A 2d Well-balanced Positivity Preserving Second Order Scheme for Shallow Water Flows on Unstructured Meshes

Emmanuel Audusse, Marie-Odile Bristeau

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

A 2d Well-balanced Positivity Preserving Second Order Scheme for Shallow Water Flows on Unstructured Meshes

Emmanuel Audusse — Marie-Odile Bristeau

N° 5260

Juillet 2004

Thème NUM

ISSN 0249-6399 ISRN INRIA/RR--5260--FR+ENG

INRIA
ROCQUENCOURT

A 2d Well-balanced Positivity Preserving Second Order Scheme for Shallow Water Flows on Unstructured Meshes

Emmanuel Audusse * , Marie-Odile Bristeau *

Thème NUM — Systèmes numériques
Projet BANG

Rapport de recherche n° 5260 — Juillet 2004 — 31 pages

Abstract: We consider the solution of the Saint-Venant equations with topographic source terms on 2D unstructured meshes by a finite volume approach. We first present a stable and positivity preserving homogeneous solver issued from a kinetic representation of this hyperbolic conservation laws system. This water depth positivity property is important when dealing with wet-dry interfaces. Then we introduce a local hydrostatic reconstruction that preserves the positivity properties of the homogeneous solver and leads to a well-balanced scheme satisfying the steady state condition of still water. Finally a second order extension based on limited reconstructed values on both sides of each interface and on an enriched interpretation of the source terms satisfies the same properties and gives a noticeable accuracy improvement. Numerical examples on academic and real problems are presented.

Key-words: Saint-Venant system, Shallow water flow, Hyperbolic systems, Finite volumes, Kinetic solver, Hydrostatic reconstruction, Well-balanced scheme, Positivity preserving scheme, Second order extension

* Project BANG

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau, Rocquencourt, BP 105, 78153 Le Chesnay Cedex (France)
Téléphone : +33 1 39 63 55 11 — Télécopie : +33 1 39 63 53 30

Un schéma 2d de second ordre préservant les équilibres et la positivité pour le système de Saint-Venant

Résumé : On s'intéresse ici à la résolution, par une méthode volumes finis, du système de Saint-Venant avec termes sources topographiques sur des domaines 2d en maillages non strucrturés. Grace à une interprétation cinétique de ce système hyperbolique de lois de conservation, nous dérivons tout d'abord un solveur homogène stable et qui préserve la positivité de la hauteur d'eau - propriété indispensable lorsqu'on s'intéresse à des problèmes de bancs (dé)couvrants. Nous introduisons ensuite une reconstruction hydrostatique locale, qui préserve les propriétés de positivité du solveur homogène, et qui permet de construire un schéma adapté au système de Saint-Venant avec termes sources topographiques, puisqu'il préserve l'équilibre du lac au repos. Enfin, une extension au second ordre, basée sur un processus reconstruction-limitation des variables de chaque côté de l'interface et sur une modélisation plus fine des termes sources, permet des gains de précision remarquables, tout en préservant les propriétés de positivité et d'équilibre. Des applications numériques sur des cas-tests académiques et sur des problèmes réels sont présentées.

Mots-clés : Système de Saint-Venant, Eaux peu profondes, Systèmes hyperboliques, Volumes finis, Solveur cinétique, Reconstruction hydrostatique, Schéma équilibre, Schéma positif, Ordre deux

1 INTRODUCTION

We consider in this article the 2D Saint-Venant system with topographic source term. This system introduced in [41] is very commonly used for the numerical simulation of various geophysical shallow-water flows, such as rivers, lakes or coastal areas, or even oceans, atmosphere or avalanches [8, 23] when completed with appropriate terms. It can be derived as a formal first order approximation of the three-dimensional free surface incompressible Navier-Stokes equations, using the so-called shallow water assumption [15, 20]. Usually, several other terms are added in order to take into account frictions on the bottom and the surface and other physical features. One can also describe the evolution of a temperature (or a concentration of a pollutant) advected by the flow by adding a third equation to the system [4].

The difficulty to define accurate numerical schemes is related to the deep mathematical structure of such hyperbolic systems; the first existence proof of weak solutions after shocks in the large is due to [33] in 1995. It is based on the kinetic interpretation of the system which is also a method to derive numerical schemes with good properties.

Even if some authors use a finite element framework [25, 29], a classical approach for solving hyperbolic systems of conservation laws consists in using finite volume schemes (see [21, 31, 7]). In the context of the discretization of the Saint-Venant system, it is important to get schemes that satisfy very natural properties such as positivity of the water depth, ability to compute dry areas, preservation of steady states such as still water equilibrium (see [3] and the references therein), and eventually satisfy a discrete entropy inequality [38]. For hyperbolic systems with source terms, the difficulty to build stable schemes which preserve equilibria was pointed out by several authors, see Bermudez and Vasquez [6], Greenberg and Leroux [19], LeVeque [32] and led to the notion of *well-balanced* schemes.

