K$ the potential energy at time n at the level of the cell K . Setting $\lambda = \frac{\Delta t}{\Delta x}$, we have the following estimate:*

$$\mathcal{E}_K^{n+1} - \mathcal{E}_K^n + \Delta t \widetilde{\partial}_K(\phi, hu^*) - \Delta t (hu)_K^n \partial_K^c \phi \leq 2g\lambda^2 \mathcal{S}_K([hu]^2) + 2g\lambda^2 \mathcal{S}_K(\Pi^2) - \lambda \mathcal{S}_K(\Pi, [\phi]), \quad (20)$$

where the operator \mathcal{S}_K is defined according to (7). The flux term $\widetilde{\partial}_K(\phi, hu^*)$ is defined by

$$\widetilde{\partial}_K(\phi, hu^*) = \partial_K(\phi, hu^*) - \partial_K([\phi], [hu]),$$

where $\partial_K(\phi, hu^*)$ and $\partial_K([\phi], [hu])$ are defined by (5). The term $\partial_K([\phi], [hu])$ is a second order contribution with respect to the mesh size, seen as a bias on the leading flux term $\partial_K(\phi, hu^*)$.

Proof. A basic computation give:

$$\mathcal{E}_K^{n+1} - \mathcal{E}_K^n = (h_K^{n+1} - h_K^n) \phi_K^n + \frac{1}{2}g(h_K^{n+1} - h_K^n)^2. \quad (21)$$

On a first hand we have:

$$\begin{aligned} (h_K^{n+1} - h_K^n) \phi_K^n &= -\frac{\Delta t}{\Delta x} (hu_{K+1/2}^* - hu_{K-1/2}^*) \phi_K^n \\ &= -\frac{\Delta t}{\Delta x} (hu_{K+1/2}^* \bar{\phi}_{K+1/2} - hu_{K-1/2}^* \bar{\phi}_{K-1/2}) \\ &\quad + \frac{\Delta t}{\Delta x} (hu_{K+1/2}^* [\phi]_{K+1/2} + hu_{K-1/2}^* [\phi]_{K-1/2}) \\ &= -\Delta t \partial_K(\phi, hu^*) + \frac{\Delta t}{\Delta x} (\bar{hu}_{K+1/2} [\phi]_{K+1/2} + \bar{hu}_{K-1/2} [\phi]_{K-1/2}) - \frac{\Delta t}{\Delta x} \mathcal{S}_K(\Pi, [\phi]) \\ &= -\Delta t \partial_K(\phi, hu^*) + \frac{\Delta t}{\Delta x} (hu)_K^n ([\phi]_{K+1/2} + [\phi]_{K-1/2}) \\ &\quad + \Delta t \partial_K([\phi], [hu]) - \frac{\Delta t}{\Delta x} \mathcal{S}_K(\Pi, [\phi]) \\ &= -\Delta t \widetilde{\partial}_K(\phi, hu^*) + \Delta t (hu)_K^n \partial_K^c \phi - \lambda \mathcal{S}_K(\Pi, [\phi]). \end{aligned}$$

Using standard convexity inequalities, the remaining term of the right hand side in (21) can be estimated as follows:

$$\begin{aligned} \frac{g}{2} (h_K^{n+1} - h_K^n)^2 &= \frac{g}{2} \lambda^2 (\overline{hu}_{K+1/2} - \overline{hu}_{K-1/2} - (\Pi_{K+1/2} - \Pi_{K-1/2}))^2 \\ &= \frac{g}{2} \lambda^2 ([hu]_{K+1/2} + [hu]_{K-1/2} - (\Pi_{K+1/2} - \Pi_{K-1/2}))^2 \\ &\leq 2g\lambda^2 \mathcal{S}_K ([hu]^2) + 2g\lambda^2 \mathcal{S}_K (\Pi^2) . \end{aligned}$$

Gathering the two previous contributions, we get the announced result. \square

We now turn to kinetic energy estimates:

Proposition 2.2. *Define $\mathcal{K}_K^n = \frac{1}{2} h_K^n (u_K^n)^2$ the local kinetic energy at time n . Let (ξ_K) a real sequence verifying $\xi_K \in]0, 1[$ for every $K \in \mathbb{Z}$, and assume that the local CFL condition (17) is ensured with $\rho = \xi_K$ at the level of the cell K . We have the following estimation:*

$$\begin{aligned} \mathcal{K}_K^{n+1} - \mathcal{K}_K^n + \Delta t \widetilde{\partial}_K^{up} \left(\frac{1}{2} u^2, hu^* \right) + \Delta t (hu)_K^n \partial_K^c \phi \\ \leq -\lambda \mathcal{S}_K ([hu], \Lambda) + 2\lambda^2 \mathcal{S}_K ([\phi]^2, h\nu) + 2\lambda^2 \mathcal{S}_K (\Lambda^2, h\nu) . \end{aligned} \quad (22)$$

Having in mind definition (6), the upwind flux is of the form

$$\widetilde{\partial}_K^{up} \left(\frac{1}{2} u^2, hu^* \right) = \partial_K^{up} \left(\frac{1}{2} u^2, hu^* \right) + \mathcal{A}_K ,$$

where term \mathcal{A}_K gathers $\mathcal{O}(\Delta t, \Delta x)$ antisymmetric terms, corresponding to a consistent perturbation of the leading flux. The sequence (ν_K) implied in estimate (22) is defined by:

$$\nu_K = \frac{1 + \xi_K}{1 - \xi_K} . \quad (23)$$

Proof. We start writing:

$$\begin{aligned} h_K^{n+1} (u_K^{n+1} - u_K^n) u_K^n &= h_K^{n+1} \left(\frac{1}{2} (u_K^{n+1})^2 - \frac{1}{2} (u_K^n)^2 - \frac{1}{2} (u_K^{n+1} - u_K^n)^2 \right) \\ &= \mathcal{K}_K^{n+1} - \mathcal{K}_K^n - \frac{1}{2} (u_K^n)^2 (h_K^{n+1} - h_K^n) - \frac{1}{2} h_K^{n+1} (u_K^{n+1} - u_K^n)^2 \\ &= \mathcal{K}_K^{n+1} - \mathcal{K}_K^n + \frac{1}{2} \Delta t (u_K^n)^2 \partial_K (hu^*) - \frac{1}{2} h_K^{n+1} (u_K^{n+1} - u_K^n)^2 . \end{aligned}$$

