HAL
open science

A large time-step and well-balanced Lagrange-Projection type scheme for the shallow-water equations

Christophe Chalons, Pierre Kestener, Samuel Kokh, Maxime Stauffert

To cite this version:

Christophe Chalons, Pierre Kestener, Samuel Kokh, Maxime Stauffert. A large time-step and well-balanced Lagrange-Projection type scheme for the shallow-water equations. 2016. hal-01297043v1

HAL Id: hal-01297043

https://hal.science/hal-01297043v1

Preprint submitted on 2 Apr 2016 (v1), last revised 2 Jul 2016 (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.

A large time-step and well-balanced Lagrange-Projection type scheme for the shallow-water equations

C. Chalons[1], P. Kestener[2], S. Kokh[2,3] and M. Stauffert[1,2]

Abstract

This work focuses on the numerical approximation of the Shallow Water Equations (SWE) using a Lagrange-Projection type approach. We propose to extend to this context the recent implicit-explicit schemes developed in [16], [18] in the framework of compressible flows, with or without stiff source terms. These methods enable the use of time steps that are no longer constrained by the sound velocity thanks to an implicit treatment of the acoustic waves, and maintain accuracy in the subsonic regime thanks to an explicit treatment of the material waves. In the present setting, a particular attention will be also given to the discretization of the non-conservative terms in SWE and more specifically to the well-known well-balanced property. We prove that the proposed numerical strategy enjoys important non linear stability properties and we illustrate its behaviour past several relevant test cases.

1 Introduction

We are interested in the design of a numerical scheme for the well-known Shallow Water Equations (SWE), given by

$$\begin{cases} \partial_t h + \partial_x (hu) = 0, & \text{(1a)} \\ \partial_t (hu) + \partial_x \left(hu^2 + g\dfrac{h^2}{2} \right) = -gh\partial_x z, & \text{(1b)} \end{cases}$$

where $z(x)$ denotes a given smooth topography and $g > 0$ is the gravity constant. The primitive variables are the water depth $h \geq 0$ and its velocity u, which both depend on the space and time variables, respectively $x \in \mathbb{R}$ and $t \in [0, \infty)$. At time $t = 0$, we assume that the initial water depth $h(x, t = 0) = h_0(x)$ and velocity $u(x, t = 0) = u_0(x)$ are given. In order to shorten the notations, we will use the following condensed form of (1), namely

$$\partial_t \mathbf{U} + \partial_x \mathbf{F}(\mathbf{U}) = \mathbf{S}(\mathbf{U}, z), \tag{2}$$

where $\mathbf{U} = (h, hu)^T$, $\mathbf{F}(\mathbf{U}) = (hu, hu^2 + gh^2/2)^T$ and $\mathbf{S}(\mathbf{U}, z) = (0, -gh\partial_x z)^T$. This system is supplemented with the validity of entropy inequalities which can be written either in a non-conservative form as follows,

$$\partial_t \mathcal{U}(\mathbf{U}) + \partial_x \mathcal{F}(\mathbf{U}) \leqslant -ghu\partial_x z, \tag{3}$$

with the non-conservative entropy $\mathcal{U}$ and the associated flux $\mathcal{F}$ defined by

$$\mathcal{U}(\mathbf{U}) = \frac{hu^2}{2} + \frac{gh^2}{2}, \quad \mathcal{F}(\mathbf{U}) = \left(\frac{u^2}{2} + gh \right) hu,$$

or in conservative form as follows,

$$\partial_t \tilde{\mathcal{U}}(\mathbf{U}, z) + \partial_x \tilde{\mathcal{F}}(\mathbf{U}, z) \leqslant 0, \tag{4}$$

where the conservative entropy $\tilde{\mathcal{U}}$ and the associated flux $\tilde{\mathcal{F}}$ now depend on z and are defined by,

$$\tilde{\mathcal{U}}(\mathbf{U}, z) = \mathcal{U}(\mathbf{U}) + ghz \quad \text{and} \quad \tilde{\mathcal{F}}(\mathbf{U}, z) = \mathcal{F}(\mathbf{U}) + ghuz.$$

[1]Laboratoire de Mathématiques de Versailles, UVSQ, CNRS, Université Paris-Saclay, 78035 Versailles, France.
[2]Maison de la Simulation USR 3441, Digiteo Labs, bât. 565, PC 190, CEA Saclay, 91191 Gif-sur-Yvette, France.
[3]CEA/DEN/DANS/DM2S/STMF - CEA Saclay, 91191 Gif-sur-Yvette, France.

The proposed numerical scheme should be consistent with (2) and should satisfy a discrete form of (at least) one of these entropy inequalities.

The steady states of (2) are governed by the ordinary differential system $\partial_x F(\mathbf{U}) = \mathbf{S}(\mathbf{U}, z)$, namely

$$hu = \text{constant}, \quad \frac{u^2}{2} + g(h+z) = \text{constant}.$$

In this paper, we will be more specifically interested in the so-called "lake at rest" steady solution defined by

$$h + z = \text{constant}, \quad u = 0. \tag{5}$$

The proposed numerical scheme should be able to preserve discrete initial conditions matching (5), which corresponds to the very well-known well-balanced property (see for instance the recent book [28] for a review).

A third objective of the method is to ensure the positivity of the water height if the initial water height is positive.

Last but not least, we are especially interested in this work in subsonic or near low-Froude number flows. In this case, it turns out that the usual CFL time step limitation of Godunov-type numerical schemes is driven by the acoustic waves and can thus be very restrictive. We are thus interested in the design of a mixed implicit-explicit large time-step strategy following the lines of the pionneering work [22] and the more recent ones [15], [16], [17], [18]. By large time-step, we mean that the scheme should be stable under a CFL stability condition driven by the (slow) material waves, and not by the (fast) acoustic waves as it is customary in Godunov-type schemes. Numerical evidences will show a gain in efficiency.

There is a huge amount of works about the design of numerical schemes for the SWE, and most of the schemes intended to satisfy the first three properties above. To mention only a few of them, we refer the reader to the following well-known contributions [3], [29], [27], [30], [32], [24], [25], [1], [21], [35], [11], [21], [10], [37], [36], [6], [7], [34], [2], [4], [5]. We also refer to the books [9] and [28] which provide additional references and very nice overviews.

The design of mixed implicit-explicit (IMEX) schemes based on a Lagrange-Projection type approach which are stable under a CFL restriction driven by the slow material waves and not the acoustic waves has been given a first interest in the pionneering work [22] and was further developed for the computation of large friction or low-Mach regimes in [15], [16], [17], [18], [19] for single or two-phase flow models. It is the purpose of this paper to adapt these IMEX strategies to the shallow-water equations while preserving the first three properties above, namely the lake-at-rest well-balanced property, the positivity of the water height, and the validity of a discrete form of the entropy inequality. Another new large time step method for the shallow water flows in the low Froude number limit has been proposed in [7]. The strategy is also mixed implicit-explicit considering the fast acoustic waves and the slow transport waves respectively, but does not rely on the natural Lagrange-Projection like decomposition proposed here. Note also that we focus here on subsonic or low Froude number flows, but we do not consider the low Froude number limit which is the purpose of a current work in progress. We also refer the reader to the recent contribution [38] which proves rigorously that the IMEX Lagrange projection scheme is AP for one-dimensional low-Mach isentropic Euler and low-Froude shallow water equations.

2 Operator splitting Lagrange-Projection approach and relaxation procedure

In this section we adapt the so-called operator splitting Lagrange-Projection strategy presented in [16] to the Shallow Water Equations (1). This splitting involves a so-called Lagrange step system that accounts for the acoustic waves and topography variations for which we shall propose an approximation

based on a Suliciu [33] relaxation approach using the notion of consistency in the integral sense [24, 25], and a so-called transport step accounting for the (slow) transport phenomenon.

Before describing the numerical method, we introduce classic notations pertaining to our discretization context. Space and time are discretized using a space step Δx and a time step Δt into a set of cells $[x_{j-1/2}, x_{j+1/2})$ and instants $t^{n+1} = n\Delta t$, where $x_{j+1/2} = j\Delta x$ and $x_j = (x_{j-1/2} + x_{j+1/2})/2$ are respectively the cell interfaces and cell centers, for $j \in \mathbb{Z}$ and $n \in \mathbb{N}$. For a given initial condition $x \mapsto \mathbf{U}^0(x)$, we consider a discrete initial data $\mathbf{U}_j^0$ defined by $\mathbf{U}_j^0 = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \mathbf{U}^0(x)\,\mathrm{d}x$, for $j \in \mathbb{Z}$. The algorithm proposed in this paper aims at computing an approximation $\mathbf{U}_j^n$ of $\frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \mathbf{U}(x, t^n)\,\mathrm{d}x$ where $x \to \mathbf{U}(x, t^n)$ is the exact solution of (1) at time t^n.

2.1 Acoustic/transport operator decomposition

We describe here a procedure that allows to approximate the evolution of the system (1) over a time interval $[t^n, t^n + \Delta t)$. The guideline of the method consists in decoupling the terms responsible for the acoustic, the topography variations and the transport phenomena. In the sequel if $p^{\mathrm{EOS}} : 1/h \mapsto gh^2/2$ we shall note $p = p^{\mathrm{EOS}}(1/h)$, and define the sound velocity c of (1) by $c^2 = \frac{\mathrm{d}}{\mathrm{d}h}[(p^{\mathrm{EOS}})(1/h)] = gh$. By using the chain rule for the space derivatives we split up the operators of system (1) so that it reads for smooth solutions

$$\begin{cases} \partial_t h + u\partial_x h + h\partial_x u = 0 \\ \partial_t(hu) + u\partial_x(hu) + hu\partial_x u + \partial_x p = -gh\partial_x z. \end{cases}$$

Consequently, we propose to approximate the solutions of (1) by approximating the solutions of the following two subsystems, namely

$$\begin{cases} \partial_t h + h\partial_x u = 0, & \text{(6a)} \\ \partial_t(hu) + hu\partial_x u + \partial_x p = -gh\partial_x z, & \text{(6b)} \end{cases}$$

and

$$\begin{cases} \partial_t h + u\partial_x h = 0, & \text{(7a)} \\ \partial_t(hu) + u\partial_x(hu) = 0, & \text{(7b)} \end{cases}$$

one after the other. System (6) deals with the acoustic effects and the topography variation, while system (7) involves the material transport. In the following we shall refer to (6) as the acoustic or Lagrangian system and (7) as the transport or projection system.

The overall algorithm can be described as follows: for a given discrete state $\mathbf{U}_j^n = (h, hu)_j^n$, $j \in \mathbb{Z}$ that describes the system at instant t^n, the update to the $\mathbf{U}_j^{n+1} = (h, hu)_j^{n+1}$ is a two-step process defined by

1. Update $\mathbf{U}_j^n$ to $\mathbf{U}_j^{n+1-}$ by approximating the solution of (6),
2. Update $\mathbf{U}_j^{n+1-}$ to $\mathbf{U}_j^{n+1}$ by approximating the solution of (7).

2.2 Relaxation approximation of the acoustic system

If we note $\tau = 1/h$ the specific volume, the acoustic system (6) takes the form

$$\begin{cases} \partial_t \tau - \tau\partial_x u = 0, & \text{(8a)} \\ \partial_t u + \tau\partial_x p = -g\partial_x z, & \text{(8b)} \\ \partial_t z = 0. & \text{(8c)} \end{cases}$$

It is straightforward to check that the quasilinear system (8) is strictly hyperbolic over the space $\{(\tau, u, z)^T \in \mathbb{R}^3 \mid \tau > 0\}$ and the eigenstructure of the system is composed by three fields associated

with the eigenvalues $\{-c, 0, c\}$. The wave associated with $\pm c$ (resp. 0) is genuinely nonlinear (resp. a stationary contact discontinuity). Let us underline that the material velocity u is not involved in the characteristic velocities of (8) but only the sound velocity c.