The difficulty to allow h to vanish in a Roe solver is treated in [34]. Different approaches to satisfy the well-balanced property have been proposed. The Roe solver [39] has been modified in order to preserve steady states in Bermudez and Vasquez [6]. A two-dimensional extension is performed in Bermudez et al [5] and recent extensions to other types of homogeneous solvers can be found in [12, 13]. In [27] Jin proposes an other way to adapt the Roe solver to the non homogeneous case. Following the idea of Leroux et al. [19] for the scalar case, Gosse [17, 18] or Gallouët et al. [16] construct numerical schemes based on the solution of the Riemann problem associated with a larger system where a third equation on the variable describing the bottom topography is added. Another approach by LeVeque [32] is based on using the Godunov scheme for reconstructed variables. Approaches based on central schemes are used in Kurganov and Levy [30] or Russo [40]. Notice also that for the one-dimensional system, Perthame and Simeoni propose in [38] a kinetic method that includes the source term in the kinetic formulation and so that allows to ensure a discrete in-cell entropy inequality. For the simpler case of a scalar conservation law, a kinetic scheme that preserves equilibrium and which is proved to be convergent, is presented in Botchorishvili et al [11]. Excepted in the two last references, preservation of steady states and positivity of the water depth are never treated at the same time. In this paper we describe a possible solution method for the 2D Saint-Venant system using a *kinetic solver*

and a *hydrostatic reconstruction* procedure. We first present the kinetic formulation of the Saint-Venant system and how to deduce a macroscopic scheme - for the homogeneous case - with good stability properties as the built-in preservation of the water depth positivity even when applications with dry areas are considered and existence of a discrete in-cell entropy inequality. We refer to Perthame [36] for a survey of the theoretical properties of the kinetic schemes. Then we present a *well-balanced* extension that is based on a *hydrostatic reconstruction* of the water depth and preserves the still water equilibrium while satisfying the stability properties of the homogeneous solver. We finally present and detail a formally second order extension which increases the accuracy of the results while also preserving the equilibrium and stability properties of the scheme.

The outline of the paper is the following. After recalling the 2D shallow water equations and their main properties in Section 2, we introduce the kinetic representation of the system in Section 3. In Section 4 we deduce a standard kinetic scheme, then in Section 5 we introduce the hydrostatic reconstruction and we deduce a *well-balanced* kinetic scheme which is adapted to the still water equilibrium. In Section 6, a second order extension is developed. The Section 7 illustrates the possibilities of the method by numerical results.

2 THE SAINT-VENANT SYSTEM

2.1 Equations

We consider the 2D Saint-Venant system, written in its physical conservative form

$$\frac{\partial h}{\partial t} + \mathrm{div}(h\mathbf{u}) = 0, \tag{2.1}$$

$$\frac{\partial h\mathbf{u}}{\partial t} + \mathrm{div}(h\mathbf{u}\otimes\mathbf{u}) + \nabla(\frac{g}{2}h^2) + gh\nabla Z = 0, \tag{2.2}$$

where we denote $h(t,x,y) \geq 0$, the water depth, $\mathbf{u}(t,x,y) = (u,v)^T$ the flow velocity, g the acceleration due to gravity intensity and $Z(x,y)$ the bottom depth, and therefore $h+Z$ is the water surface level (see Fig. 2.1). We denote also $\mathbf{q}(t,x,y) = (q_x,q_y)^T = h(t,x,y)\mathbf{u}(t,x,y)$ the flux of water.

To obtain a well-posed problem we add to this system some initial conditions

$$h(0,x,y) = h^0(x,y), \qquad \mathbf{u}(0,x,y) = \mathbf{u}^0(x,y), \tag{2.3}$$

and boundary conditions. In this paper we consider only the following types of boundaries:

- Solid walls on which we prescribe a slip condition

$$\mathbf{u}\cdot\mathbf{n} = 0 \tag{2.4}$$

with $\mathbf{n}$ the unit outward normal to the boundary,

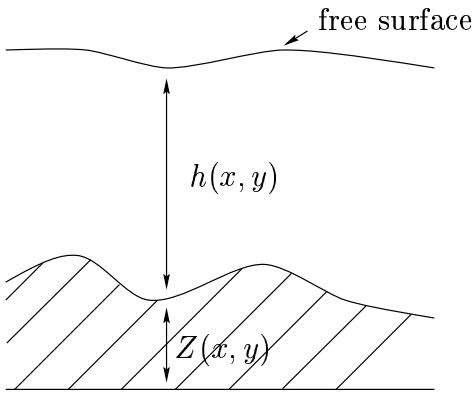


Figure 2.1:

- Fluid boundaries on which we prescribe zero, one or two of the following conditions depending of the type of the flow (fluvial or torrential)

  –Water level $h_g + Z$ given,

  –Flux $\mathbf{q}_g$ given.