Hence, with the velocity equation (16):

$$\begin{aligned} \mathcal{K}_K^{n+1} - \mathcal{K}_K^n + \frac{1}{2} \Delta t (u_K^n)^2 \partial_K (hu^*) - \frac{1}{2} h_K^{n+1} (u_K^{n+1} - u_K^n)^2 \\ = u_K^n (-\Delta t (hu^* \partial u)_K^{up} - \Delta t h_K^n \partial_K \phi^*) , \end{aligned}$$

or equivalently:

$$\begin{aligned} \mathcal{K}_K^{n+1} - \mathcal{K}_K^n + \Delta t \left(\frac{1}{2} (u_K^n)^2 \partial_K (hu^*) + u_K^n (hu^* \partial u)_K^{up} \right) + \Delta t (hu)_K^n \partial_K \phi^* \\ = \frac{1}{2} h_K^{n+1} (u_K^{n+1} - u_K^n)^2 . \end{aligned} \quad (24)$$

We focus on the terms of the left hand side. With the help of Lemma 1.3, we write:

$$\begin{aligned} \frac{1}{2}(u_K^n)^2 \partial_K(hu^*) + u_K^n (hu^* \partial u)_K^{up} &= \frac{1}{2} \left((u_K^n)^2 \partial_K(hu^*) + 2u_K^n (hu^* \partial u)_K^{up} \right) \\ &= \partial_K^{up} \left(\frac{1}{2} u^2, hu^* \right) + \frac{2}{\Delta x} \mathcal{R}_K^+([u]^2, hu^*) . \end{aligned} \quad (25)$$

The last term of the left hand side in (24) can be subject to the following reformulation:

$$\begin{aligned} (hu)_K^n \partial_K \phi^* &= (hu)_K^n \partial_K^c \phi - \frac{1}{\Delta x} (hu)_K^n (\Lambda_{K+1/2} - \Lambda_{K-1/2}) \\ &= (hu)_K^n \partial_K^c \phi - \frac{1}{\Delta x} (\bar{h}u_{K+1/2} \Lambda_{K+1/2} - \bar{h}u_{K-1/2} \Lambda_{K-1/2}) \\ &\quad + \frac{1}{\Delta x} ([hu]_{K+1/2} \Lambda_{K+1/2} + [hu]_{K-1/2} \Lambda_{K-1/2}) \\ &= (hu)_K^n \partial_K^c \phi - \partial_K(\Lambda, hu) + \frac{1}{\Delta x} \mathcal{S}_K([hu], \Lambda) . \end{aligned} \quad (26)$$

Injecting (25) and (26) in (24), we get:

$$\mathcal{K}_K^{n+1} - \mathcal{K}_K^n + \Delta t \widetilde{\partial}_K^{up} \left(\frac{1}{2} u^2, hu^* \right) + \Delta t (hu)_K^n \partial_K^c \phi = -\lambda \mathcal{S}_K([hu], \Lambda) + \mathcal{U}_K , \quad (27)$$

where

$$\mathcal{U}_K = -2 \frac{\Delta t}{\Delta x} \mathcal{R}_K^+([u]^2, hu^*) + \frac{1}{2} h_K^{n+1} (u_K^{n+1} - u_K^n)^2 , \quad (28)$$

and $\widetilde{\partial}_K^{up} \left(\frac{1}{2} u^2, hu^* \right) = \partial_K^{up} \left(\frac{1}{2} u^2, hu^* \right) - \partial_K(\Lambda, hu)$. Remark that, as will be established rigorously later on, the viscous term Λ is scaled as $\mathcal{O}(\Delta t)$, so that the bias on the flux $\partial_K(\Lambda, hu)$ is of the expected form, according to the statement of the Proposition.

The following step relies on the estimation of the term $\frac{1}{2} h_K^{n+1} (u_K^{n+1} - u_K^n)^2$. Using again the velocity equation (16), we have:

$$\frac{1}{2} h_K^{n+1} (u_K^{n+1} - u_K^n)^2 = \frac{1}{2} h_K^{n+1} \Delta t^2 \left[\frac{1}{h_K^{n+1}} (hu^* \partial u)_K^{up} + \frac{h_K^n}{h_K^{n+1}} \partial_K \phi^* \right]^2 .$$

Using the equality

$$\partial_K \phi^* = \frac{1}{\Delta x} (\phi_{K+1/2}^* - \phi_{K-1/2}^*) = \frac{1}{\Delta x} ([\phi]_{K+1/2} + [\phi]_{K-1/2}) - \frac{1}{\Delta x} (\Lambda_{K+1/2} - \Lambda_{K-1/2}) ,$$

we write

$$\frac{1}{2} h_K^{n+1} (u_K^{n+1} - u_K^n)^2 = \frac{1}{2} \Delta t^2 h_K^{n+1} (C_1 + C_2 + C_3)^2 ,$$

with

$$\begin{aligned} C_1 &= \frac{1}{h_K^{n+1}} (hu^* \partial u)_K^{up} , \\ C_2 &= \frac{1}{\Delta x} \frac{h_K^n}{h_K^{n+1}} ([\phi]_{K+1/2} + [\phi]_{K-1/2}) , \\ C_3 &= -\frac{1}{\Delta x} \frac{h_K^n}{h_K^{n+1}} (\Lambda_{K+1/2} - \Lambda_{K-1/2}) . \end{aligned}$$

We now use the convexity of the square function with the weights $\left(\xi_K, \frac{1-\xi_K}{2}, \frac{1-\xi_K}{2}\right)$, so that:

$$\frac{1}{2}h_K^{n+1}(u_K^{n+1} - u_K^n)^2 \leq \mathcal{C}_1 + \mathcal{C}_2 + \mathcal{C}_3,$$

where

$$\begin{aligned} \mathcal{C}_1 &= \frac{1}{2}\Delta t^2 \frac{1}{\xi_K h_K^{n+1}} [(hu^* \partial u)_K^{up}]^2, \\ \mathcal{C}_2 &= \left(\frac{\Delta t}{\Delta x}\right)^2 \frac{h_K^n}{1-\xi_K} \left(\frac{h_K^n}{h_K^{n+1}}\right) ([\phi]_{K+1/2} + [\phi]_{K-1/2})^2, \\ \mathcal{C}_3 &= \left(\frac{\Delta t}{\Delta x}\right)^2 \frac{h_K^n}{1-\xi_K} \left(\frac{h_K^n}{h_K^{n+1}}\right) (\Lambda_{K+1/2} - \Lambda_{K-1/2})^2. \end{aligned}$$