For $t \in [t^n, t^n+\Delta t)$, we propose to approximate $\tau(x,t)\partial_x \cdot$ by $\tau(x,t^n)\partial_x \cdot$ and $\partial_x z = \tau(x,t)\partial_x z/\tau(x,t)$ by $\tau(x,t^n)\partial_x z/\tau(x,t)$ in (8). If one introduces the mass variable m defined by $\mathrm{d}m(x) = \tau(x,t^n)^{-1}\mathrm{d}x$, up to a slight abuse of notations, system (8) can be recast into

$$\begin{cases} \partial_t \tau - \partial_m u = 0, & \text{(9a)} \\ \partial_t u + \partial_m p = -g\partial_m z/\tau, & \text{(9b)} \\ \partial_t z = 0. & \text{(9c)} \end{cases}$$

Let us note that, when the topography is flat, system (9) is consistent with the usual form of the barotropic gas dynamics equations in Lagrangian coordinates with a mass space variable (see for instance [26]).

We carry on with the approximation process of the acoustic system (6) by using a Suliciu-type relaxation approximation of (9), see [33]. We will see in the sequel that this strategy will allow us to design a simple and not expensive time implicit treatment of (6) in order to remove the usual CFL restriction associated with the fast acoustic waves $\pm c$. The design principle of the so-called pressure relaxation methods is now very well-known, see for instance [33, 31, 23, 12, 14, 8, 9] and the references therein and consists in introducing a larger system with linearly degenerate characteristic fields so that the underlying Riemann problem is easy to solve. To do so, we introduce a new independent variable pressure Π that can be seen as a linearization of the pressure p. While the pressure p verifies $\partial_t p + (c/\tau)^2\partial_m u = 0$ when τ and u are smooth solutions of (8), the surrogate pressure Π is evolved according to its own partial differential equation. Within the time interval $t \in [t^n, t^n+\Delta t)$, we propose to consider the following relaxation system

$$\begin{cases} \partial_t \tau - \partial_m u = 0, & \text{(10a)} \\ \partial_t u + \partial_m \Pi = -g\partial_m z/\tau, & \text{(10b)} \\ \partial_t \Pi + a^2\partial_m u = \lambda(p(\mathcal{T}) - \Pi), & \text{(10c)} \\ z_t = 0, & \text{(10d)} \end{cases}$$

where $a > 0$ is a constant whose choice will be specified later on, $\lambda > 0$ is the relaxation parameter, and $\mathcal{T}$ obeys the well-defined (under appropriate conditions on a) implicit relation

$$\Pi = p(\mathcal{T}) + a^2(\mathcal{T} - \tau).$$

System (10) is indeed an approximation of (9) in the sense that in the asymptotic regime $\lambda \to +\infty$ we have, at least formally that $\Pi \to p$ and we recover (9), see [14] for a rigourous proof for both smooth and discontinuous solutions. Let us also briefly recall that this relaxation model can be endowed with a relaxation entropy defined by

$$\Sigma = \frac{u^2}{2} - \int^{\mathcal{T}} p^{EOS}(\tau)d\tau + \frac{\Pi^2 - (p^{EOS})^2(\mathcal{T})}{2\,a^2}\,, \tag{11}$$

which is such that $h\Sigma$ coincides with the entropy $\mathcal{U}$ at equlibrium $\mathcal{T} = \tau$, and by the chain rule and for smooth solutions easily satisfies

$$\partial_t \Sigma + \partial_m \Pi u = -\frac{\lambda}{a^2}(p(\mathcal{T}) - \Pi)^2 - g\frac{u}{\tau}\partial_m z, \tag{12}$$

which is nothing but a relaxation and Lagrangian form of (3). Note that the first term of the right-hand side is negative so that the proposed relaxation process is entropy consistent in the sense of [20].

We adopt the classic method that allows to reach the $\lambda \to \infty$: at each time step, we enforce the equilibrium relation $\Pi_i^n = p^{\mathrm{EOS}}(\tau_i^n)$ and solve (10) with $\lambda = 0$. In order to prevent this relaxation procedure from generating instabilities, it is now well established that a must be chosen sufficiently large in agreement with the Whitham subcharacteristic condition

$$a > \max\left(c(\tau)/\tau\right), \tag{13}$$

when τ spans the values of the solution of (10) for $t \in [t^n, t^n + \Delta t)$ (see again the above references). For $\lambda = 0$, system (10) can take the compact form

$$\partial_t \mathbf{W} + \partial_m \mathbf{G}(\mathbf{W}) = \left(-\frac{g}{\tau}\partial_m z\right)\mathbf{E}_2, \tag{14}$$

where $\mathbf{W} = (\tau, u, \Pi, z)^T$, $\mathbf{G}(\mathbf{W}) = (-u, \Pi, a^2 u, 0)^T$, $\mathbf{E}_2 = (0, 1, 0, 0)^T$. Let us discuss a few properties of (14). First, it can be easily proved that (14) is strictly hyperbolic and involves four linearly degenerate characteristic fields associated with the characteristic velocities $\{-a, 0, +a\}$ that are nothing but approximations of the eigenvalues of (9). The jump relations involved with each field are detailed in appendix A. The non-conservative product that features in (9) is well defined for smooth z under consideration here.

Before going any further, let us observe that (14) can be recast into the following equivalent form

$$\begin{cases} \partial_t \tau - \partial_m u = 0, & \text{(15a)} \\ \partial_t \overrightarrow{w} + a\partial_m \overrightarrow{w} = -ag\partial_m z/\tau & \text{(15b)} \\ \partial_t \overleftarrow{w} - a\partial_m \overleftarrow{w} = +ag\partial_m z/\tau, & \text{(15c)} \\ \partial_t z = 0 & \text{(15d)} \end{cases}$$

where the new variables $\overrightarrow{w}$ and $\overleftarrow{w}$ are defined by $\overrightarrow{w} = \Pi + au$, $\overleftarrow{w} = \Pi - au$. These quantities are nothing but the strong Riemann invariants associated with the characteristic speeds $\pm a$ of the relaxation system (15) when the topography terms are omitted. The closure relations for (15) are naturally given by

$$u = \frac{\overrightarrow{w} - \overleftarrow{w}}{2a}, \quad \Pi = \frac{\overrightarrow{w} + \overleftarrow{w}}{2}.$$

This new formulation will be used in the sequel to study a time-implicit discretization of (14).

We now need to propose a discretization strategy for (14). Unfortunately, the classic relaxation solver strategy cannot be carried on here since the solution of the Riemann problem associated with (14) cannot be defined easily. Indeed it is not possible to properly define the non-conservative term $\partial_m z/\tau$ with a piecewise constant initial value for z. However we will see in the next section that it is possible to derive an approximate Riemann solver for (14) using a discretization of the non-conservative product that is consistent (in a sense to be specified later) with the smooth term $(g/\tau)\partial_m z$.

2.3 Approximate Riemann solver for the acoustic system

Let $\Delta m_L > 0$, $\Delta m_R > 0$ and suppose given a smooth function $m \mapsto z(m)$. If $\overline{m} \in \mathbb{R}$, we consider a piecewise initial data defined by

$$\mathbf{W}(m, t = 0) = \begin{cases} \mathbf{W}_L = (\tau_L, u_L, \Pi_L, z_L)^T, & \text{if } m \leq \overline{m}, \\ \mathbf{W}_R = (\tau_R, u_R, \Pi_R, z_R)^T, & \text{if } m > \overline{m}, \end{cases} \tag{16}$$

where Π_k and z_k, $k = L, R$ are defined by

$$\Pi_L = p^{\mathrm{EOS}}(\tau_L), \quad \Pi_R = p^{\mathrm{EOS}}(\tau_R), \quad z_L = \frac{1}{\Delta m_L}\int_{-\Delta m_L}^{0} z(\overline{m} + r)\mathrm{d}r, \quad z_R = \frac{1}{\Delta m_R}\int_{0}^{\Delta m_R} z(\overline{m} + r)\mathrm{d}r.$$

Note that Π_L and Π_R are at equilibrium. Let us now build an approximate Riemann solver for the relaxed acoustic system (14). We seek for a function $\mathbf{W}_{\mathrm{RP}}$ composed by four states separated by discontinuities as follows

$$\mathbf{W}_{\mathrm{RP}}\left(\frac{m-\overline{m}}{t};\mathbf{W}_L,\mathbf{W}_R\right)=\begin{cases}\mathbf{W}_L, & \text{if } \frac{m-\overline{m}}{t}\leq -a,\\ \mathbf{W}_L^*, & \text{if } -a<\frac{m-\overline{m}}{t}\leq 0,\\ \mathbf{W}_R^*, & \text{if } 0<\frac{m-\overline{m}}{t}\leq a,\\ \mathbf{W}_R, & \text{if } a<\frac{m-\overline{m}}{t},\end{cases} \tag{17}$$

where the intermediate states are such that the following three consistency properties hold true (see [24], [25]):

a) $\mathbf{W}_{\mathrm{RP}}$ is consistent in the integral sense with the Shallow-Water Equations, more specifically in our context: if Δt is such that $a\Delta t\leq \min(\Delta m_L,\Delta m_R)/2$, then

$$\mathbf{G}(\mathbf{W}_R)-\mathbf{G}(\mathbf{W}_L)=-a(\mathbf{W}_L^*-\mathbf{W}_L)+a(\mathbf{W}_R-\mathbf{W}_R^*)-\frac{\Delta m_L+\Delta m_R}{2}\left\{\frac{g}{\tau}\partial_m z\right\}, \tag{18}$$

where $\left\{\frac{g}{\tau}\partial_m z\right\}$ is consistent with the influence of the source term, in the sense that

$$\lim_{\substack{\Delta m_L,\Delta m_R\to 0\\ \mathbf{W}_L,\mathbf{W}_R\to(\overline{\tau},\overline{u},\overline{\Pi},z(\overline{m}))}}\left\{\frac{g}{\tau}\partial_m z\right\}=\frac{g}{\overline{\tau}}(\partial_m z)(\overline{m}); \tag{19}$$

b) In the case of constant bottom, i.e. $z_L=z_R$, $\mathbf{W}_{\mathrm{RP}}$ must degenerate towards the classic solution of the Riemann problem of the acoustic relaxed system (14) for a flat bottom $\partial_m z=0$;

c) If $\mathbf{W}_L$ and $\mathbf{W}_R$ verify the lake at rest condition

$$u_L=u_R=0,\qquad 1/\tau_L+z_L=1/\tau_R+z_R, \tag{20}$$

then $\mathbf{W}_L^*=\mathbf{W}_L$ and $\mathbf{W}_R^*=\mathbf{W}_R$.

Defining a proper function $\mathbf{W}_{\mathrm{RP}}$ thus simply boils down to proposing intermediate states $\mathbf{W}_L^*$ and $\mathbf{W}_R^*$ that comply with a), b) and c). We proceed as follows: first we impose that $\mathbf{W}_L$ and $\mathbf{W}_L^*$ (resp. $\mathbf{W}_R^*$ and $\mathbf{W}_R$) verify the jump conditions

$$a(\mathbf{W}_L^*-\mathbf{W}_L)+\mathbf{G}(\mathbf{W}_L^*)-\mathbf{G}(\mathbf{W}_L)=0,\quad -a(\mathbf{W}_R-\mathbf{W}_R^*)+\mathbf{G}(\mathbf{W}_R)-\mathbf{G}(\mathbf{W}_R^*)=0. \tag{21}$$

This amounts to say that the discontinuity of velocity $\pm a$ of $\mathbf{W}_{\mathrm{RP}}$ behaves like the $(\pm)a$-wave of system (14) for a flat bottom. Similarly, across the discontinuity of velocity 0 we impose that

$$u_L^*=u_R^*=:u^*. \tag{22}$$

Relations (21) and (22) does not provide enough information to determine the intermediate states $\mathbf{W}_L^*$ and $\mathbf{W}_R^*$. Indeed, they provide us with only seven independent relations while we aim at defining eight quantities, namely the four components of each $\mathbf{W}_L^*$ and $\mathbf{W}_R^*$.