2.2 Properties of the system

The system (2.1)-(2.2) is a first order conservation laws system and can be written in the general form

$$\frac{\partial \mathbf{U}}{\partial t} + \text{div } \mathbf{F}(\mathbf{U}) = \mathbf{B}(\mathbf{U}), \tag{2.5}$$

with $\mathbf{U} = (h, q_x, q_y)^T$ and

$$\mathbf{F}(\mathbf{U}) = \begin{pmatrix} q_x & q_y \\ \frac{q_x^2}{h} + \frac{g}{2}h^2 & \frac{q_x q_y}{h} \\ \frac{q_x q_y}{h} & \frac{q_y^2}{h} + \frac{g}{2}h^2 \end{pmatrix}, \qquad \mathbf{B}(\mathbf{U}) = \begin{pmatrix} 0 \\ -gh\partial_x Z \\ -gh\partial_y Z \end{pmatrix}. \tag{2.6}$$

This system is *strictly hyperbolic* for $h > 0$ (see [9]). It admits an *invariant region* $h(t,x) \geq 0$, the water depth h can indeed vanish (flooding zones, dry regions, tidal flats) and the system loses hyperbolicity at $h = 0$ which generates theoretical and numerical difficulties.

The system is also concerned with a fundamental *entropy property* which is given in the following theorem.

Theorem 2.1 *The system (2.1)–(2.2) admits a mathematical entropy (which is also the energy)*

$$E(h,u,Z) = h\frac{|\mathbf{u}|^2}{2} + \frac{gh^2}{2} + ghZ, \tag{2.7}$$

which satisfies

$$\frac{\partial}{\partial t}(E) + \mathrm{div}[\mathbf{u}\,(E + \frac{gh^2}{2})] \leq 0. \tag{2.8}$$

We do not prove this theorem which relies on the classical theory of hyperbolic equations and simple algebraic calculation, see Serre [42], Dafermos [14]. We just recall that for smooth solutions the inequality in (2.8) is an equality.

Another important property is that the Saint-Venant system admits *steady states*. They are characterized by

$$\begin{aligned} \mathrm{div}(h\mathbf{u}) &= 0, \\ \nabla P - \mathrm{curl}\ \mathbf{u}\begin{pmatrix} v \\ u \end{pmatrix} &= 0, \end{aligned}$$

where

$$P(x,y) = \frac{|\mathbf{u}|^2}{2} + g(h+Z).$$

It follows in particular that P is constant along streamlines and in the irrotationnal areas. For 2D problems, we are interested particularly by the so-called *lake at rest steady state*

$$\mathbf{u} = 0, \qquad h + Z = H, \tag{2.9}$$

where H is a constant. It will be important that the numerical scheme also satisfies this property.

For 1D flows, steady states with non vanishing velocity are also characterized by simple relations

$$hu = C_1, \qquad \frac{|u|^2}{2} + g(h+Z) = C_2, \tag{2.10}$$

where C_1 and C_2 are two constants. These relations (2.10) are used to compute the exact solution of the 1D flows in a 2D channel with a bump at the bottom, which are presented in Sec. 7.

3 KINETIC REPRESENTATION

We introduce a kinetic approach to system (2.1)–(2.2) and in the next section we deduce from the discretization of the corresponding kinetic equation, a kinetic scheme for this system.

Let $\chi(w)$ be a positive, even function defined on $\mathbb{R}^2$ i.e.

$$\chi(w) = \chi(-w) \geq 0 \tag{3.1}$$

and satisfying

$$\int_{\mathbb{R}^2} \begin{pmatrix} 1 \\ w_i w_j \end{pmatrix} \chi(w)\ dw = \begin{pmatrix} 1 \\ \delta_{ij} \end{pmatrix} \tag{3.2}$$

with δ_{ij} the Kronecker symbol.
In addition we assume that $\chi(w)$ is compactly supported, i.e.

$$\exists w_M \in \mathbb{R},\ \text{such that} \quad \chi(w) = 0 \text{ for } \|w\| \geq w_M. \tag{3.3}$$

An example of function χ satisfying these properties is

$$\chi(w) = \frac{1}{12} \mathbb{1}_{|w_i| \leq \sqrt{3}},\ i = 1, 2. \tag{3.4}$$

We introduce a microscopic density of particles $M(t, x, y, \xi)$ defined by a so-called *Gibbs equilibrium*

$$M(t, x, y, \xi) = M(h, \xi - \mathbf{u}) = \frac{h(t, x, y)}{\tilde{c}^2}\ \chi(\frac{\xi - \mathbf{u}(t, x, y)}{\tilde{c}}), \tag{3.5}$$

with $\tilde{c}$ defined by

$$\tilde{c}^2 = \frac{g\, h}{2}. \tag{3.6}$$

With these definitions we can write a kinetic interpretation of the system (2.1)–(2.2) and we have the following statement:

Theorem 3.1 *The functions $(h, \mathbf{q})$ are strong solutions to the system (2.1)–(2.2) or (2.5)–(2.6) if and only if $M(t, x, y, \xi)$ satisfies the kinetic equation*

$$\frac{\partial M}{\partial t} + \xi \cdot \nabla_{\mathbf{x}} M - g \nabla Z \cdot \nabla_{\xi} M = Q(t, x, y, \xi), \tag{3.7}$$

for some "collision term" $Q(t, x, y, \xi)$ which satisfies for a.e. $(t,\ x, y)$,

$$\int_{\mathbb{R}^2} \begin{pmatrix} 1 \\ \xi \end{pmatrix} Q\ d\xi = 0. \tag{3.8}$$