Regarding the first term \mathcal{C}_1 , a direct use of Lemma 1.4 give:

$$\mathcal{C}_1 \leq 2 \left(\frac{\Delta t}{\Delta x}\right)^2 \frac{1}{\xi_K h_K^{n+1}} ((-hu_{K+1/2}^*)^+ + (hu_{K-1/2}^*)^+) \mathcal{R}_K^+([u]^2, hu^*), \quad (29)$$

so that, considering the first term of (28):

$$\begin{aligned} &\mathcal{C}_1 - 2 \frac{\Delta t}{\Delta x} \mathcal{R}_K^+([u]^2, hu^*) \\ &\leq 2 \frac{\Delta t}{\Delta x} \mathcal{R}_K^+([u]^2, hu^*) \left(\frac{\Delta t}{\Delta x} ((-hu_{K+1/2}^*)^+ + (hu_{K-1/2}^*)^+) / (\xi_K h_K^{n+1}) - 1\right). \end{aligned} \quad (30)$$

The right hand side is negative under the condition (17) with $\rho = \xi_K$ in virtue of (19). Now, considering that (18) holds, the ratio h_K^n/h_K^{n+1} can be bounded by $1 + \xi_K$ in \mathcal{C}_2 and \mathcal{C}_3 , leading to:

$$\mathcal{C}_2 \leq 2\lambda^2 h_K^n \frac{1 + \xi_K}{1 - \xi_K} \mathcal{S}_K([\phi]^2).$$

Using the notation $\nu_K = \frac{1 + \xi_K}{1 - \xi_K}$ and (8), this estimation can be reformulated as:

$$\mathcal{C}_2 \leq 2\lambda^2 \mathcal{S}_K([\phi]^2, h\nu) - 2\lambda^2 \Delta x \partial_K([\phi]^2, [h\nu]). \quad (31)$$

In a similar way, we obtain:

$$\mathcal{C}_3 \leq 2\lambda^2 \mathcal{S}_K(\Lambda^2, h\nu) - 2\lambda^2 \Delta x \partial_K(\Lambda^2, [h\nu]). \quad (32)$$

Gathering estimates (30), (31) and (32) we get, going back to (28):

$$\begin{aligned} \mathcal{U}_K &\leq 2\lambda^2 \mathcal{S}_K([\phi]^2, h\nu) + 2\lambda^2 \mathcal{S}_K(\Lambda^2, h\nu) \\ &\quad - 2\lambda^2 \Delta x \partial_K(\Lambda^2, [h\nu]) - 2\lambda^2 \Delta x \partial_K([\phi]^2, [h\nu]). \end{aligned} \quad (33)$$

Plugging this estimate in (27), we recognize the right hand side of (22) up to antisymmetric contributions scaled as $\mathcal{O}(\Delta t^2)$, that are incorporated in the fluxes $\widetilde{\partial}_K^{up}\left(\frac{1}{2}u^2, hu^*\right)$. \square

3 Global stability result

3.1 Choice of the stabilization terms

Denoting $E_K^n = \mathcal{E}_K^n + \mathcal{K}_K^n$ the local mechanical energy, the sum of energy estimates (20) and (22), can be written under the compact form:

$$E_K^{n+1} - E_K^n + \Delta t \mathfrak{d}_K \left(e + \frac{1}{2}gh, hu^* \right) \leq \mathcal{G}_K + \mathcal{L}_K, \quad (34)$$

where the energy flux is given by (we remark that $e + \frac{1}{2}gh = \frac{1}{2}u^2 + \phi$):

$$\mathfrak{d}_K \left(e + \frac{1}{2}gh, hu^* \right) = \widetilde{\partial}_K^{up} \left(\frac{1}{2}u^2, hu^* \right) + \widetilde{\partial}_K(\phi, hu^*),$$

and

$$\mathcal{G}_K = 2g\lambda^2 \mathcal{S}_K(\Pi^2) - \lambda \mathcal{S}_K(\Pi, [\phi]) + 2\lambda^2 \mathcal{S}_K([\phi]^2, h\nu), \quad (35)$$

$$\mathcal{L}_K = 2g\lambda^2 \mathcal{S}_K([q]^2) - \lambda \mathcal{S}_K([q], \Lambda) + 2\lambda^2 \mathcal{S}_K(\Lambda^2, h\nu), \quad (36)$$

We thus recognize at the left hand side a discrete equivalent of the continuous energy equation (2). We now discuss the conditions ensuring the negativity of the terms appearing in the right hand side.

Proposition 3.1 (*Control of the terms \mathcal{G}_K and \mathcal{L}_K*). *Suppose that the viscous terms involved in formulas (14), (15) have the following form:*

$$\Pi_{K+1/2} = 2\gamma \left(\frac{\Delta t}{\Delta x} \right) (\overline{h\nu})_{K+1/2} [\phi]_{K+1/2}, \quad (37)$$

and

$$\Lambda_{K+1/2} = 2\alpha g \left(\frac{\Delta t}{\Delta x} \right) [hu]_{K+1/2}, \quad (38)$$

where γ, α are positive constants. Assume that the following time constraint holds:

$$\frac{\Delta t}{\Delta x} \left(\sqrt{g(\overline{h\nu})_{K+1/2}} \right) \leq \frac{1}{4}, \quad (39)$$

where we recall that $\nu_K = \frac{1 + \xi_K}{1 - \xi_K}$ (see (23), Proposition 2.2). Then, if α, γ satisfy the following bounds:

$$\gamma, \alpha \in [r^-, r^+] \quad , \quad \text{with } r^\pm(\vartheta) = \frac{1 \pm \sqrt{1 - \vartheta}}{\vartheta/2}, \quad (40)$$

where $\vartheta = 16\lambda^2 g(\overline{h\nu})_{K+1/2}$, the terms \mathcal{G}_K and \mathcal{L}_K in (34) are both negative.