We choose to add another jump relation across the stationary discontinuity of $\mathbf{W}_{\mathrm{RP}}$ that complies with condition b): we impose that

$$\Pi_R^*-\Pi_L^*+\mathscr{M}=0, \tag{23}$$

where $\mathscr{M}$ is a function to be specified such that $\mathscr{M}=0$ if $z_L=z_R$. Relations (14), (22) and (23) allow

to solve for $\mathbf{W}_L^*$ and $\mathbf{W}_R^*$ and we obtain

$$\begin{cases} \tau_L^* = \tau_L + \dfrac{1}{2a}\left(u_R - u_L\right) - \dfrac{1}{2a^2}\left(\Pi_R - \Pi_L\right) - \dfrac{\mathscr{M}}{2a^2}, \\ \tau_R^* = \tau_R + \dfrac{1}{2a}\left(u_R - u_L\right) + \dfrac{1}{2a^2}\left(\Pi_R - \Pi_L\right) + \dfrac{\mathscr{M}}{2a^2}, \\ u^* = u_R^* = u_L^* = \dfrac{u_R + u_L}{2} - \dfrac{1}{2a}\left(\Pi_R - \Pi_L\right) - \dfrac{\mathscr{M}}{2a}, \\ \Pi^* = \dfrac{\Pi_R + \Pi_L}{2} - \dfrac{a}{2}\left(u_R - u_L\right), \\ \Pi_L^* = \Pi^* + \dfrac{\mathscr{M}}{2}, \\ \Pi_R^* = \Pi^* - \dfrac{\mathscr{M}}{2}, \\ z_L^* = z_L, \\ z_R^* = z_R. \end{cases} \tag{24}$$

We now only need to determine $\mathscr{M}$ such that conditions a), b) and c) are satisfied. It is straightforward to see that the integral consistency requirement of condition a) implies by (18) that

$$\mathscr{M} = \left\{\frac{g}{\tau}\partial_m z\right\}\frac{\Delta m_L + \Delta m_R}{2}. \tag{25}$$

A simple mean to comply with conditions a) and b) is to choose

$$\mathscr{M} = \frac{g}{\tau_\Delta(\mathbf{W}_L, \mathbf{W}_R)}(z_R - z_L), \tag{26}$$

where $\tau_\Delta(\mathbf{W}_L, \mathbf{W}_R)$ has to be chosen such that $\tau_\Delta(\mathbf{W}_L, \mathbf{W}_R) \to \overline{\tau}$ if $\tau_L, \tau_R \to \overline{\tau}$. At last, we need to ensure condition c): if we have (20), then $\mathbf{W}_L^* = \mathbf{W}_L$ and $\mathbf{W}_R = \mathbf{W}_R^*$ imply that

$$\frac{1}{\tau_\Delta(\mathbf{W}_L, \mathbf{W}_R)} = \frac{1}{2}\left(\frac{1}{\tau_L} + \frac{1}{\tau_R}\right). \tag{27}$$

As a conclusion, we choose to adopt (27) as a definition of τ_Δ for any $\mathbf{W}_L$ and $\mathbf{W}_R$. This yields that

$$\left\{\frac{g}{\tau}\partial_m z\right\}(\mathbf{W}_L, \mathbf{W}_R, \Delta m_L, \Delta m_R) = \frac{g}{\tau_\Delta(\mathbf{W}_L, \mathbf{W}_R)}\frac{2}{\Delta m_L + \Delta m_R}(z_R - z_L). \tag{28}$$

It is then straightforward to check that the approximate Riemann solver defined by (24) and (28) verifies the three conditions a), b) and c). We sump up in the following proposition the properties of our Riemann solver.

Proposition 1. *Consider the approximate Riemann solver $\boldsymbol{W}_{RP}$ defined by* (17), (24), (26) *and* (27).

(i) $\boldsymbol{W}_{RP}$ is consistent in the integral sense with the Shallow-Water Equations (1).

(ii) In the case of a constant bottom $z_L = z_R$, $\boldsymbol{W}_{RP}$ degenerates to a classic approximate Riemann solver for the barotropic Euler equations in Lagrange coordinates.

(iii) If $\boldsymbol{W}_L$ and $\boldsymbol{W}_R$ verify the lake at rest relation (20)*, then $\boldsymbol{W}_k^* = \boldsymbol{W}_k$, $k = L, R$.*

3 Numerical method

In this section, we now give the details of the two-step process proposed in Section 2.1 for solving the Shallow Water Equations. Let us briefly recall that this two-step process is defined by

1. Update $\mathbf{U}_j^n$ to $\mathbf{U}_j^{n+1-}$ by approximating the solution of (6),

2. Update $\mathbf{U}_j^{n+1-}$ to $\mathbf{U}_j^{n+1}$ by approximating the solution of (7).

In the sequel we shall note $\Delta m_j = \Delta x_j h_j^n$, $\Delta m_{j+1/2} = (\Delta m_j + \Delta m_{j+1})/2$, and if we assume as given the approximate solution $\{\mathbf{U}_j^n\}_j$ at time t^n, we introduce the approximate solution $\{\mathbf{W}_j^n\}_j$ at equilibrium in the $\mathbf{W}$ variable with a clear and natural definition. We begin with a fully explicit discretization of the Shallow Water Equations, which means that both steps of the process are solved with a time-explicit procedure, and we will go on with a mixed implicit-explicit strategy for which the solutions of (6) are solved implicitly in time and the solutions of (7) are solved explicitly. The latter strategy allows to get rid of the strong CFL restriction coming from the acoustic waves in the subsonic regime and corresponds to the very motivation of the present study.

3.1 Time-explicit discretization

Let us begin with the time-explicit discretization of the acoustic system (6), or equivalently (8).

Acoustic step. The acoustic update is achieved thanks to the proposed relaxation approximation and the corresponding approximate Riemann solver detailed in Section 2.3. More precisely, we propose to simply use a Godunov-type method based on this approximate Riemann solver. As it is customary and starting from the piecewise constant initial data defined by the sequence $\{\mathbf{W}_j^n\}_j$, it consists in averaging after a Δt-long time evolution, the juxtaposition of the approximate Riemann solutions defined locally at each interface $x_{j+1/2}$. Following the same lines as in [16] and [18], see also [24], [25], [13] and the references therein, this update procedure can be easily expressed as follows after simple calculations,

$$\begin{cases} \tau_j^{n+1-} = \tau_j^n + \dfrac{\Delta t}{\Delta m_j}(u^*_{j+1/2} - u^*_{j+1/2}), & \text{(29a)} \\ u_j^{n+1-} = u_j^n - \dfrac{\Delta t}{\Delta m_j}(\Pi^*_{j+1/2} - \Pi^*_{j+1/2}) - \Delta t \left\{ \dfrac{g}{\tau}\partial_m z \right\}_j^n, & \text{(29b)} \\ \Pi_j^{n+1-} = \Pi_j^n - \dfrac{\Delta t}{\Delta m_j} a^2 (u^*_{j+1/2} - u^*_{j+1/2}). & \text{(29c)} \end{cases}$$

where $\Pi_j^n = p^{\mathrm{EOS}}(\tau_j^n)$ and

$$\begin{cases} u^*_{j+1/2} = \dfrac{1}{2}(u_j^n + u_{j+1}^n) - \dfrac{1}{2a}(\Pi_{j+1}^n - \Pi_j^n) - \dfrac{\Delta m_{j+1/2}}{2a}\left\{ \dfrac{g}{\tau}\partial_m z \right\}_{j+1/2}^n, & \text{(30a)} \\ \Pi^*_{j+1/2} = \dfrac{1}{2}(\Pi_j^n + \Pi_{j+1}^n) - \dfrac{a}{2}(u_{j+1}^n - u_j^n), & \text{(30b)} \\ \left\{ \dfrac{g}{\tau}\partial_m z \right\}_{j+1/2}^n = \left\{ \dfrac{g}{\tau}\partial_m z \right\}(\mathbf{W}_j^n, \mathbf{W}_j^n, \Delta m_j, \Delta m_{j+1}) = \dfrac{g}{2}\left(\dfrac{1}{\tau_j^n} + \dfrac{1}{\tau_{j+1}^n} \right) \dfrac{z_{j+1} - z_j}{\Delta m_{j+1/2}}, & \text{(30c)} \\ \left\{ \dfrac{g}{\tau}\partial_m z \right\}_j^n = \dfrac{1}{2}\left(\dfrac{\Delta m_{j+1/2}}{\Delta m_j}\left\{ \dfrac{g}{\tau}\partial_m z \right\}_{j+1/2}^n + \dfrac{\Delta m_{j-1/2}}{\Delta m_j}\left\{ \dfrac{g}{\tau}\partial_m z \right\}_{j-1/2}^n \right). & \text{(30d)} \end{cases}$$

If we focus now on the conservative variable $\mathbf{U} = (h, hu)$, the discretization (29) yields the following formula for the update sequence $\{\mathbf{U}_j^{n+1-}\}_j$, namely

$$\begin{cases} L_j h^{n+1-} = h_j^n, & \text{(31a)} \\ L_j (hu)^{n+1-} = (hu)_j^n - \dfrac{\Delta t}{\Delta x_j}(\Pi^*_{j+1/2} - \Pi^*_{j-1/2}) - \Delta t\, h_j^n \left\{ \dfrac{g}{\tau}\partial_m z \right\}_j^n, & \text{(31b)} \\ L_j = 1 + \dfrac{\Delta t}{\Delta x_j}(u^*_{j+1/2} - u^*_{j-1/2}). & \text{(31c)} \end{cases}$$

Let us now continue with the discretization of the transport equations (7).

Transport step. Denoting $\varphi \in \{h, hu\}$ and following again the same lines of [16] and [18], see again also [22], we use a standard time-explicit upwind discretization for the transport step by setting

$$\varphi_j^{n+1} = \varphi_j^{n+1-} - \frac{\Delta t}{\Delta x_j}(u^*_{j+1/2}\varphi^{n+1-}_{j+1/2} - u^*_{j-1/2}\varphi^{n+1-}_{j-1/2}) + \frac{\Delta t}{\Delta x_j}\varphi_j^{n+1-}(u^*_{j+1/2} - u^*_{j-1/2}), \tag{32}$$

where

$$\varphi^{n+1-}_{j+1/2} = \begin{cases} \varphi_j^{n+1-}, \text{ if } u^*_{j+1/2} \geq 0, \\ \varphi_{j+1}^{n+1-}, \text{ if } u^*_{j+1/2} < 0. \end{cases}$$

Let us note that the transport update (32) equivalently reads

$$\varphi_j^{n+1} = \varphi_j^{n+1-} L_j + \frac{\Delta t}{\Delta x_j}\left(u^*_{j+1/2}\varphi^{n+1-}_{j+1/2} - u^*_{j-1/2}\varphi^{n+1-}_{j-1/2}\right), \tag{33}$$

and that the interface value of the velocity $u^*_{j+1/2}$ coincides with the one proposed in the first step, which is actually crucial in order for the whole scheme to be conservative. The next statement gather the main properties satisfied by our explicit in time and two-step algorithm.

Overall Discretization. After injecting (31c) into (33) one obtains the complete update procedure from t^n to t^{n+1}. For the conservative variables it reads

$$\begin{cases} h_j^{n+1} = h_j^n + \frac{\Delta t}{\Delta x_j}\left(u^*_{j+1/2}h^{n+1-}_{j+1/2} - u^*_{j-1/2}h^{n+1-}_{j-1/2}\right), \\ (hu)_j^{n+1} = (hu)_j^n + \frac{\Delta t}{\Delta x_j}\Big[u^*_{j+1/2}(hu)^{n+1-}_{j+1/2} + \Pi^*_{j+1/2} - u^*_{j-1/2}(hu)^{n+1-}_{j-1/2} - \Pi^*_{j-1/2}\Big] + \Delta t h_j^n \left\{\frac{g}{\tau}\partial_m z\right\}_j^n. \end{cases} \tag{34}$$

We sum up the properties of our explicit scheme (29)-(32) in the following proposition.

Proposition 2. *The fully explicit scheme* (29)-(32) *satisfies the following:*

(i) it is a conservative scheme for the water height h. It is also a conservative scheme for hu when the topography source term vanishes.