Proof. The proof relies on a very obvious computation. The two Saint-Venant equations are equivalent with the equation (3.7) once integrated in ξ against 1 and ξ. These are consequences of the usual relations deduced from the properties of χ and from (3.8):

$$\begin{pmatrix} h \\ \mathbf{q} \\ \dfrac{\mathbf{q} \otimes \mathbf{q}}{h} + \dfrac{g}{2}\, h^2 \mathbf{Id} \end{pmatrix} = \int_{\mathbb{R}^2} \begin{pmatrix} 1 \\ \xi \\ \xi \otimes \xi \end{pmatrix} M(\xi)\ d\xi, \tag{3.9}$$

and

$$\begin{pmatrix} 0 \\ h \end{pmatrix} = -\int_{\mathbb{R}^2} \begin{pmatrix} 1 \\ \xi \end{pmatrix} \nabla_{\xi} M(\xi)\ d\xi. \tag{3.10}$$

□

This theorem produces a very useful consequence: the non-linear system (2.1)–(2.2) can be viewed as a linear transport equation on a nonlinear quantity M, for which it is easier to find a simple numerical scheme with good theoretical properties.

The mathematical entropy property, for the energy, can also be considered in terms of the kinetic approach. Indeed if the "collision term" $Q(t,x,y,\xi)$ satisfies also for a.e. $(t,\ x,y)$,

$$\int_{\mathbb{R}^2} |\xi|^2 Q\ d\xi \leq 0,$$

the entropy inequality (2.8) is equivalent with the equation (3.7) once integrated in ξ against $|\xi|^2/2 + gZ$. It is a consequence of the following relation between the kinetic density M and the macroscopic energy E

$$E(t,x,y) = \int_{\mathbb{R}^2} \left(\frac{|\xi|^2}{2} + gZ \right)\ M(t,x,y,\xi)\ d\xi. \tag{3.11}$$

Remark 3.1 *The relation between M and E is more complex in the one-dimensional case since it requires to take into account the transverse translational energy through a cubic term [38].*

Among the functions that verify (3.1)-(3.3) one of them can be computed through a minimization problem. In particular it justifies the interpretation of such a density $M(t,x,y,\xi)$ as the microscopic equilibrium of the system.

Proposition 3.1 *The minimum of the energy*

$$\epsilon(f) = \int_{\mathbb{R}^2} \left(\frac{|\xi|^2}{2} + gZ \right) f(\xi) d\xi,$$

under the constraints

$$0 \leq f \leq \frac{1}{4\pi}, \qquad \int_{\mathbb{R}^2} f(\xi) d\xi = h, \qquad \int_{\mathbb{R}^2} \xi f(\xi) d\xi = h\mathbf{u},$$

is achieved by the function $M(t,x,y,\xi)$ defined by (3.5) with χ given by

$$\chi(w) = \frac{1}{4\pi} \mathbb{1}_{\|w\| \leq 2}. \tag{3.12}$$

Proof. The form of the constraints and of the functional that we minimize leads to search solutions of the form

$$M(t,x,y,\xi) = A(t,x,y) \mathbb{1}_{\|\xi - \mathbf{u}\| \leq B(t,x,y)}$$

and the results follows obviously. □

Remark 3.2 *The minimum $M(t,x,y,\xi)$ computed in Proposition 3.1 is of the form (3.1)-(3.3), (3.5) and then it verifies (3.11). It follows that $\epsilon(M(t,x,y,\xi)) = E(t,x,y)$ where $E(t,x,y)$ is the macroscopic energy of the Saint-Venant system defined in (2.7). This justifies the choice of the constants in the constraints in Proposition 3.1.*

Remark 3.3 *The function χ defined by (3.12) is also the only choice such that $M(t,x,y,\xi)$ defined by (3.5) preserves the microscopic lake at rest steady state, i.e. satisfies the kinetic equation (3.7) with $Q(t,x,y,\xi) = 0$ on all macroscopic steady states associated to a lake at rest defined by (2.9). For the one-dimensional Saint-Venant system, the authors of [38] use the equivalent 1d property to derive a well-balanced kinetic scheme by integrating the topographic source term in the kinetic fluxes.*

4 FINITE VOLUMES / KINETIC SOLVER

A classical approach for solving hyperbolic systems consists in using finite volume schemes (see [21]) which are defined by the fluxes computed at the control volume interfaces. We show in Section 4.2 how the fluxes of the kinetic scheme are deduced from the discretization of the kinetic equation (3.7). In this section we do not take into account the topographic source term, i.e. we first consider the homogeneous system of equations (2.5) with $\mathbf{B} = 0$. In Section 5 we adapt the scheme to the non flat bottom case.