Proof. Considering the definition of \mathcal{S}_K (7), the term \mathcal{G}_K may be written as the sum of two interface contributions. More explicitly, we have $\mathcal{G}_K = \mathfrak{g}_{k+1/2} + \mathfrak{g}_{k-1/2}$ with

$$\mathfrak{g}_{K+1/2} = 2g\lambda^2 (\Pi_{K+1/2})^2 - \lambda \Pi_{K+1/2} [\phi]_{K+1/2} + 2\lambda^2 (\overline{h\nu})_{K+1/2} ([\phi]_{K+1/2})^2,$$

wich leads to identify conditions ensuring $\mathbf{g}_{K+1/2} \leq 0$ at the level of each cell interface. In what follows, we will focus at a generic interface $K + 1/2$ and remove the corresponding subscripts for the sake of clarity. We thus rewrite $\Pi = 2\gamma\lambda(\overline{h\nu})[\phi]$, to get:

$$\mathbf{g} = 2\lambda^2(\overline{h\nu})[\phi]^2 [1 - \gamma + (4\lambda^2 g(\overline{h\nu})) \gamma^2] .$$

We easily verify that the condition to have real roots leads to the CFL condition (39) and the negativity is ensured with the bounds (40) for the constant γ . In a similar way, we can write the term (36) as $\mathcal{L}_K = \mathfrak{l}_{K+1/2} + \mathfrak{l}_{K-1/2}$ (36), where:

$$\mathfrak{l}_{K+1/2} = 2g\lambda^2 ([q]_{K+1/2})^2 - \lambda[q]_{K+1/2}\Lambda_{K+1/2} + 2\lambda^2(h\nu)_{K+1/2} (\Lambda_{K+1/2})^2 .$$

This motivates a definition of the form (38), from which we extract exactly the same conditions, namely (39) and $\alpha \in [r^-, r^+]$. \square

Remark 3.2. *We have $\lim_{\vartheta \rightarrow 0} r^-(\vartheta) = 1$, $r^\pm(1) = 2$ and $\lim_{\vartheta \rightarrow 0} r^+(\vartheta) = +\infty$. The function $\vartheta \mapsto r^-(\vartheta)$ being increasing and $\vartheta \mapsto r^+(\vartheta)$ decreasing on the interval $[0, 1]$, we get that the value $\gamma = \alpha = 2$ is always admissible, independently from K . Note also the possibility of taking γ, α close to 1 by diminishing ϑ (that is the time step).*

3.2 Choice of parameters ξ_K and global CFL condition

We start by recalling that the negativity of the right hand side of the discrete energy budget (34) is ensured under the advective time constraint (17) with $\rho = \xi_K \in]0, 1[$, supplemented by the condition (39), implying the local quantity $\nu_K = \frac{1 + \xi_K}{1 - \xi_K}$. A first basic approach consists in choosing $\xi_K = \xi$ as constant over the mesh and consider separately the associated time constraints (17), (39), as done in [1]. This strategy appears to be relevant in low Froude regimes, where the time step is mainly governed by (39), which suggests a low value for ξ (so that ν is close to 1) to minimize diffusive losses. More generally, it is possible to exploit the local character of ξ_K to express an explicit balance between (17) and (39), depending on the flow regime. First remark that the condition (17) can be written as:

$$2\lambda\mathbf{u}_K \leq \frac{\xi_K}{1 + \xi_K}, \quad (41)$$

where we have set $\mathbf{u}_K = \frac{1}{2} (|hu_{K+1/2}^*| + |hu_{K-1/2}^*|) / h_K^n$. The condition (39) is

$$\lambda \left(\sqrt{g(\overline{h\nu})_{K+1/2}} \right) \leq \frac{1}{4} .$$

Using the fact that $\sqrt{\frac{1}{2}(a+b)} \leq \max(\sqrt{a}, \sqrt{b})$, this can be expressed through the local sufficient condition

$$\lambda\mathbf{c}_K \leq \frac{1}{4\sqrt{\nu_K}},$$

where $\mathbf{c}_K = \sqrt{gh_K^n}$. Since $\nu_K > 1$ we have $\frac{1}{\nu_K} < \frac{1}{\sqrt{\nu_K}}$, so that this condition is ensured under:

$$4\lambda\mathbf{c}_K \leq \frac{1 - \xi_K}{1 + \xi_K}. \quad (42)$$

Assume now that the following CFL condition holds:

$$\lambda \max_K(\mathbf{u}_K + \mathbf{c}_K) \leq 1/4. \quad (43)$$

This condition allows to define ξ_K according to (41), that is assume the case of equality $2\lambda\mathbf{u}_K = \frac{\xi_K}{1 + \xi_K}$. Indeed, this latter equality is equivalent to:

$$2\lambda\mathbf{u}_K(1 + \xi_K) = \xi_K \Leftrightarrow \xi_K(1 - 2\lambda\mathbf{u}_K) = 2\lambda\mathbf{u}_K. \quad (44)$$

Under (43), we have $1 - 2\lambda\mathbf{u}_K > 0$, so that we can define:

$$\xi_K = \frac{2\lambda\mathbf{u}_K}{1 - 2\lambda\mathbf{u}_K}, \quad (45)$$

with $\xi_K \in]0, 1[$. This gives

$$\frac{1 - \xi_K}{1 + \xi_K} = 1 - 4\lambda\mathbf{u}_K,$$

so that condition (42) is in fact equivalent to (43). We finally remark that under this condition:

$$4\lambda\mathbf{u}_K \leq \frac{\mathbf{u}_K}{\mathbf{u}_K + \mathbf{c}_K},$$

leading to $1/\nu_K = \frac{1 - \xi_K}{1 + \xi_K} = 1 - 4\lambda\mathbf{u}_K \geq 1 - \frac{\mathbf{u}_K}{\mathbf{u}_K + \mathbf{c}_K} = \frac{\mathbf{c}_K}{\mathbf{u}_K + \mathbf{c}_K}$. This implies:

$$\nu_K \leq 1 + (Fr)_K,$$

where $(Fr)_K = \frac{\mathbf{u}_K}{\mathbf{c}_K}$ stands for a local Froude number. To summarize, equipped with definitions (37), (38) and the notations

$$\mathbf{u}_K = \frac{1}{2} (|hu_{K+1/2}^*| + |hu_{K-1/2}^*|) / h_K^n, \quad \mathbf{c}_K = \sqrt{gh_K^n},$$

the scheme (13) is stable under the CFL condition (43) in the sense of existence of local energy estimate of the form:

$$E_K^{n+1} - E_K^n + \Delta t \mathfrak{d}_K \left(e + \frac{1}{2}gh, hu^* \right) \leq 0. \quad (46)$$

The constants γ, α involved in (37), (38) have to satisfy the bounds (40), and the quantity ν_K is defined by $\nu_K = \frac{1}{1 - 4\lambda\mathbf{u}_K}$.