Under the Whitham subcharacteristic condition and the Courant-Friedrichs-Lewy (CFL) conditions,

$$\max_j \frac{\Delta t}{h_j^n \Delta x_j} \leq \frac{1}{2a} \quad \textit{and} \quad \max_j \frac{\Delta t}{\Delta x_j}\Big((u^*_{j-\frac{1}{2}})^+ - (u^*_{j+\frac{1}{2}})^-\Big) < 1, \tag{35}$$

(ii) the water height h_j^n is positive for all j and $n > 0$ provided that h_j^0 is positive for all j,

(iii) it is well-balanced, with respect to the lake at rest condition (5)*,*

(iv) it degenerates to the classic Lagrange-Projection scheme when the bottom is flat.

Proof.

(i) This is a straightforward consequence of (34).

(ii) Thanks to (31c) and (31a), the CFL condition (35) ensures that $h_j^{n+1-} > 0$ for $j \in \mathbb{Z}$. The CFL condition (35) yields that h_j^{n+1} is a convex combination of $(h_k^{n+1-})_{k=j\pm1,j}$ and therefore $h_j^{n+1} > 0$.

(iii) Consider a discrete fluid state at instant t^n that matches the lake at rest condition, namely: $u_j^n = 0$, $h_j^n + z_j = h_{j+1}^n + z_{j+1}$, for all $j \in \mathbb{Z}$. Thanks to the condition c) verified by the approximated Riemann solver of the acoustic step, we know that $u_j^{n+1-} = 0$, $h_j^{n+1-} + z_j = h_{j+1}^{n+1-} + z_{j+1}$, for all $j \in \mathbb{Z}$. And thus, the transport step (32) boils down to $h_j^{n+1} = h_j^{n+1-}$ and $u_j^{n+1} = 0$.

(iv) This is consequence of condition b) imposed on the approximate Riemann solver for the acoustic step. □

Remark. Following the theory proposed by Gallice in [24] and [25] for non conservative systems with source terms, it is also possible to prove that our time-explicit Godunov-type scheme based on the definition of a consistent approximate Riemann solver satisfies a discrete version of the non conservative entropy (3) under additional assumptions on the intermediate states and the propagation speed a. We refer for instance the reader to [2] and [4] for detailed calculations. Note that we are not able to prove at present that the scheme satisfies a discrete version of the conservative entropy inequality (4).

3.2 Implicit in time Lagrange-Projection method

Let us now consider the ultimate algorithm of this paper, which consists in considering a time-implicit scheme for the Lagrangian step and keeping unchanged the transport step. As we will see in the next theorem, this strategy will allow us to obtain a non linearly stable algorithm under a CFL restriction based on the material velocity u and not on the sound velocity c. In order to derive a time-implicit scheme for the Lagrangian step, we follow the following standard approach where the numerical fluxes are now evaluated at time t^{n+1-}, which gives here the same update formulas as in the explicit case which are

$$\begin{cases} \tau_j^{n+1-} = \tau_j^n + \dfrac{\Delta t}{\Delta m_j}(u^*_{j+1/2} - u^*_{j+1/2}), & \text{(36a)} \\ u_j^{n+1-} = u_j^n - \dfrac{\Delta t}{\Delta m_j}(\Pi^*_{j+1/2} - \Pi^*_{j+1/2}) - \Delta t \left\{ \dfrac{g}{\tau}\partial_m z \right\}_j^n, & \text{(36b)} \\ \Pi_j^{n+1-} = \Pi_j^n - \dfrac{\Delta t}{\Delta m_j} a^2 (u^*_{j+1/2} - u^*_{j+1/2}), & \text{(36c)} \end{cases}$$

but where the numerical fluxes now involve quantities at time t^{n+1-} apart from the term consistent with $\left\{ \frac{g}{\tau}\partial_m z \right\}$, which writes

$$\begin{cases} u^*_{j+1/2} = \dfrac{1}{2}(u_j^{n+1-} + u_{j+1}^{n+1-}) - \dfrac{1}{2a}(\Pi_{j+1}^{n+1-} - \Pi_j^{n+1-}) - \dfrac{\Delta m_{j+1/2}}{2a} \left\{ \dfrac{g}{\tau}\partial_m z \right\}_{j+1/2}^n, & \text{(37a)} \\ \Pi^*_{j+1/2} = \dfrac{1}{2}(\Pi_j^{n+1-} + \Pi_{j+1}^{n+1-}) - \dfrac{a}{2}(u_{j+1}^{n+1-} - u_j^{n+1-}), & \text{(37b)} \end{cases}$$

with $\left\{ \frac{g}{\tau}\partial_m z \right\}_j^n$ and $\left\{ \frac{g}{\tau}\partial_m z \right\}_{j+1/2}^n$ given by (30c) and (30d).

Let us observe that we suggest here to keep on evaluating the topography source term at time t^n. This choice is motivated by the fact that this implicit system to be solved turns out to be a linear system with a significantly reduced coupling of the variables. More precisely, it is interesting to see that it is equivalent to the following one written in characteristic variables, namely

$$\begin{cases} \tau_j^{n+1-} = \tau_j^n + \dfrac{\Delta t}{\Delta m_j}\left(u^*_{j+1/2} - u^*_{j-1/2} \right), \\ \overrightarrow{w}_j^{n+1-} = \overrightarrow{w}_j^n - a\dfrac{\Delta t}{\Delta m_j}\left(\overrightarrow{w}_j^{n+1-} - \overrightarrow{w}_{j-1}^{n+1-} \right) - a\Delta t \dfrac{\Delta m_{j-1/2}}{\Delta m_j} \left\{ \dfrac{g}{\tau}\partial_m z \right\}_{j-1/2}^n, \\ \overleftarrow{w}_j^{n+1-} = \overleftarrow{w}_j^n + a\dfrac{\Delta t}{\Delta m_j}\left(\overleftarrow{w}_{j+1}^{n+1-} - \overleftarrow{w}_j^{n+1-} \right) + a\Delta t \dfrac{\Delta m_{j+1/2}}{\Delta m_j} \left\{ \dfrac{g}{\tau}\partial_m z \right\}_{j+1/2}^n, \\ z_j^{n+1} = z_j^n, \end{cases} \tag{38}$$

where of course $u^*_{j+1/2}$ means here $u^{*,n+1-}_{j+1/2}$ (the notation has been lightened for the sake of clarity). Notice that the coupling between the four variables is actually weak in (38) since we can easily first solve the linear system given by the second and the third equations, which are nothing but

$$\begin{cases} \left(I_N + a\Delta t\, A_+^n \right) \overrightarrow{w}^{n+1-} = \overrightarrow{w}^n - a\Delta t\, b_+^n, \\ \left(I_N - a\Delta t\, A_-^n \right) \overleftarrow{w}^{n+1-} = \overleftarrow{w}^n + a\Delta t\, b_-^n, \end{cases}$$

where we have set

$$A_+^n = \begin{pmatrix} \frac{1}{\Delta m_1} & 0 & \cdots & 0 \\ \frac{-1}{\Delta m_2} & \frac{1}{\Delta m_2} & \ddots & \vdots \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & \frac{-1}{\Delta m_N} & \frac{1}{\Delta m_N} \end{pmatrix}, \qquad b_+^n = g \cdot \begin{pmatrix} \frac{1}{\Delta m_1} \frac{h_1^n + h_0^n}{2} (z_1 - z_0) \\ \frac{1}{\Delta m_2} \frac{h_2^n + h_1^n}{2} (z_2 - z_1) \\ \vdots \\ \frac{1}{\Delta m_N} \frac{h_N^n + h_{N-1}^n}{2} (z_N - z_{N-1}) \end{pmatrix},$$

and

$$A_-^n = \begin{pmatrix} \frac{-1}{\Delta m_1} & \frac{1}{\Delta m_1} & 0 & 0 \\ 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \frac{-1}{\Delta m_{N-1}} & \frac{1}{\Delta m_{N-1}} \\ 0 & \vdots & 0 & \frac{-1}{\Delta m_N} \end{pmatrix}, \qquad b_-^n = g \cdot \begin{pmatrix} \frac{1}{\Delta m_1} \frac{h_2^n + h_1^n}{2} (z_2 - z_1) \\ \vdots \\ \frac{1}{\Delta m_{N-1}} \frac{h_N^n + h_{N-1}^n}{2} (z_N - z_{N-1}) \\ \frac{1}{\Delta m_N} \frac{h_{N+1}^n + h_N^n}{2} (z_{N+1} - z_N) \end{pmatrix}.$$

Let us of course notice that a few coefficients of the matrices A_+^n and A_-^n, and vectors b_+^n and b_-^n might be modified depending on the boundary conditions, but the purpose is to highlight that the characteristic variables $\overleftarrow{w}$ and $\overrightarrow{w}$ can be solved independently. Once this is done, the τ variable can be updated explicitly since $u_{j+1/2}^*$, or let us say $u_{j+1/2}^{*,n+1-}$, is explicitly known from the knowledge of $\overleftarrow{w}^{n+1-}$ and $\overrightarrow{w}^{n+1-}$ by the formulas

$$u_j^{n+1-} = \frac{1}{2a}(\overrightarrow{w}_j^{n+1-} - \overleftarrow{w}_j^{n+1-}), \quad \Pi_j^{n+1-} = \frac{1}{2}(\overrightarrow{w}_j^{n+1-} + \overleftarrow{w}_j^{n+1-}).$$

At last, notice that the matrices $\left(I_N + a\Delta t\, A_+^n\right)$ and $\left(I_N - a\Delta t\, A_-^n\right)$ are clealry triangular with positive diagonal coefficients, so that the system (38) has a unique solution whatever the time step $\Delta t > 0$ is.

It is quite natural at this stage to wonder whether the proposed time-implicit treatment of the Lagrangian step is well-balanced, which was true for the time-explicit version and was the key property leading to the well-balanced property of the global Explicit-Explicit Lagrange-Projection scheme in the previous section. It is the purpose of the next lemma.

Lemma 1. *Under the assumption of the lake at rest at the initial time, i.e. :*

$$\forall j \in \{1, \ldots, N\}, \ \begin{cases} u_j^0 = 0, \\ h_j^0 + z_j^0 = \textit{constant}, \end{cases}$$

the implicit scheme for the Lagrangian step keeps this initial state unchanged, which means that the time-implicit Lagrangian step as well as the global Implicit-Explicit Lagrange-Projection scheme is still well-balanced.