4.1 Finite volume formalism

We recall here the general formalism of finite volumes. Let Ω denote the computational domain with boundary Γ, which we assume polygonal. Let $\mathcal{T}_h$ be a triangulation of Ω which vertices are denoted P_i with $\mathbf{S_i}$ the set of interior nodes and $\mathbf{G_i}$ the set of boundary nodes. The dual cells C_i are obtained by joining the centers of mass of the triangles surrounding each vertex P_i. We use the following notations (see Fig. 4.1):

- K_i, set of subscripts of nodes P_j surrounding P_i,
- $|C_i|$, area of C_i,
- Γ_{ij}, boundary edge between the cells C_i and C_j,
- L_{ij}, length of Γ_{ij},
- $\mathbf{n}_{ij}$, unit normal to Γ_{ij}, outward to C_i ($\mathbf{n}_{ji}$=-$\mathbf{n}_{ij}$).

If P_i is a node belonging to the boundary Γ, we join the centers of mass of the triangles adjacent to the boundary to the middle of the edge belonging to Γ (see Fig. 4.2) and we denote

- Γ_i, the two edges of C_i belonging to Γ,
- L_i, length of Γ_i (for sake of simplicity we assume in the following that $L_i = 0$ if P_i does not belong to Γ),
- $\mathbf{n}_i$, the unit outward normal defined by averaging the two adjacent normals.

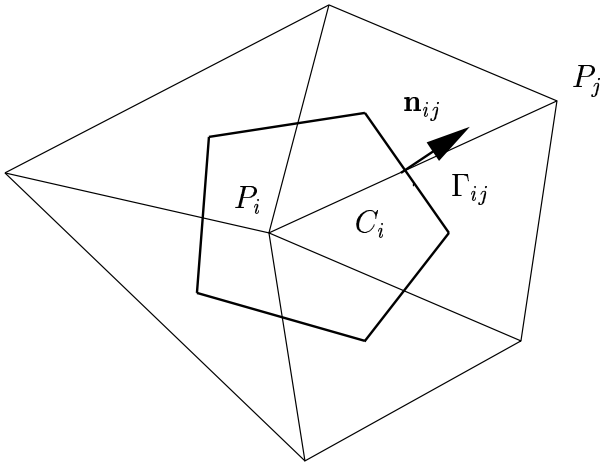


Figure 4.1: Dual cell C_i.

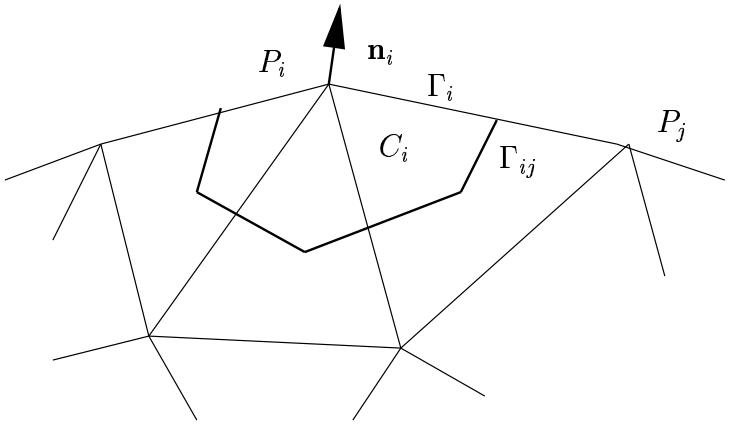


Figure 4.2: Boundary cell C_i.

Let Δt be the timestep, we set $t^n = n\ \Delta t$. We denote by $\mathbf{U}_i^n$ the approximation of the cell average of the exact solution at time t^n

$$\mathbf{U}_i^n \simeq \frac{1}{|C_i|}\int_{C_i} \mathbf{U}(t^n, \mathbf{x})\ d\mathbf{x}. \tag{4.1}$$

We integrate in space and time the equation (2.5) on the set $C_i \times (t^n, t^{n+1})$, and, integrating by parts the divergence term, we obtain

$$\int_{C_i} \mathbf{U}(t^{n+1}, \mathbf{x})\ d\mathbf{x} - \int_{C_i} \mathbf{U}(t^n, \mathbf{x})\ d\mathbf{x} + \int_{t^n}^{t^{n+1}} \int_{\partial C_i} \mathbf{F}(\mathbf{U}) \cdot \mathbf{n}\ d\mathbf{x}dt = 0. \tag{4.2}$$

So we can write

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \sum_{j \in K_i} \sigma_{ij} \mathcal{F}(\mathbf{U}_i^n, \mathbf{U}_j^n, \mathbf{n}_{ij}) - \sigma_i \mathcal{F}(\mathbf{U}_i^n, \mathbf{U}_{e,i}^n, \mathbf{n}_i) \tag{4.3}$$

with

$$\sigma_{ij} = \frac{\Delta t\, L_{ij}}{|C_i|}, \qquad \sigma_i = \frac{\Delta t\, L_i}{|C_i|}, \tag{4.4}$$

In (4.3) the term $\mathcal{F}(\mathbf{U}_i, \mathbf{U}_j, \mathbf{n}_{ij})$ denotes an interpolation of the normal component of the flux $\mathbf{F}(\mathbf{U}) \cdot \mathbf{n}_{ij}$ along the edge Γ_{ij}. This interpolation is usually performed using a one-dimensional solver since locally the problem looks like a planar discontinuity. In the next section we define $\mathcal{F}(\mathbf{U}_i, \mathbf{U}_j, \mathbf{n}_{ij})$ using the kinetic interpretation of the system. The computation of the value $\mathbf{U}_{e,i}$ which denotes a value outside C_i defined such that the boundary conditions are satisfied and the definition of the boundary flux $\mathcal{F}(\mathbf{U}_i, \mathbf{U}_{e,i}, \mathbf{n}_i)$ are detailed in Section 4.5.