3.3 Addition of passive transport

To complete the picture, we show how the method naturally extends to passive transport considering the following model:

$$\begin{cases} \partial_t h + \partial_x(hu) = 0, \\ \partial_t(hu) + \partial_x\left(hu^2 + \frac{gh^2}{2}\right) = -gh\partial_x z, \\ \partial_t(h\omega) + \partial_x(hu\omega) = 0, \end{cases} \quad (47)$$

ω referring to any scalar variable. The augmented energy relies on the quantity:

$$\underline{e} = \frac{1}{2}u^2 + \frac{1}{2}gh + gb + \kappa\omega^2,$$

where $\kappa > 0$, and the energy equation remains:

$$\partial_t(h\underline{e}) + \partial_x\left(\left(\underline{e} + \frac{1}{2}gh\right)hu\right) = 0. \quad (48)$$

The last equation in (47) can be integrated using the upwind fluxes defined in (6):

$$\begin{cases} h_K^{n+1} = h_K^n - \Delta t \partial_K(hu^*), \\ (hu)_K^{n+1} = (hu)_K^n - \Delta t \partial_K^{up}(u, hu^*) - \Delta t h_K^n \partial_K \phi^*, \\ (h\omega)_K^{n+1} = (h\omega)_K^n - \Delta t \partial_K^{up}(\omega, hu^*). \end{cases} \quad (49)$$

More precisely, we have the following result:

Proposition 3.3. *Let $\mathcal{W}_K^n = \kappa h_K^n (\omega_K^n)^2$ the local energy at time n associated with the variable ω . Under the CFL condition (17) with $\rho = 1$, we have:*

$$\mathcal{W}_K^{n+1} - \mathcal{W}_K^n + \Delta t \partial_K^{up}(\kappa\omega^2, hu^*) \leq 0. \quad (50)$$

Proof. The proof follows the first steps of the one of Proposition 2.2 for the kinetic energy. First, by similar arguments as those employed to get (16), we easily get:

$$\omega_K^{n+1} = \omega_K^n - \Delta t \frac{1}{h_K^{n+1}} (hu^* \partial \omega)_K^{up}. \quad (51)$$

Using the mass equation and the relation (12), we get:

$$2\kappa h_K^{n+1} (\omega_K^{n+1} - \omega_K^n) \omega_K^n = \mathcal{W}_K^{n+1} - \mathcal{W}_K^n + \kappa \Delta t (\omega_K^n)^2 \partial_K(hu^*) - \kappa h_K^{n+1} (\omega_K^{n+1} - \omega_K^n)^2,$$

and using the scheme (51), together with Lemma 1.3:

$$\begin{aligned} \mathcal{W}_K^{n+1} - \mathcal{W}_K^n &= -\kappa \Delta t (2\omega_K^n (hu^* \partial \omega)_K^{up} + (\omega_K^n)^2 \partial_K hu^*) + \kappa h_K^{n+1} (\omega_K^{n+1} - \omega_K^n)^2 \\ &= -\kappa \Delta t \partial_K^{up}(\omega^2, hu^*) - 4\kappa \frac{\Delta t}{\Delta x} \mathcal{R}_K^+([\omega]^2, hu^*) + \kappa h_K^{n+1} (\omega_K^{n+1} - \omega_K^n)^2. \end{aligned}$$

This previous equality can be reformulated as

$$\mathcal{W}_K^{n+1} - \mathcal{W}_K^n + \Delta t \partial_K^{up}(\kappa\omega^2, hu^*) = \kappa \mathcal{T}_K, \quad (52)$$

with $\mathcal{T}_K = h_K^{n+1} (\omega_K^{n+1} - \omega_K^n)^2 - 4 \frac{\Delta t}{\Delta x} \mathcal{R}_K^+([\omega]^2, hu^*)$. Thus it remains to establish that $\mathcal{T}_K \leq 0$ under the CFL condition (17). We have, using (51) and Lemma 1.4 :

$$\begin{aligned} h_K^{n+1} (\omega_K^{n+1} - \omega_K^n)^2 &= \frac{(\Delta t)^2}{h_K^{n+1}} [(hu^* \partial \omega)_K^{up}]^2 \\ &\leq 4 \left(\frac{\Delta t}{\Delta x} \right)^2 \frac{1}{h_K^{n+1}} ((-hu_{K+1/2}^*)^+ + (hu_{K-1/2}^*)^+) \mathcal{R}_K^{up}([\omega]^2, hu^*) \end{aligned}$$

Hence:

$$\mathcal{T}_K \leq 4 \frac{\Delta t}{\Delta x} (\mathcal{R}_K^+([\omega]^2, hu^*)) \left(\frac{\Delta t}{\Delta x} ((-hu_{K+1/2}^*)^+ + (hu_{K-1/2}^*)^+) / h_K^{n+1} - 1 \right),$$

which is negative under the CFL (17) in the case $\rho = 1$, according to (19). \square

It directly follows that the scheme stability is ensured under the same conditions (discussed in Section §3.2), with this time a discrete energy estimate of the form:

$$\underline{E}_K^{n+1} - \underline{E}_K^n + \Delta t \mathfrak{d}_K \left(\underline{e} + \frac{1}{2} gh, hu^* \right) \leq 0, \quad (53)$$

where $\underline{E}_K^n = \mathcal{E}_K^n + \mathcal{K}_K^n + \mathcal{W}_K^n$, and the energy flux is given by:

$$\mathfrak{d}_K \left(\underline{e} + \frac{1}{2} gh, hu^* \right) = \widetilde{\partial}_K^{up} \left(\frac{1}{2} u^2, hu^* \right) + \widetilde{\partial}_K(\phi, hu^*) + \partial_K^{up}(\kappa \omega^2, hu^*).$$

4 Two dimensional extension

To extend the present approach in the 2d case, we consider a 2d domain meshed with polygonal cells indexed by $K \in \mathbf{Z}$. In what follows, m_K and $m_{\partial K}$ will stand for the area and perimeter of cell K . For every edge $e \in \partial K$, we will employ the notations m_e and $\vec{n}_{e,K}$ respectively for the length and the outward normal of the corresponding boundary interface (see Fig 1). The set of equations we now consider is the following:

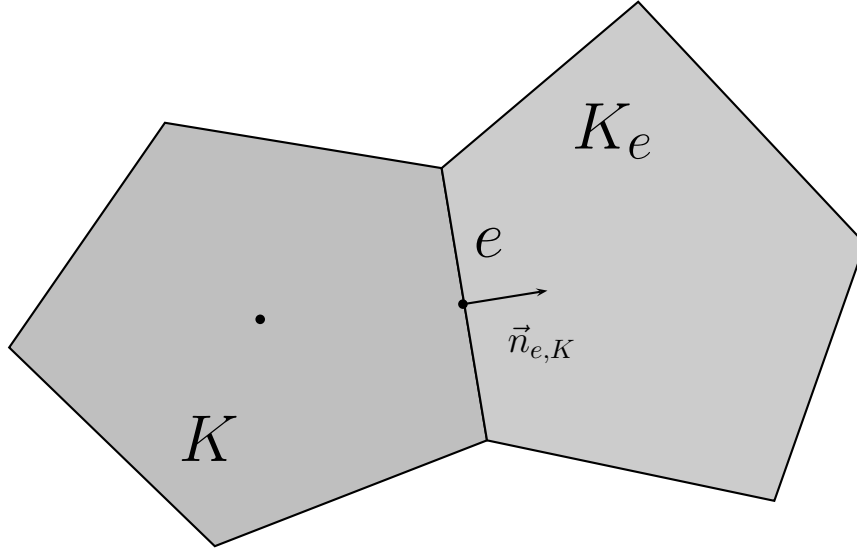


Figure 1: Geometric settings - focus on the interface e .

$$\begin{cases} \partial_t h + \operatorname{div}(\mathbf{h}\mathbf{u}) = 0, \\ \partial_t(\mathbf{h}\mathbf{u}) + \operatorname{div}\left(\mathbf{h}\mathbf{u} \otimes \mathbf{u} + \frac{gh^2}{2}\right) = -gh\nabla z, \end{cases} \quad (54)$$

which can be written in the compact form:

$$\partial_t W + \operatorname{div}\mathbf{F}(W) + hS(W) = 0, \quad (55)$$

where $\mathbf{F}(W) = \begin{pmatrix} \mathbf{h}\mathbf{u} \\ \mathbf{h}\mathbf{u} \otimes \mathbf{u} \end{pmatrix}$ and $S(W) = \begin{pmatrix} 0 \\ \nabla\phi \end{pmatrix}$, where we recall that $\phi = g(h + z)$.

4.1 Derivation of the scheme

Denoting W_K^n the approximate solution on the cell K at time t^n , we seek for a two-dimensional formulation in terms of convex combination of one dimensional schemes, that is:

$$W_K^{n+1} = \sum_{e \in \partial K} \alpha_{e,K} W_{e,K}^{n+1}, \quad (56)$$

where $W_{e,K}^{n+1}$ results from the application of the one dimensional approach in the normal direction $\vec{n}_{e,K}$, and $\alpha_{e,K}$ are positive coefficients such that $\sum_{e \in \partial K} \alpha_{e,K} = 1$. Before going

further, we first remark that the one dimensional scheme with passive transport (49) can be rewritten in the following form:

$$\begin{aligned} W_K^{n+1} = W_K^n &- \frac{\Delta t}{\Delta x} (\mathcal{F}(W_K^n, W_{K+1}^n) - \mathcal{F}(W_{K-1}^n, W_K^n)) \\ &- \frac{\Delta t}{\Delta x} h_K^n (\mathcal{S}(W_K^n, W_{K+1}^n) - \mathcal{S}(W_{K-1}^n, W_K^n)), \end{aligned} \quad (57)$$

where, using notation (6):

$$\mathcal{F}(W_K^n, W_{K+1}^n) = \begin{pmatrix} (hu)_{K+1/2}^* \\ \mathcal{F}_{K+1/2}^{up}(u, (hu)^*) \\ \mathcal{F}_{K+1/2}^{up}(\omega, (hu)^*) \end{pmatrix}, \quad \mathcal{S}(W_K^n, W_{K+1}^n) = \begin{pmatrix} 0 \\ \phi_{K+1/2}^* \\ 0 \end{pmatrix}. \quad (58)$$

Going back to the 2d formalism, let's now consider a generic interface e at the level of a cell K and apply this scheme to the states W_K, W_K, W_{K_e} in the reference associated to the outward normal $\vec{n}_{e,K}$. The velocity vector is subject to a change of coordinates which can be operated through the following mapping (we note $\vec{n}_{e,K} = {}^t(n_x, n_y)$):

$$\tau: W = \begin{pmatrix} h \\ hu \end{pmatrix} \mapsto \widehat{W} = \begin{pmatrix} h \\ hu^\tau \end{pmatrix}, \quad (59)$$

where

$$\begin{aligned} u^\tau &= un_x + vn_y, \\ v^\tau &= -un_y + vn_x. \end{aligned}$$

It can easily be checked that the return to original (x, y) coordinates through τ^{-1} is characterized by:

$$\begin{aligned} u &= u^\tau n_x - v^\tau n_y, \\ v &= u^\tau n_y + v^\tau n_x. \end{aligned}$$

Applying the basis scheme (57) in this reference to the states $\widehat{W}_K, \widehat{W}_K, \widehat{W}_{K_e}$, with a space step $\Delta_K = \frac{m_K}{m_{\partial K}}$ and v^τ as passive transport, we get the following auxiliary update:

$$\begin{aligned} \widehat{W}_{e,K}^{n+1} &= \widehat{W}_K^n - \frac{\Delta t}{\Delta_K} \left(\mathcal{F}(\widehat{W}_K^n, \widehat{W}_{K_e}^n) - \mathcal{F}(\widehat{W}_K^n, \widehat{W}_K^n) \right) \\ &\quad - \frac{\Delta t}{\Delta_K} h_K^n \left(\mathcal{S}(\widehat{W}_K^n, \widehat{W}_{K_e}^n) - \mathcal{S}(\widehat{W}_K^n, \widehat{W}_K^n) \right). \end{aligned} \quad (60)$$