Proof. Under the assumption of the lake at rest, and thanks to the initialisation of the relaxation pressure, namely $\forall j \in \{1, \ldots, N\}$, $\Pi_j^0 = \frac{g}{2}\left(h_j^0\right)^2$, we get

$$g\frac{h_j^0 + h_{j-1}^0}{2}(z_j - z_{j-1}) = -\frac{g}{2}\left((h_j^0)^2 - (h_{j-1}^0)^2\right) = -(\Pi_j^0 - \Pi_{j-1}^0).$$

Thus one can write

$$b_+^0 = -A_+^0 \Pi^0,$$

and, in this special case where $u^0 \equiv 0$,

$$\left(I_N + a\Delta t\, A_+^n\right)(\Pi + au)^{1-} = \overrightarrow{w}^0 - a\Delta t\, b_+^0 = (\Pi + au)^0 + a\Delta t\, A_+^0 \Pi^0 = \left(I_N + a\Delta t\, A_+^n\right)(\Pi + au)^0,$$

which finally yields to

$$\overrightarrow{w}^{1-} = (\Pi + au)^{1-} = (\Pi + au)^0 = \overrightarrow{w}^0.$$

Similarly one can prove that

$$\overrightarrow{w}^{1-} = (\Pi - au)^{1-} = (\Pi - au)^0 = \overrightarrow{w}^0,$$

so that

$$\begin{cases} u_j^{1-} = \dfrac{\overrightarrow{w}_j^{1-} - \overleftarrow{w}_j^{1-}}{2a} = \dfrac{\overrightarrow{w}_j^0 - \overleftarrow{w}_j^0}{2a} = u_j^0 = 0, \\ \Pi_j^{1-} = \dfrac{\overrightarrow{w}_j^{1-} + \overleftarrow{w}_j^{1-}}{2} = \dfrac{\overrightarrow{w}_j^0 + \overleftarrow{w}_j^0}{2} = \Pi_j^0 = \dfrac{g}{2}(h_j^0)^2, \end{cases}$$

and

$$u_{j+1/2}^{*,1-} = \frac{1}{2}(u_{j+1}^{1-} - u_j^{1-}) - \frac{1}{2a}(\Pi_{j+1}^{1-} - \Pi_j^{1-}) - \frac{g}{2a}\frac{h_{j+1}^0 + h_j^0}{2}(z_{j+1} - z_j) = 0,$$

for all j, and then

$$\tau_j^{1-} = \tau_j^0 + \frac{\Delta t}{\Delta m_j}\left(u_{j+1/2}^{*,1-} - u_{j-1/2}^{*,1-}\right) = \tau_j^0,$$

for all j. Finally we get that the lake is also at rest at the end of the Lagrangian step and, since the transport step is trivial because $u^{1^-} \equiv 0$, the global implicit-explicit scheme is well-balanced. □

Proposition 3. *Under the Whitham subcharacteristic condition and the CFL condition*

$$\max_j \frac{\Delta t}{\Delta x_j}\left((u_{j-\frac{1}{2}}^*)^+ - (u_{j+\frac{1}{2}}^*)^-\right) < 1, \tag{39}$$

the implicit-explicit scheme satisfies the following stability properties:

(i) it is a conservative scheme for the water height h. It is also a conservative scheme for hu when the topography source term vanishes,

(ii) the water height h_j^n is positive for all j and $n > 0$ provided that h_j^0 is positive for all j,

(iii) it is well-balanced,

(iv) it satisfies a discrete entropy inequality,

(iv) and it gives the usual implicit-explicit Lagrange-Projection scheme when the bottom is flat.

Proof. The properties are obtained in the same way as in the explicit case, except for the well-balanced property which has already been proved in the previous Lemma, and the validity of the entropy inequality which is proved in appendix B.

4 Numerical results

The aim of this section is to illustrate the behaviour of our Lagrange-Projection like strategies in one space dimension. We will also compare the results with the simple, well-balanced, positive and entropy-satisfying scheme recently proposed in [2] (the scheme will be referred to as the HLLACU scheme) and the very well-known hydrostatic reconstruction scheme [1] based on a classic HLL scheme and referred to as HRHLL in the following.

Let us first notice that two (classic) options will be considered in order to evaluate the artificial sound speed a involved in the acoustic step. Let $\kappa > 1$. The first one is based on a local definition of the Lagrangian sound speed in agreement with a local evaluation of the subcharacteristic condition, namely

$$a_{j+1/2} = \kappa \max\left(h_j^n\sqrt{gh_j^n}, h_{j+1}^n\sqrt{gh_{j+1}^n}\right), \quad \forall\, j \in \mathbb{Z}, \tag{40}$$

while the second one considers an uniform estimate by setting

$$a_{j+1/2} = \kappa \max_{i\in\mathbb{Z}} \left(h_i^n \sqrt{g h_i^n} \right), \quad \forall\, j \in \mathbb{Z}. \tag{41}$$

In practice, we set $\kappa = 1.01$. For the sake of conciseness, the full-explicit scheme will be referred to as $\mathrm{EXEX_{loc}}$ (resp. $\mathrm{EXEX_{glob}}$) and the semi-implicit scheme will be referred to as $\mathrm{IMEX_{loc}}$ (resp. $\mathrm{IMEX_{glob}}$) when (40) (resp. (41)) is used.

Let us also mention for all the test cases, uniform space steps Δx will be considered and the time steps Δt will be chosen in agreement with the CFL conditions (35) and (39). More precisely, we will set (unless otherwise stated)

$$\Delta t = \frac{\Delta x}{2 \max_{j\in\mathbb{Z}}(\sqrt{g h_j^n}, |u^*_{j+1/2}|)}, \tag{42}$$

for the explicit schemes, and

$$\Delta t = \frac{\Delta x}{2 \max_{j\in\mathbb{Z}}(|u^*_{j+1/2}|)}, \tag{43}$$

for the implicit ones, where $u^*_{j+1/2}$ is calculated at time t^n for the sake of simplicity.

Before starting, let us finally mention that initial data matching the lake at rest condition (5) are preserved by construction by the $\mathrm{EXEX_k}$ and $\mathrm{IMEX_k}$ schemes, $k = \mathrm{loc}, \mathrm{glob}$. Therefore, such test cases will not be considered hereafter.

4.1 Dam break problem

We first consider the classical dam break. The space domain $[0, 1500]$ is divided into two parts with the same length and such that the water height is higher on the left side,

$$h(x, t = 0) = 20, \quad \text{if } x \leq 750, \quad h(x, t = 0) = 15, \quad \text{if } x > 750.$$

The velocity is set to be zero on both side at the initial time when the dam breaks and the water starts flowing. Importantly, the topography is not flat but given by the regularized two-step function

$$z(x) = \begin{cases} 4e^{2-\frac{150}{x-487.5}}, & \text{if } 487.5 < x <= 562.5, \\ 8 - 4e^{2-\frac{150}{637.5-x}}, & \text{if } 562.5 < x <= 637.5, \\ 8, & \text{if } 637.5 < x <= 862.5, \\ 8 - 4e^{2-\frac{150}{x-862.5}}, & \text{if } 862.5 < x <= 937.5, \\ 4e^{2-\frac{150}{1012.5-x}}, & \text{if } 937.5 < x <= 1012.5, \\ 0 & \text{otherwise.} \end{cases}$$

At last, the spatial domain is discretized over a 1500-cell grid and Neumann boundary conditions are used. Figures 1 and 1 show the solutions at final times $T = 10$ and $T = 50$ with different numerical strategies. The following comments are in order. We first observe that the implicit schemes are the most diffusive, which was clearly expected from the implicit treatment of the acoustic step. Note also that our Lagrangian-Projection schemes are intrinsically made of two averaging steps, which is necessary to separate the acoustic and transport effects, but at the price of additional numerical diffusion compared to a direct Eulerian approach like the one proposed in the HLLACU scheme. We also observe that a local definition of parameter a is preferable to the global one in order to reduce the numerical diffusion. For this reason, we will only consider the local evaluation in the following test cases ($k = \mathrm{loc}$). As far as the time step step is concerned, we observed for this test case that the averaged value (calculated from the time iterations needed to reach the final time $T = 50$) is about five times larger for the $\mathrm{IMEX_{loc}}$ than for the $\mathrm{EXEX_{loc}}$ schemes.

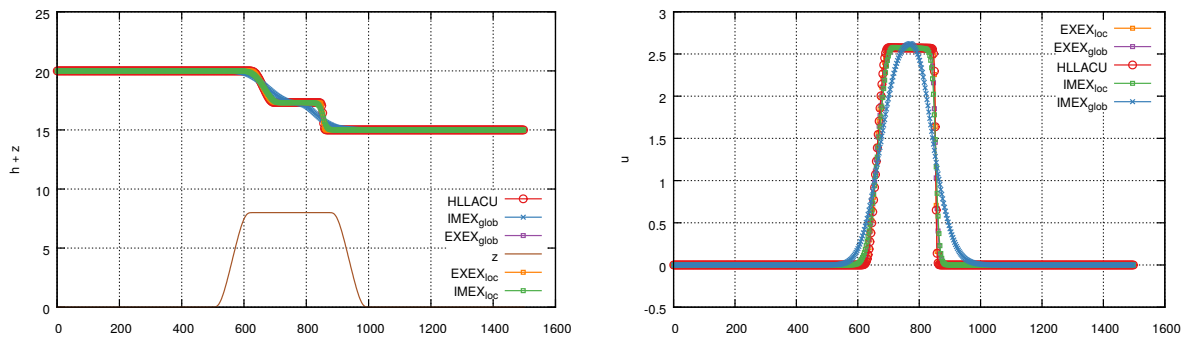

Figure 1: Dam Break problem. Profile of $z + h$ and u at time $T = 10$.

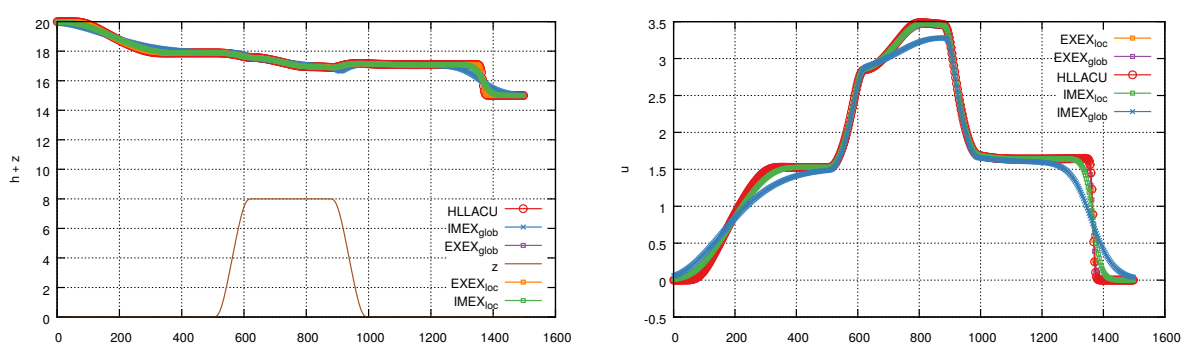

Figure 2: Dam Break problem. Profile of $z + h$ and u at time $T = 50$.

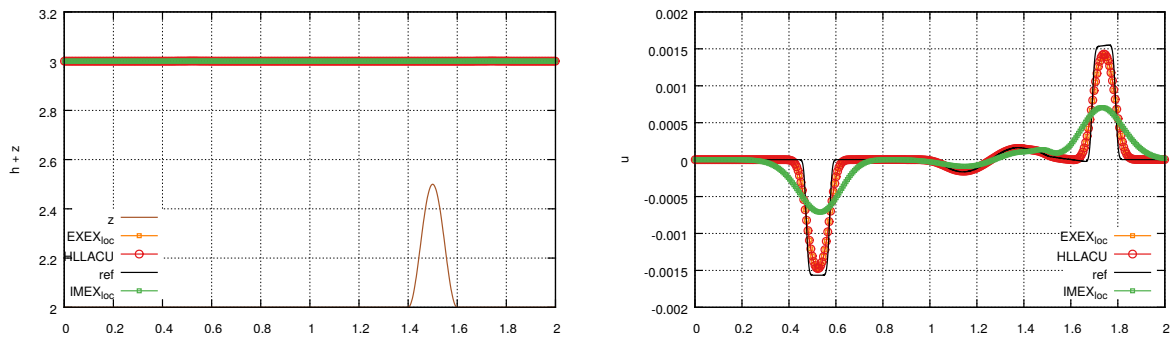

Figure 3: Propagation of perturbations test at final time $T = 0.2$. On the left : total heights $h + z$, on the right : velocities u, with $\Delta x = 1/500$.

4.2 Propagation of perturbations

This test case focuses on the perturbation of a steady state solution by a pulse that splits into two opposite waves. More precisely, the space domain is reduced to the interval $[0, 2]$, the bottom topography is defined by $z(x) = 2 + 0.25(\cos(10\pi(x - 0.5)) + 1)$ if $1.4 < x < 1.6$, and 2 otherwise, and the initial state is such that $u(0, x) = 0$ and $h(0, x) = 3 - z(x) + \Delta h$ if $1.1 < x < 1.2$, and $3 - z(x)$ otherwise, where $\Delta h = 0.001$ is the height of the perturbation. The CFL parameter is set to 0.9 (instead of 1/2 in (42) and (43)), the final time is $T = 0.2$, the space step equals $\Delta x = 1/500$ and Neumann boundary conditions are used.