4.2 Kinetic solver

In the following of the present section we assume that P_i is an interior point. Being given the solution $\mathbf{U}_i^n$ at time t^n for each cell, we compute $\mathbf{U}_i^{n+1}$ by the following algorithm with three steps:

- We define $M_i^n = M(h_i^n, \xi - \mathbf{u}_i^n)$ with M defined by (3.5).
- We use the microscopic equation (3.7). Since this equation is linear, we can apply a simple upwind scheme [21] which defines a density function $f_i^{n+1}(\xi)$

  $$f_i^{n+1}(\xi) - M_i^n(\xi) + \frac{\Delta t}{|C_i|} \sum_{j \in K_i} L_{ij}\ \xi \cdot \mathbf{n}_{ij}\ M_{ij}^n(\xi) = 0, \tag{4.5}$$

  with the fluxes $M_{ij}^n(\xi)$ computed by the upwind formula

  $$M_{ij}^n(\xi) = \begin{cases} M_i^n(\xi) & \text{for } \xi \cdot \mathbf{n}_{ij} \geq 0, \\ M_j^n(\xi) & \text{for } \xi \cdot \mathbf{n}_{ij} \leq 0. \end{cases}$$

  Notice however that the density function $f(\xi)$ is not an equilibrium (see remark (4.1)).
- Nevertheless, by analogy with the computations in the proof of Theorem (3.1), we can recover the macroscopic quantities $\mathbf{U}_i^{n+1}$ at time t^{n+1} by integration

  $$\mathbf{U}_i^{n+1} = \int_{\mathbb{R}^2} \begin{pmatrix} 1 \\ \xi \end{pmatrix}\ f_i^{n+1}(\xi)\ d\xi. \tag{4.6}$$

Remark 4.1 *The interpretation is that, as usual, the collision term Q, which forces the relaxation of f to Gibbs equilibrium M, is neglected in the advection scheme (4.5). And at each timestep we deduce $M_i^{n+1}(\xi)$ from $\mathbf{U}_i^{n+1}$ which is a way to perform all collisions at once and to recover the Gibbs equilibrium without computing them explicitely.*

The numerical conssitency of the kinetic solver relies on the fact that if we consider the exact solution of the homogeneous kinetic transport equation - (3.7) with $\nabla Z = 0$ and $Q = 0$ - we can prove that the macroscopic quantities obtained through the integration process describes previously are first order approximations in time of the solutions of the Saint-Venant system - see [21].

The numerical feasability of the kinetic solver relies on the possibility to write directly a finite volume formula, which therefore avoids using the extra variable ξ in the actual implementation. Indeed, the equation (4.6) can be written with the form (4.3) with

$$\mathcal{F}(\mathbf{U}_i, \mathbf{U}_j, \mathbf{n}_{ij}) = \mathbf{F}^+(\mathbf{U}_i, \mathbf{n}_{ij}) + \mathbf{F}^-(\mathbf{U}_j, \mathbf{n}_{ij}), \tag{4.7}$$

and

$$\mathbf{F}^+(\mathbf{U}_i, \mathbf{n}_{ij}) = \int_{\xi\cdot\mathbf{n}_{ij}\geq 0} \xi\cdot\mathbf{n}_{ij} \begin{pmatrix} 1 \\ \xi \end{pmatrix} M_i(\xi)\, d\xi, \tag{4.8}$$

$$\mathbf{F}^-(\mathbf{U}_j, \mathbf{n}_{ij}) = \int_{\xi\cdot\mathbf{n}_{ij}\leq 0} \xi\cdot\mathbf{n}_{ij} \begin{pmatrix} 1 \\ \xi \end{pmatrix} M_j(\xi)\, d\xi. \tag{4.9}$$

Notice that (4.8)-(4.9) imply

$$\mathrm{F}_h^+(\mathbf{U}_i, \mathbf{n}_{ij}) \geq 0, \qquad \mathrm{F}_h^-(\mathbf{U}_j, \mathbf{n}_{ij}) \leq 0. \tag{4.10}$$