By consistency of the fluxes \mathcal{F} with the 1d physical advective fluxes $F = {}^t(hu, hu^2, hw)$, we have $\mathcal{F}(\widehat{W}_K^n, \widehat{W}_K^n) = F(\widehat{W}_K^n)$, and it can easily be checked that $\tau^{-1}F(\widehat{W}_K^n) = \mathbf{F}(W_K^n) \cdot \vec{n}_{e,K}$. In a second hand, according to (58) and the definition (14), the computation of $\mathcal{F}(\widehat{W}_K^n, \widehat{W}_{K_e}^n)$ implies the following interface value:

$$(hu^\tau)_e^* = (\overline{hu^\tau})_e - \Pi_e,$$

where $(\overline{hu^\tau})_e = \frac{1}{2} (hu_{K_e}^\tau + hu_K^\tau)$, and, following definition (37):

$$\Pi_e = 2\gamma \left(\frac{\Delta t}{\Delta_K} \right) (\overline{h\nu})_e \left(\frac{\phi_{K_e} - \phi_K}{2} \right).$$

Remarking that $u^\tau = \mathbf{u} \cdot \vec{n}_{e,K}$, we get:

$$(hu^\tau)_e^* = (\mathbf{hu})_e^* \cdot \vec{n}_{e,K},$$

where

$$(\mathbf{hu})_e^* = (\overline{\mathbf{hu}})_e - \Pi_e, \quad \Pi_e = 2\gamma \left(\frac{\Delta t}{\Delta_K} \right) (\overline{h\nu})_e (\delta_e \phi), \quad (61)$$

with the jump operator $\delta_e \phi = \frac{1}{2}(\phi_{K_e} - \phi_K) \vec{n}_{e,K}$. As a result, we can write:

$$\mathcal{F}(\widehat{W}_K^n, \widehat{W}_{K_e}^n) = \begin{pmatrix} (\mathbf{h}\mathbf{u})_e^* \cdot \vec{n}_{e,K} \\ \mathcal{F}_e^{up}(u^\tau, (\mathbf{h}\mathbf{u})_e^* \cdot \vec{n}_{e,K}) \\ \mathcal{F}_e^{up}(v^\tau, (\mathbf{h}\mathbf{u})_e^* \cdot \vec{n}_{e,K}) \end{pmatrix}, \quad (62)$$

so that going back to the (x, y) coordinates we get:

$$\mathcal{F}(W_K^n, W_{K_e}^n, \vec{n}_{e,K}) := \tau^{-1} \mathcal{F}(\widehat{W}_K^n, \widehat{W}_{K_e}^n) = \begin{pmatrix} (\mathbf{h}\mathbf{u})_e^* \cdot \vec{n}_{e,K} \\ \mathcal{F}_e^{up}(\mathbf{u}, (\mathbf{h}\mathbf{u})_e^* \cdot \vec{n}_{e,K}) \end{pmatrix}, \quad (63)$$

where

$$\mathcal{F}_e^{up}(\mathbf{u}, (\mathbf{h}\mathbf{u})_e^* \cdot \vec{n}_{e,K}) = \mathbf{u}_K ((\mathbf{h}\mathbf{u})_e^* \cdot \vec{n}_{e,K})^+ + \mathbf{u}_{K_e} ((\mathbf{h}\mathbf{u})_e^* \cdot \vec{n}_{e,K})^-.$$

Remains the second component of $\mathcal{S}(\widehat{W}_K^n, \widehat{W}_{K_e}^n)$, which is, still according to the basis scheme (57):

$$\phi_e^* = \bar{\phi}_e - \Lambda_e, \quad (64)$$

with $\bar{\phi}_e = \frac{1}{2}(\phi_{K_e} - \phi_K)$ and the stabilization term (see (38)):

$$\begin{aligned} \Lambda_e &= 2\alpha g \left(\frac{\Delta t}{\Delta_K} \right) \left(\frac{hu_{K_e}^\tau - hu_K^\tau}{2} \right) \\ &= 2\alpha g \left(\frac{\Delta t}{\Delta_K} \right) (\delta_e h\mathbf{u}), \end{aligned} \quad (65)$$

where $(\delta_e h\mathbf{u}) = \frac{1}{2}(\mathbf{h}\mathbf{u}_{K_e} - \mathbf{h}\mathbf{u}_K) \cdot \vec{n}_{e,K}$. Note finally that we have $\tau^{-1} \mathcal{S}(\widehat{W}_K^n, \widehat{W}_{K_e}^n) = \begin{pmatrix} 0 \\ \phi_K \vec{n}_{e,K} \end{pmatrix}$ and $\tau^{-1} \mathcal{S}(\widehat{W}_K^n, \widehat{W}_{K_e}^n) = \begin{pmatrix} 0 \\ \phi_e^* \vec{n}_{e,K} \end{pmatrix}$. Putting this all together, we are ready to use formula (56) with $\alpha_{e,K} = \frac{m_e}{m_{\partial K}}$, which gives, with the support of Green formula:

$$\begin{aligned} W_K^{n+1} &= \sum_{e \in \partial K} \frac{m_e}{m_{\partial K}} \tau^{-1} (\widehat{W}_{e,K}^{n+1}) \\ &= W_K^n - \frac{\Delta t}{m_K} \sum_{e \in \partial K} \mathcal{F}(W_K^n, W_{K_e}^n, \vec{n}_{e,K}) m_e - \frac{\Delta t}{m_K} h_K^n \sum_{e \in \partial K} \begin{pmatrix} 0 \\ \phi_e^* \vec{n}_{e,K} \end{pmatrix} m_e. \end{aligned}$$

This leads to the scheme proposed in [1]:

$$\left\{ \begin{aligned} h_K^{n+1} &= h_K^n - \frac{\Delta t}{m_K} \sum_{e \in \partial K} (\mathbf{h}\mathbf{u})_e^* \cdot \vec{n}_{e,K} m_e \\ (\mathbf{h}\mathbf{u})_K^{n+1} &= (\mathbf{h}\mathbf{u})_K^n - \frac{\Delta t}{m_K} \sum_{e \in \partial K} (\mathbf{u}_K ((\mathbf{h}\mathbf{u})_e^* \cdot \vec{n}_{e,K})^+ + \mathbf{u}_{K_e} ((\mathbf{h}\mathbf{u})_e^* \cdot \vec{n}_{e,K})^-) m_e \\ &\quad - \frac{\Delta t}{m_K} h_K^n \sum_{e \in \partial K} \phi_e^* \vec{n}_{e,K} m_e, \end{aligned} \right. \quad (66)$$