It turns out that since the perturbation is small, the values of the velocity u keeps a small amplitude during the whole computation. As an immediate consequence, considering the natural CFL condition (43) gives very large time steps which naturally induces much numerical diffusion. In order to reduce the numerical diffusion and improve the overall accuracy of the numerical solution, the time step Δt_{imp} given by (43) was first limited to ten times the time step Δt_{exp} given by (42). In other words, we chose

$$\Delta t = \min(10\Delta t_{exp}, \Delta t_{imp})$$

for this test case. Figure 3 compares the numerical solutions given by the EXEX$_{\text{loc}}$, IMEX$_{\text{loc}}$ and HLLACU schemes. The implicit scheme is clearly more diffusive than the explicit ones. Note that the so-called reference solution is given by the solution of the HLLACU scheme on a 10000-cell grid.

Figure 4 shows that same solutions but the implicit scheme is now run using the explicit CFL restriction (42). As expected, the numerical approximation is more accurate and the numerical diffusion is significantly reduced. At last, Figure 5 shows the numerical solutions using a 10000-cell grid. The schemes converge to the same solution.

4.3 Steady flow over a bump.

The aim of this test case is to test the ability of the schemes to converge to some moving water equilibrium. Let us remind that the steady states are governed by the equations $hu = K_1$ and $\dfrac{u^2}{2} + g(h + z) = K_2$.

Fluvial regime : In this test case, we set $K_1 = 1$ and $K_2 = 25$, we denote $h_{eq}(x), u_{eq}(x)$ the values of h and u at this equilibrium. The domain is $[0, 4]$ and the bottom topography is defined by $z(x) = (\cos(10\pi(x - 1)) + 1)/4$ if $1.9 \leq x \leq 2.1$ and 0 elsewhere. The CFL parameter is equal to 0.5 and the space step to $\Delta x = 1/400$. The initial condition is chosen out of equilibrium and given by

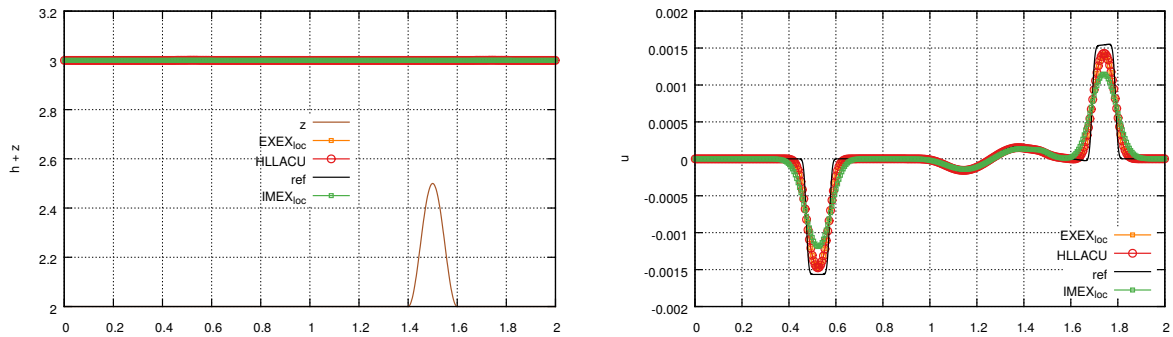

Figure 4: Propagation of perturbations test at final time $T = 0.2$. On the left : total heights $h + z$, on the right : velocities u, with $\Delta x = 1/500$. Here, the implicit scheme is run using the explicit CFL restriction (42).

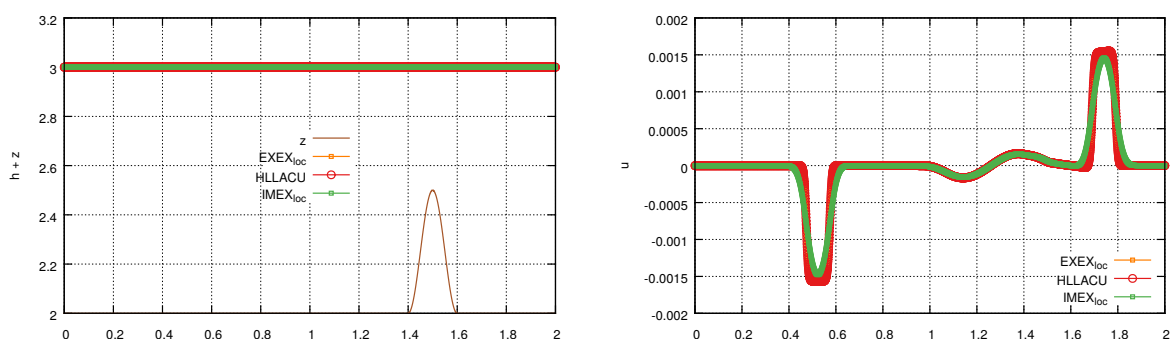

Figure 5: Propagation of perturbations test at final time $T = 0.2$. On the left : total heights $h + z$, on the right : velocities u, with $\Delta x = 1/5000$.

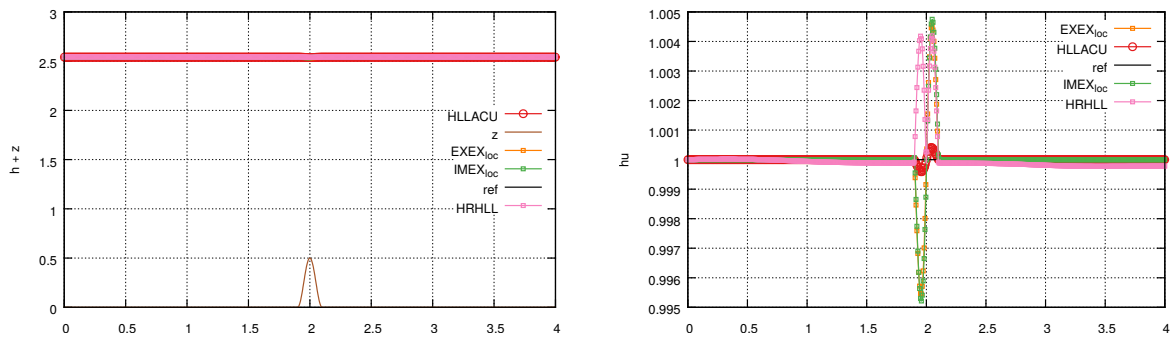

Figure 6: Fluvial regime at time $T = 200$. On the left : total heights $h + z$, on the right : discharge hu.

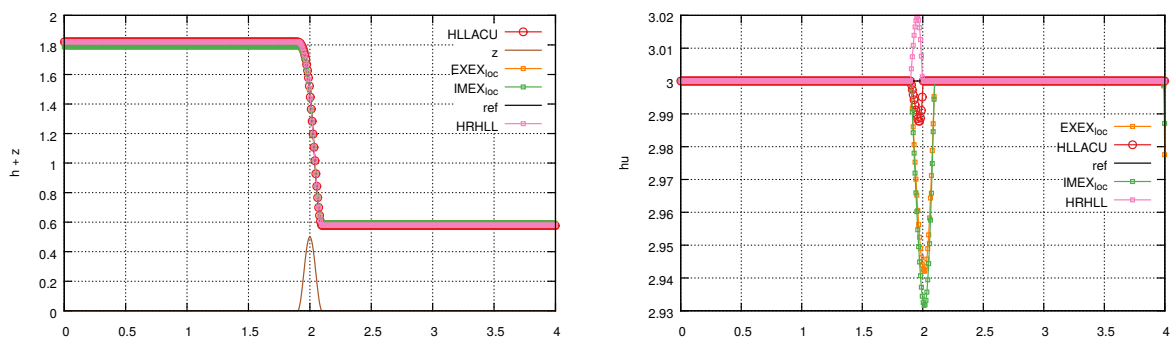

Figure 7: Transcritical regime without shock. On the left : total heights $h+z$, on the right : discharge hu.

$h = h_{eq}$ and $u = 0$. The boundary conditions are set to be

$$\begin{cases} \partial_x h(x=0) = 0, \\ (hu)(x=0) = K_1, \end{cases} \quad \text{and} \quad \begin{cases} h(x=4) = h_{eq}(x=4), \\ \partial_x (hu)(x=4) = 0. \end{cases}$$

Figure 6 shows the solution at the final time $t = 200$. We can observe that the solutions are close to the expected equilibrium, except near the mid domain where the momentum is not yet constant for the mesh size under consideration. The Lagrange-Projection schemes give numerical solutions very close to the one obtained with the HRHLL scheme based on the hydrostatic reconstruction, while the ACU scheme is clearly more accurate. Note also that on this test case, the implicit CFL condition (43) allows to use time steps up to ten times larger than the explicit condition (42).

Transcritical regime without shock : In this test case, we set $K_1 = 3$, $K_2 = \frac{3}{2}(K_1 g)^{2/3} + \frac{g}{2}$. We used the same boundary conditions and started from the same initial condition as in the previous simulation. The solutions are shown at time $t = 10$ on Figure 7.

Transcritical regime with shock : This test has been proposed by Castro et al. [11]. The parameters are described hereafter: the space domain is the interval $[0, 25]$, the bottom topography is defined by $z(x) = 3 - 0.005(x - 10)^2$, if $8 < x < 12$, and 2.8 otherwise. The initial state is defined by $h(0, x) = 3.13 - z(x)$, $q(0, x) = 0.18$ and the boundary conditions are $q(t, 0) = 0.18$, $\partial_x q(t, 25) = 0$, $h(t, 25) = 0.33$ and $\partial_x h(t, 0) = 0$. The final time is set to $t = 200$, the space step to $\Delta x = 1/64$ and the CFL to 0.9. We can see on the Figure 8 that we obtain similar results with the different schemes.

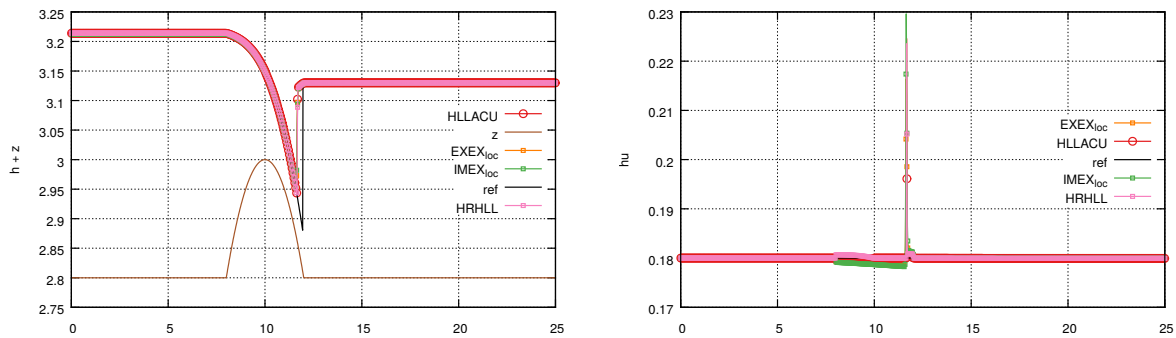

Figure 8: Transcritical regime with shock at final time $T = 200$. On the left : total heights $h + z$, on the right : discharge hu.

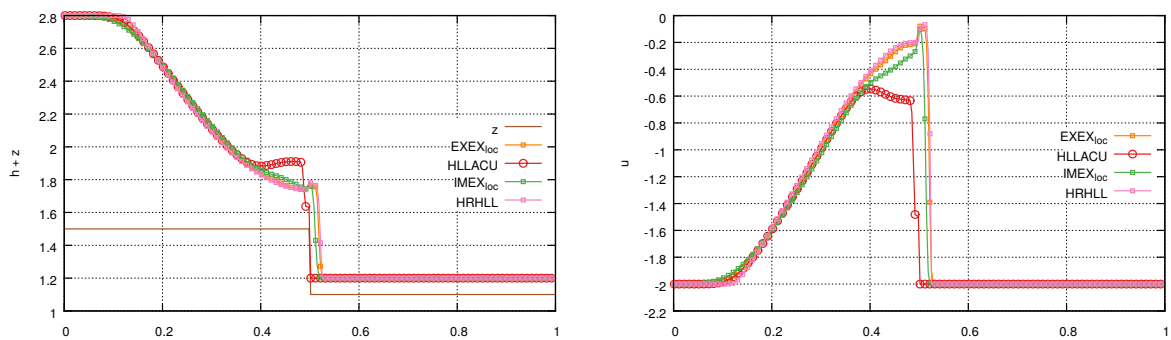

Figure 9: Non-unique solution test case at final time $T = 0.1$. On the left : total heights $h + z$, on the right : velocities u.