4.3 Numerical implementation

We give here some details on the implementation of the kinetic scheme defined by (4.3), (4.7)-(4.9). For the efficiency of the method, we code in fact a variant where the choice of the function χ depends on the interface under consideration. For an interface with unit normal $\mathbf{n} = (n_x, n_y)^T$, we define a local basis (n, τ) associated to the normal direction and to the tangential one. We denote $\hat{\mathbf{U}}_\mathbf{n} = (h, q_n, q_\tau)^T$, the vector deduced from $\mathbf{U}$ by the rotation in this new basis and $\hat{\mathbf{u}} = (u_n, u_\tau)^T = (\frac{q_n}{h}, \frac{q_\tau}{h})^T$. So we have $\hat{\mathbf{U}}_\mathbf{n}$ defined by

$$\hat{\mathbf{U}}_\mathbf{n} = \mathbf{R}_\mathbf{n}\ \mathbf{U} \quad \text{with} \quad \mathbf{R}_\mathbf{n} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & n_x & n_y \\ 0 & -n_y & n_x \end{pmatrix} \tag{4.11}$$

and

$$\mathbf{F}^+(\mathbf{U}, \mathbf{n}) = \mathbf{R}_\mathbf{n}^{-1} \hat{\mathbf{F}}^+(\hat{\mathbf{U}}_\mathbf{n}). \tag{4.12}$$

Using (4.8), we give the detailed expression of $\hat{\mathbf{F}}^+(\hat{\mathbf{U}}_i)$ related to the interface Γ_{ij}

$$\hat{\mathbf{F}}^+(\hat{\mathbf{U}}_{i,\mathbf{n}_{ij}}) = \frac{h_i}{\tilde{c}_i^2} \int_{\{\xi_n \geq 0\}\times\mathbb{R}} \xi_n \begin{pmatrix} 1 \\ \xi \end{pmatrix} \chi\left(\frac{\xi - \hat{\mathbf{u}}_i}{\tilde{c}_i}\right) d\xi \tag{4.13}$$

or, after the change of variables $w = \dfrac{\xi - \hat{\mathbf{u}}_i}{\tilde{c}_i}$,

$$\hat{\mathbf{F}}^+(\hat{\mathbf{U}}_{i,\mathbf{n}_{ij}}) = h_i \int_{\{w_n \geq \frac{-u_{i,n}}{\tilde{c}_i}\}\times\mathbb{R}} (u_{i,n} + w_n\,\tilde{c}_i) \begin{pmatrix} 1 \\ u_{i,n} + w_n\,\tilde{c}_i \\ u_{i,\tau} + w_\tau\,\tilde{c}_i \end{pmatrix} \chi(w)\,dw, \tag{4.14}$$

we have dropped here the subscript ij for the components u_n, u_τ.

Due to the fact that $\chi(w)$ is even, the term with w_τ disappears in (4.14) and we obtain the simpler formula

$$\hat{\mathrm{F}}^+_{u_\tau}(\hat{\mathbf{U}}_{i,\mathbf{n}_{ij}}) = \hat{u}_{i,\tau}\ \hat{\mathrm{F}}^+_h(\hat{\mathbf{U}}_{i,\mathbf{n}_{ij}}). \tag{4.15}$$

We obtain $\mathbf{F}^-$ by an analogous computation and so we have the same property for $\hat{\mathcal{F}}_{u_\tau}$. We will use this property to deduce a modified scheme with better accuracy.

4.4 Upwind kinetic scheme

In order to reduce the diffusion of the scheme, we modify the computation of the flux related to the tangential component. For the computation of $u_{i,\tau}^{n+1}$ we replace the expression of $\hat{\mathcal{F}}$ by the following:

$$\hat{\mathcal{F}}_{u_\tau}(\hat{\mathbf{U}}_{i,\mathbf{n}_{ij}}, \hat{\mathbf{U}}_{j,\mathbf{n}_{ij}}) = u_{ij,\tau}\ \hat{\mathcal{F}}_h(\hat{\mathbf{U}}_{i,\mathbf{n}_{ij}}, \hat{\mathbf{U}}_{j,\mathbf{n}_{ij}}) \tag{4.16}$$

with

$$u_{ij,\tau} = \begin{cases} u_{i,\tau} & \text{for } \hat{\mathcal{F}}_h \geq 0, \\ u_{j,\tau} & \text{for } \hat{\mathcal{F}}_h \leq 0. \end{cases} \tag{4.17}$$

Formula (4.17) introduces some upwinding depending on the sign of the total flux. We give in [10] a numerical result showing the efficiency of (4.16)-(4.17).