where

$$(\mathbf{h}\mathbf{u})_e^* = (\overline{h\mathbf{u}})_e - \mathbf{\Pi}_e, \quad \mathbf{\Pi}_e = 2\gamma \left(\frac{\Delta t}{\Delta_K} \right) (\overline{h\mathbf{v}})_e (\delta_e \phi), \quad (67)$$

$$\phi_e^* = \bar{\phi}_e - \Lambda_e \quad , \quad \Lambda_e = 2\alpha g \left(\frac{\Delta t}{\Delta_K} \right) (\delta_e h \mathbf{u}) , \quad (68)$$

with $\delta_e \phi = \frac{1}{2}(\phi_{K_e} - \phi_K) \vec{n}_{e,K}$ and $(\delta_e h \mathbf{u}) = \frac{1}{2}((h \mathbf{u})_{K_e} - (h \mathbf{u}_K)) \cdot \vec{n}_{e,K}$. We now discuss the conditions ensuring fully discrete energy estimates.

4.2 Stability issues

Denoting $\eta(W) = h \underline{e}(W) = \frac{1}{2}gh^2 + ghb + \frac{1}{2}hu^2 + \frac{1}{2}hv^n$ the energy functional, the energy equation provided by (53) and resulting from (60) can be rewritten as:

$$\eta(\widehat{W}_{e,K}^{n+1}) \leq \eta(\widehat{W}_K^n) - \frac{\Delta t}{\Delta_K} \left(\mathcal{G}(\widehat{W}_K^n, \widehat{W}_{K_e}^n) - \mathcal{G}(\widehat{W}_K^n, \widehat{W}_K^n) \right) , \quad (69)$$

with \mathcal{G} a numerical entropy flux. These flux are defined according to the discrete operators involved in (53), which take the following form:

$$\partial_K \left(\underline{e} + \frac{1}{2}gh, (hu^\tau)^* \right) = \widetilde{\partial}_K^{up} \left(\frac{1}{2} \|\mathbf{u}\|^2, (hu^\tau)^* \right) + \widetilde{\partial}_K(\phi, (hu^\tau)^*) . \quad (70)$$

we recall here that the main part of these operators are given by (5) and (6), while the superscripts " \sim " refer to low order corrections (see Proposition 2.1, formula (20) and proof of Proposition 2.2 for an explicit expression of these terms). Since the quantity \underline{e}_K^n is invariant with the change of coordinates (59), we note that $\eta(\widehat{W}_{e,K}^{n+1}) = \eta(W_{e,K}^{n+1})$, and the consistency with the (1d) continuous entropy flux G allows to write:

$$\mathcal{G}(\widehat{W}_K^n, \widehat{W}_K^n) = G(\widehat{W}_K^n) = \left(\underline{e}(W_K^n) + \frac{1}{2}gh_K^n \right) (h \mathbf{u})_K^n \cdot \vec{n}_{e,K} , \quad (71)$$

which corresponds to the projection of the 2d entropy flux on the normal vector, namely $\mathbf{G}(W_K^n) \cdot \vec{n}_{e,K}$. On the other hand, returning to the discrete operators involved in (53), we can define the main part of the entropy fluxes as:

$$\mathcal{G}(W_K^n, W_{K_e}^n, \vec{n}_{e,K}) := \mathcal{G}(\widehat{W}_K^n, \widehat{W}_{K_e}^n) = \mathcal{F}_K^{up} \left(\frac{1}{2} \|\mathbf{u}\|^2, (h \mathbf{u})_e^* \cdot \vec{n}_{e,K} \right) + \phi_e (h \mathbf{u})_e^* \cdot \vec{n}_{e,K} + \widetilde{\mathcal{A}}_K , \quad (72)$$

where $\widetilde{\mathcal{A}}_K$ contains the lower order corrections (see (70) and discussion below). The convexity of η leads to:

$$\begin{aligned} \eta(W_K^{n+1}) &= \eta \left(\sum_{e \in \partial K} \alpha_{e,K} W_{e,K}^{n+1} \right) \\ &\leq \sum_{e \in \partial K} \alpha_{e,K} \eta(W_{e,K}^{n+1}) \\ &\leq \sum_{e \in \partial K} \alpha_{e,K} \left[\eta(W_K^n) - \frac{\Delta t}{\Delta_K} \left(\mathcal{G}(\widehat{W}_K^n, \widehat{W}_{K_e}^n) - \mathcal{G}(\widehat{W}_K^n, \widehat{W}_K^n) \right) \right] \\ &\leq \eta(W_K^n) - \frac{\Delta t}{m_K} \sum_{e \in \partial K} \mathcal{G}(W_K^n, W_{K_e}^n, \vec{n}_{e,K}) m_e . \end{aligned} \quad (73)$$

Note that this discrete energy estimate is ensured provided that inequality (69) holds, that is the stability conditions of the 1d-like scheme (60). According to Section 3, this is ensured by condition (43), which becomes in the present context:

$$4 \frac{\Delta t}{\Delta_K} (\mathbf{u}_K + \mathbf{c}_K) \leq 1, \quad (74)$$

where we recall that $\Delta_K = \frac{m_K}{m_{\partial K}}$, $\mathbf{c}_K = \sqrt{gh_K^n}$. Recalling that $(hu^\tau)_e^* = (h\mathbf{u})_K^* \cdot \vec{n}_{e,K}$, we take:

$$\mathbf{u}_K = \max_{e \in \partial K} \left(\frac{1}{2} (|(h\mathbf{u})_e^* \cdot \vec{n}_{e,K}| + |(h\mathbf{u})_K^n \cdot \vec{n}_{e,K}|) / h_K^n \right).$$

This allows to define the parameter ν in (67) as in the 1d case : $\nu_K = \frac{1}{1 - 4 \frac{\Delta t}{\Delta_K} \mathbf{u}_K}$. The

constants α, γ involved in (67) and (68) are still subject to the bound conditions (40) of Proposition 3.1

References

- [1] F. Couderc, A. Duran, and J.-P. Vila. An explicit asymptotic preserving low froude scheme for the multilayer shallow water model with density stratification. *Journal of Computational Physics*, 343:235–270, 2017.
- [2] M. Parisot and J.-P. Vila. Centered-potential regularization of advection upstream splitting method : Application to the multilayer shallow water model in the low Froude number regime. *SIAM Journal on Numerical Analysis*, 54:3083 – 3104, 2016.