4.4 Non-unique solution to the Riemann problem

This aim of this test case is to consider a Riemann problem for which the entropy solution is not unique, in order to see whether the numerical schemes capture the same solution or not. The spatial domain is $[0, 1]$, the gravitational acceleration g is set to 2 and the CFL coefficient equals 0.9. Note however that considering the mixed implicit-explicit scheme, the time step Δt was restricted to three times the explicit time step, namely

$$\Delta t = \min(3\Delta t_{exp}, \Delta t_{imp})$$

where we have used the same notations as in the propagation of perturbations test case. The final time $T = 0.1$ and the space step is $\Delta x = 1/300$. The initial data is given by

$$(z, h, u)^T = \begin{cases} (1.5, 1.3, -2)^T & \text{if } x \leq 0.5, \\ (1.1, 0.1, -2)^T & \text{if } x > 0.5, \end{cases}$$

and we used Neumann boundary conditions. It is quite interesting to observe on Figure 9 that the methods proposed in the present paper and the hydrostatic scheme seem to converge to the same solution, while the HLLACU scheme capture a quite different solution.

Conclusion

We have proposed a large time step and well-balanced scheme for the shallow-water equations and proved stability properties under a time step CFL restriction based on the material velocity u and not on the sound speed c as it is customary. The Lagrangian-Projection decomposition proved to be efficient on a variety of test cases, but may be more diffusive than a direct Eulerian approach. We believe that the proposed implicit-explicit strategy is especially well adapted for subsonic flows but even more for large Froude numbers, which is our very motivation and the purpose of an ongoing work in several space dimensions. Works in progress also include a high-order accuracy extension using discontinuous Galerkin strategies for the space variable and Runge-Kunta techniques for the time variable.

Acknowledgement. This work was partially supported by a public grant as part of the Investissement d'avenir project, reference ANR-11-LABX-0056-LMH, LabEx LMH. The authors also thank S. Noelle and H. Zakerzadeh for useful and interesting discussions on the topic.

References

[1] E. Audusse, F. Bouchut, M. O. Bristeau, R. Klein, B. Perthame, A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows, SIAM J. Sci. Comp., 25, 2050–2065 (2004).

[2] E. Audusse, C. Chalons, P. Ung, A very simple well-balanced positive and entropy-satisfying scheme for the shallow-water equations, to appear in Commun. in Math. Sci. (2015).

[3] A. Bermudez, M.E. Vazquez-Cendon, Upwind Methods for Hyperbolic Conservation Laws with Source Terms, Comp. & Fluids. **23** 1049–1071 (1994).

[4] C. Berthon, C. Chalons, A fully well-balanced, positive and entropy-satisfying Godunov-type method for the Shallow-Water Equations, Math. of Comp., vol 85, pp 1281-1307, 2016.

[5] C. Berthon, C. Chalons, S. Cornet, and G. Sperone. Fully well-balanced, positive and simple approximate Riemann solver for shallow-water equations. *Bull. Braz. Math. Soc.*, New Series 47(1), 1-14, 2016.

[6] C. Berthon, F. Foucher, Efficient wellbalanced hydrostatic upwind schemes for shallowwater equations, J. Comput. Phys., 231 4993–5015 (2012).

[7] G. Bispen, K. R. Arun, M. Lukacova-Medvidova, S. Noelle, IMEX large time step finite volume methods for low Froude number shallow water flows. Communications in Computational Physics 16 (2014), 307-347.

[8] F. Bouchut. A reduced stability condition for nonlinear relaxation to conservative laws. J. Hyp. Diff. Eq., 1(1), pp. 149-170, (2004).

[9] F. Bouchut, Non-linear stability of finite volume methods for hyperbolic conservation laws and well-balanced schemes for sources, Frontiers in Mathematics, Birkhauser (2004).

[10] F. Bouchut, T. Morales, A subsonic-well-balanced reconstruction scheme for shallow water flows, Siam J. Numer. Anal., 48(5):1733–1758 (2010).

[11] M. J. Castro, A. Pardo, C. Parés, Well-balanced numerical schemes based on a generalized hydrostatic reconstruction technique, Mathematical Models and Methods in Applied Sciences, 17, 2065–2113 (2007).

[12] F. Chalons, C. Coquel. Navier-stokes equations with several independant pressure laws and explicit predictor-corrector schemes. Numerisch Math, 101(3), pp. 451-478, (2005).

[13] C. Chalons, F. Coquel, E. Godlewski, P.-A. Raviart, N. Seguin, *Godunov-type schemes for hyperbolic systems with parameter dependent source. The case of Euler system with friction*, Mathematical Models and Methods in Applied Sciences (M3AS), 20(11), 2010.

[14] C. Chalons and J.F. Coulombel. Relaxation approximation of the Euler equations. Journal of Mathematical Analysis and Applications, 348(2), pp. 872-893, (2008).

[15] C. Chalons, S. Kokh, and N. Spillane. Large time-step numerical scheme for the seven-equation model of compressible two-phase flows. In Springer, editor, *Proceedings in Mathematics, FVCA 6*, volume 4, pages pp. 225–233, 2011.

[16] C. Chalons, M. Girardin, and S. Kokh. Large time step and asymptotic preserving numerical schemes for the gas dynamics equations with source terms. *SIAM J. Sci. Comput.*, 35(6):pp. a2874–a2902, 2013.

[17] C. Chalons, M. Girardin, and S. Kokh. Operator-splitting based AP schemes for the 1D and 2D gas dynamics equations with stiff sources. AIMS Series on Applied Mathematics, vol 8, 607–614, 2014.

[18] C. Chalons, M. Girardin, and S. Kokh. An all-regime Lagrange-Projection like scheme for the gas dynamics equations on unstructured meshes. *to appear in Communications in Computational Physics, 2016.*

[19] Chalons C., Girardin M., S. Kokh. An all-regime Lagrange-Projection like scheme for 2D homogeneous models for two-phase flows on unstructured meshes, submitted.

[20] G.-Q. Chen, C.D. Levermore, T.-P. Liu, Hyperbolic conservation laws with stiff relaxation terms and entropy, Comm. Pure Appl. Math., volume 47(6), 787–830, 1994.

[21] A. Chinnayya, A.-Y. LeRoux, N. Seguin, A well-balanced numerical scheme for the approximation of the shallow-water equations with topography: the resonance phenomenon, International Journal on Finite Volume (electronic), volume 1(1), 1–33, 2004.

[22] F. Coquel, Q.L. Nguyen, M. Postel, and Q.H. Tran. Entropy-satisfying relaxation method with large time-steps for Euler IBVPs. Math. Comp. 79, pp. 1493-1533, (2010).

[23] F. Coquel, E. Godlewski, B. Perthame, A. In and P. Rascle. Some new Godunov and relaxation methods for two-phase flow problems, In Godunov methods (Oxford, 1999) (Kluwer/Plenum, 2001) pp. 179-188.

[24] G. Gallice. Solveurs simples positifs et entropiques pour les systèmes hyperboliques avec terme source. C. R. Math. Acad. Sci. Paris 334, no. 8, 713–716 (2002).

[25] G. Gallice. Positive and entropy stable Godunov-type schemes for gas dynamics and MHD equations in Lagrangian or Eulerian coordinates. Numer. Math. 94 no.4, 673–713 (2003).

[26] E. Godlewski and P.-A. Raviart. Numerical approximation of hyperbolic systems of conservation laws. Applied Mathematical Sciences, vol. 118, Springer-Verlag, New York, (1996).

[27] L. Gosse, A well-balanced flux-vector splitting scheme designed for hyperbolic systems of conservation laws with source terms, Comput. Math. Appl. 39, 135–159 (2000).

[28] L. Gosse, Computing qualitatively correct approximations of balance laws. Exponential-fit, well-balanced and asymptotic-preserving, SEMA SIMAI Springer Series 2 (2013).

[29] J. M. Greenberg, A. Y. Leroux, A well-balanced scheme for the numerical processing of source terms in hyperbolic equations, SIAM J. Numer. Anal., 33, 1–16 (1996).

[30] S. Jin, A steady-state capturing method for hyperbolic systems with geometrical source terms, Math. Model. Numer. Anal., 35, 631–645 (2001).

[31] S. Jin and Z. P. Xin. The relaxation schemes for systems of conservation laws in arbitrary space dimension. Comm. Pure Appl. Math. 48, no. 3, pp. 235-276, (1995).

[32] B. Perthame, C. Simeoni, A kinetic scheme for the Saint-Venant system with source term, Calcolo, 38, Springer-Verlag, 201-231, (2001).

[33] I. Suliciu. Ont the thermodynamics of fluids with relaxation and phase transitions. Fluids with relaxation. Int. J. Engag. Sci. 36, pp. 921-947, (1998).

[34] Y. Xing, Exactly well-balanced discontinuous Galerkin methods for the shallow water equations with moving water equilibrium, J. Comput. Phys. 257 536-553 (2014).

[35] Y. Xing, C.-W. Shu, High order well-balanced finite volume WENO schemes and discontinuous Galerkin methods for a class of hyperbolic systems with source terms, J. Comput. Phys. 214 567-598 (2006).

[36] Y. Xing, C.-W. Shu, S. Noelle, On the advantage of well-balanced schemes for moving-water equilibria of the shallow water equations, J. Sci. Comput. 48 339-349 (2011).

[37] Y. Xing, X. Zhang, C.-W. Shu, Positivity-preserving high order well-balanced discontinuous Galerkin methods for the shallow water equations, Adv. Water Resour. 33 1476-1493 (2010).

[38] H. Zakerzadeh, On the Mach-uniformity of the Lagrange-Projection scheme, submitted (2016).

A Eigenstructure of the relaxed acoustic system

Considering smooth solutions, the homogeneous relaxed acoustic system (14) reads

$$\partial_t \mathbf{W} + A(W)\partial_m \mathbf{W} = 0, \quad A = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & g/\tau \\ 0 & a^2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{44}$$

The matrix the eigenvalues of A are $\{-a, 0, +a\}$. A basis of right eigenvectors of A is

$$\mathbf{r}_0^{(1)} = (1,0,0,0)^T, \qquad \mathbf{r}_0^{(2)} = (0,0,-g,\tau)^T, \qquad \mathbf{r}_\pm = (1, \mp a, -a^2, 0)^T$$

where $\mathbf{r}_0^{(1)}$ and $\mathbf{r}_0^{(1)}$ are associated with the double eigenvalue 0 and $\mathbf{r}_\pm$ is associated with $\pm a$. The system (44) is thus hyperbolic. All the characteristic fields of (44) are linearly degenerate.

The $(\pm a)$-field possesses three Riemann invariants

$$I_\pm^1 = \pi \mp au, \qquad I_\pm^2 = u \pm a\tau, \qquad I_\pm^3 = z.$$

As a consequence, the states $\mathbf{W}_L$ and $\mathbf{W}_R$ that can be connected by a $(\pm a)$-wave can be obtained thanks to the continuity of the $(\pm a)$-Riemann invariants, which amounts to verify the jump relations

$$\mp a(\mathbf{W}_R - \mathbf{W}_R) + \mathbf{G}(\mathbf{W}_R) - \mathbf{G}(\mathbf{W}_R) = 0. \tag{45}$$

Unfortunately, the eigenvalue 0 is of multiplicity 2 and the 0-field only has a single Riemann invariant

$$I_0 = u.$$

Therefore we can only state that if two states $\mathbf{W}_L$ and $\mathbf{W}_R$ are connected by a 0-wave then

$$u_R = u_L. \tag{46}$$

B Proof of the discrete entropy inequality of Proposition 3

The proof of the discrete entropy inequality follows exactly the same lines as the one proposed in [22] for the barotropic gas dynamics equations, but taking into account here the presence of the topography source term. It naturally leads to a non conservative version of the entropy inequality. A discrete and conservative entropy inequality for the proposed algorithm remains an open problem so far. Our result states as follows.

Lemma 2. *We have the following discrete form of the entropy inequality* (3) *for all* $j \in \mathbb{Z}$*, namely*

$$\mathcal{U}_j^{n+1} - \mathcal{U}_j^n + \frac{\Delta t}{\Delta x_j}\left(\mathcal{F}_{j+1/2}^{n+1-} - \mathcal{F}_{j-1/2}^{n+1-}\right) \leq -\Delta t\, g\left\{hu\partial_x z\right\}_j,$$

with the entropy numerical fluxes

$$\mathcal{F}_{j+1/2}^{n+1-} = \left(\Pi_{j+1/2}^* + \mathcal{U}_{j+1/2}^{n+1-}\right)\tilde{u}_{j+1/2}^*,$$

where

$$\mathcal{U}_{j+1/2}^{n+1-} = \begin{cases} \mathcal{U}_j^{n+1-} & \text{if } u_{j+1/2}^* \geq 0, \\ \mathcal{U}_j^{n+1-} & \text{if } u_{j+1/2}^* < 0, \end{cases}$$

and

$$\begin{cases} \tilde{u}_{j+1/2}^* = \dfrac{u_j^{n+1-} + u_{j+1}^{n+1-}}{2} - \dfrac{1}{2a}\left(\Pi_{j+1}^{n+1-} - \Pi_j^{n+1-}\right), \\ \Pi_{j+1/2}^* = \dfrac{\Pi_j^{n+1-} + \Pi_{j+1}^{n+1-}}{2} - \dfrac{a}{2}\left(u_{j+1}^{n+1-} - u_j^{n+1-}\right), \end{cases}$$

are consistent with $\mathcal{F}$*, and the non conservative source term*

$$\left\{hu\partial_x z\right\}_j = \frac{1}{2a\Delta x_j}\left[\frac{h_j^n + h_{j-1}^n}{2}\overrightarrow{w}_j^{n+1-}\left(z_j - z_{j-1}\right) - \frac{h_{j+1}^n + h_j^n}{2}\overleftarrow{w}_j^{n+1-}\left(z_{j+1} - z_j\right)\right]$$

is consistent with $hu\partial_x z$*.*

Proof. Let us first observe that smooth solutions of (14) satisfy

$$\partial_t(\Pi^2 + a^2u^2) + 2a^2\partial_m\Pi u = -\frac{2a^2gu}{\tau}\partial_m z. \tag{47}$$

In order to obtain a discrete version of this equality, let us define

$$\eta_j^{n+1-} := \frac{(\overleftarrow{w}_j^{n+1-})^2 + (\overrightarrow{w}_j^{n+1-})^2}{2} = (\Pi_j^{n+1-})^2 + a^2(u_j^{n+1-})^2,$$

and

$$q_{j+1/2}^{n+1-} := \frac{(\overrightarrow{w}_j^{n+1-})^2 - (\overleftarrow{w}_{j+1}^{n+1-})^2}{4a} = \Pi_{j+1/2}^*\,\tilde{u}_{j+1/2}^*.$$

The formulas (38) also read

$$\begin{cases} \tau_j^{n+1^-} - \tau_j^n = \dfrac{\Delta t}{\Delta m_j}\left[u_{j+1/2}^* - u_{j-1/2}^*\right], \\ \overleftarrow{w}_j^{n+1-} - \overleftarrow{w}_j^n = a\dfrac{\Delta t}{\Delta m_j}\left[\overleftarrow{w}_{j+1}^{n+1-} - \overleftarrow{w}_j^{n+1-} + g\dfrac{h_{j+1}^n + h_j^n}{2}\left(z_{j+1} - z_j\right)\right], \\ \overrightarrow{w}_j^{n+1-} - \overrightarrow{w}_j^n = -a\dfrac{\Delta t}{\Delta m_j}\left[\overrightarrow{w}_j^{n+1-} - \overrightarrow{w}_{j-1}^{n+1-} + g\dfrac{h_j^n + h_{j-1}^n}{2}\left(z_j - z_{j-1}\right)\right], \end{cases}$$

while adding the third equation of (36) and a^2 times the first equation of (36) also gives $I_j^{n+1-} = I_j^n$ where $I = \Pi + a^2\tau$. Multiplying the second and the third equations above by $\overleftarrow{w}_j^{n+1-}$ and $\overrightarrow{w}_j^{n+1-}$ then gives

$$\begin{cases} I_j^{n+1-} = I_j^n, \\ \overleftarrow{w}_j^{n+1-}(\overleftarrow{w}_j^{n+1-} - \overleftarrow{w}_j^n) = a\dfrac{\Delta t}{\Delta m_j}\left[\overleftarrow{w}_j^{n+1-}\left(\overleftarrow{w}_{j+1}^{n+1-} - \overleftarrow{w}_j^{n+1-}\right) + \overleftarrow{w}_j^{n+1-} g\dfrac{h_{j+1}^n + h_j^n}{2}\left(z_{j+1} - z_j\right)\right], \\ \overrightarrow{w}_j^{n+1-}(\overrightarrow{w}_j^{n+1-} - \overrightarrow{w}_j^n) = -a\dfrac{\Delta t}{\Delta m_j}\left[\overrightarrow{w}_j^{n+1-}\left(\overrightarrow{w}_j^{n+1-} - \overrightarrow{w}_{j-1}^{n+1-}\right) + \overrightarrow{w}_j^{n+1-} g\dfrac{h_j^n + h_{j-1}^n}{2}\left(z_j - z_{j-1}\right)\right], \end{cases}$$

that is to say, since

$$2b(b-a) = (b^2 - a^2) + (b-a)^2 \quad \text{and} \quad 2b(a-b) = (a^2 - b^2) - (b-a)^2,$$

$$\begin{cases} I_j^{n+1-} = I_j^n, \\ \left((\overleftarrow{w}_j^{n+1-})^2 - (\overleftarrow{w}_j^n)^2\right) - a\dfrac{\Delta t}{\Delta m_j}\left((\overleftarrow{w}_{j+1}^{n+1-})^2 - (\overleftarrow{w}_j^{n+1-})^2\right) = \\ \qquad -(\overleftarrow{w}_j^{n+1-} - \overleftarrow{w}_j^n)^2 + a\frac{\Delta t}{\Delta m_j}\left[-\left(\overleftarrow{w}_{j+1}^{n+1-} - \overleftarrow{w}_j^{n+1-}\right)^2 + 2\overleftarrow{w}_j^{n+1-} g\frac{h_{j+1}^n + h_j^n}{2}\left(z_{j+1} - z_j\right)\right], \\ \left((\overrightarrow{w}_j^{n+1-})^2 - (\overrightarrow{w}_j^n)^2\right) + a\dfrac{\Delta t}{\Delta m_j}\left((\overrightarrow{w}_j^{n+1-})^2 - (\overrightarrow{w}_{j-1}^{n+1-})^2\right) = \\ \qquad -(\overrightarrow{w}_j^{n+1-} - \overrightarrow{w}_j^n)^2 - a\frac{\Delta t}{\Delta m_j}\left[\left(\overrightarrow{w}_j^{n+1-} - \overrightarrow{w}_{j-1}^{n+1-}\right)^2 + 2\overrightarrow{w}_j^{n+1-} g\frac{h_j^n + h_{j-1}^n}{2}\left(z_j - z_{j-1}\right)\right]. \end{cases}$$

Summing the last two equations, we immediately get the following discrete version of (47), namely

$$\eta_j^{n+1-} - \eta_j^n + 2a^2\frac{\Delta t}{\Delta m_j}(q_{j+1/2}^{n+1-} - q_{j-1/2}^{n+1-}) \leq -\Delta t\, 2a^2\, g\, \tau_j^n \left\{hu\partial_x z\right\}_j.$$

The rest of the proof strictly follows the one proposed in [22]. It is given here for the sake of completeness. With this in mind, let us define the energy E such that $hE = \mathcal{U}$, which means

$$E = \frac{u^2}{2} + e(\tau) = \frac{u^2}{2} + e(\tau) + \frac{\Pi^2 - \Pi^2}{2a^2} = e(\tau) + \frac{\eta - \Pi^2}{2a^2},$$

where we have set $e(\tau) = \frac{g}{2\tau} = \frac{gh}{2}$. We clearly have

$$E_j^{n+1-} - E_j^n = e(\tau_j^{n+1-}) - e(\tau_j^n) + \frac{\eta_j^{n+1-} - \eta_j^n}{2a^2} - \frac{(\Pi_j^{n+1-})^2 - (\Pi_j^n)^2}{2a^2}$$

so that, since $a^2 - b^2 = (a-b)^2 + 2b(b-a)$, we have

$$E_j^{n+1-} - E_j^n = e(\tau_j^{n+1-}) - e(\tau_j^n) + \frac{\eta_j^{n+1-} - \eta_j^n}{2a^2} - \frac{(\Pi_j^{n+1-} - \Pi_j^n)^2}{2a^2} - \frac{\Pi_j^n(\Pi_j^{n+1-} - \Pi_j^n)}{a^2}.$$

But $I_j^{n+1-} = I_j^n$ gives $\Pi_j^{n+1-} - \Pi_j^n = -a^2(\tau_j^{n+1-} - \tau_j^n)$ so that

$$E_j^{n+1-} - E_j^n + \frac{\Delta t}{\Delta m_j}(q_{j+1/2}^{n+1-} - q_{j-1/2}^{n+1-}) \leq e(\tau_j^{n+1-}) - e(\tau_j^n) + \Pi_j^n(\tau_j^{n+1-} - \tau_j^n) - \frac{a^2}{2}(\tau_j^{n+1-} - \tau_j^n)^2 - \Delta t\, g\, \tau_j^n \left\{hu\partial_x z\right\}_j.$$

Since the solution at time t^n is at equilibrium, we have $\Pi_j^n = p(\tau_j^n) = -e'(\tau_j^n) = \dfrac{g}{2}(h_j^n)^2$, so that a Taylor expansion gives

$$E_j^{n+1-} - E_j^n + \frac{\Delta t}{\Delta m_j}(q_{j+1/2}^{n+1-} - q_{j-1/2}^{n+1-}) \leq \frac{(e''(\xi) - a^2)}{2}(\tau_j^{n+1-} - \tau_j^n)^2 - \Delta t\, g\, \tau_j^n \left\{hu\partial_x z\right\}_j$$

and

$$E_j^{n+1-} - E_j^n + \frac{\Delta t}{\Delta m_j}(q_{j+1/2}^{n+1-} - q_{j-1/2}^{n+1-}) \leq \frac{(-p^{'}(\xi) - a^2)}{2}(\tau_j^{n+1-} - \tau_j^n)^2 - \Delta t\, g\, \tau_j^n \left\{hu\partial_x z\right\}_j \leq -\Delta t\, g\, \tau_j^n \left\{hu\partial_x z\right\}_j$$

by the Whitham subcharacteristic condition. This inequality is nothing but the expected entropy inequality but in Lagrangian coordinates. At this stage, it is very usual to combine the definition of the remap step (which, setting $X = h, hu$, gives X_j^{n+1} as a convex combination of X_{j-1}^{n+1-}, X_j^{n+1-} and X_{j+1}^{n+1-} under the transport CFL condition) together with the Jensen inequality for the convex mapping $(h, hu) \mapsto \mathcal{U}(h, hu)$, in order to get the expected entropy inequality in Eulerian coordinates, namely

$$\mathcal{U}_j^{n+1} - \mathcal{U}_j^n + \frac{\Delta t}{\Delta x_j}\left(\mathcal{F}_{j+1/2}^{n+1-} - \mathcal{F}_{j-1/2}^{n+1-}\right) \leq -\Delta t\, g \left\{hu\partial_x z\right\}_j .$$

We refer the reader to [22] for more details.

□