4.5 Boundary conditions

The treatment of the boundary conditions is presented in details in [9]. Here we just recall some main features about the computation of the boundary flux $\mathcal{F}(\mathbf{U}_i^n, \mathbf{U}_{e,i}^n, \mathbf{n}_i)$ appearing in (4.3). Notice first that the variable $\mathbf{U}_{e,i}^n$ can be interpreted as an approximation of the solution in a "fictitious" cell adjacent to the boundary. As before we introduce the local coordinates and define $\hat{\mathbf{U}}_{e,i}^n = \left(h_{e,i}^n, q_{e,i,n}^n, q_{e,i,\tau}^n\right)^T$. Then we can use the local flux vector splitting form associated to the kinetic formulation

$$\hat{\mathcal{F}}(\hat{\mathbf{U}}_{i,\mathbf{n}_i}^n, \hat{\mathbf{U}}_{e,i}^n) = \hat{\mathbf{F}}^+(\hat{\mathbf{U}}_{i,\mathbf{n}_i}^n) + \hat{\mathbf{F}}^-(\hat{\mathbf{U}}_{e,i}^n).$$

On the solid wall we prescribe a continuous slip condition - see Section 2. In the numerical scheme we prescribe it weakly by defining $\hat{\mathbf{U}}^n_{e,i} = \left(h^n_i, -q^n_{i,n}, q^n_{i,\tau}\right)^T$. It follows that finally

$$\hat{\mathcal{F}}(\hat{\mathbf{U}}^n_{i,\mathbf{n}_i}, \hat{\mathbf{U}}^n_{e,i}) = (0,\ \frac{g h_i^{n\,2}}{2},\ 0)^T. \tag{4.18}$$

On the fluid boundaries, the type of the flow and then the number of boundary conditions depend on the Froude number. Here we consider a local Froude number associated to the normal component of the velocity. For the fluvial cases, we define completely $\mathbf{U}_e$ by adding to the given boundary condition, the assumption that the *Riemann invariant* related to the outgoing characteristic is constant along this characteristic (see [9]).

4.6 Properties of the scheme

It is clear from (4.3), (4.7)-(4.9) that the scheme is consistent and conservative. The CFL condition for the explicit scheme (4.5) applied to the linear microscopic equation writes

$$\Delta t \leq \min \frac{|C_i|}{(L_i + \sum_{j\in K_i} L_{ij})\ (|\mathbf{u}^n_i| + w_M \tilde{c}^n_i)}. \tag{4.19}$$

Besides we have the following stability theorem:

Theorem 4.1 *Under the CFL condition (4.19), the kinetic scheme (4.3), (4.7), (4.12), (4.14) preserves the water depth positivity.*

Proof. Suppose that we have $h^n_i \geq 0$. It follows from the definitions (4.3), (4.7), (4.12), (4.14) that - here and all along the paper, when superscripts are omitted, the quantities have to be taken at time t^n

$$h^{n+1}_i = h^n_i - \sum_{j\in K_i} \sigma_{ij}(\hat{\mathrm{F}}^+_h(\hat{\mathbf{U}}^n_{i,\mathbf{n}_{ij}}) + \hat{\mathrm{F}}^-_h(\hat{\mathbf{U}}^n_{j,\mathbf{n}_{ij}})) - \ \sigma_i(\hat{\mathrm{F}}^+_h(\hat{\mathbf{U}}^n_{i,\mathbf{n}_i}) + \hat{\mathrm{F}}^-_h(\hat{\mathbf{U}}^n_{e,i})). \tag{4.20}$$

The relations (4.10) and (4.12) imply

$$\hat{\mathrm{F}}^-_h(\hat{\mathbf{U}}^n_{j,\mathbf{n}_{ij}}) \leq 0, \qquad \hat{\mathrm{F}}^-_h(\hat{\mathbf{U}}^n_{e,i}) \leq 0, \tag{4.21}$$

and, using the detailed expression of the flux (4.14),we have

$$\begin{aligned} h^{n+1}_i \geq h^n_i \qquad \Bigg(1 - \sum_{j\in K_i} \sigma_{ij} \int_{\{w_n \geq \frac{-u_{i,nij}}{\tilde{c}_i}\}\times\mathbb{R}} (u_{i,nij} + w_n\, \tilde{c}_i)\, \chi(w)\, dw \\ - \sigma_i \int_{\{w_n \geq \frac{-u_{i,ni}}{\tilde{c}_i}\}\times\mathbb{R}} (u_{i,ni} + w_n\, \tilde{c}_i)\, \chi(w)\, dw \Bigg). \end{aligned} \tag{4.22}$$

Since χ satisfies (3.1)-(3.2), we have for each n

$$\int_{\{w_n \geq \frac{-u_{i,n}}{\tilde{c}_i}\}\times\mathbb{R}} (u_{i,n} + w_n\,\tilde{c}_i)\,\chi(w)\,dw \leq |u_{i,n}| + \tilde{c}_i \int_{\{w_n \geq 0\}\times\mathbb{R}} w_n \chi(w)\,dw, \qquad (4.23)$$

and from (3.1)-(3.3) we deduce

$$\int_{\{w_n \geq 0\}\times\mathbb{R}} w_n \chi(w)\,dw \leq \int_{\{0 \leq w_n \leq 1\}\times\mathbb{R}} \chi(w)\,dw + \int_{\{w_n \geq 1\}\times\mathbb{R}} w_n^2 \chi(w)\,dw \leq 1. \quad (4.24)$$

Finally using (4.23)-(4.24) in (4.22) we obtain

$$h_i^{n+1} \geq h_i^n \left(1 - \frac{\Delta t}{|C_i|}[\sum_{j\in K_i} L_{ij}\,(|u_{i,nij}| + \tilde{c}_i) + L_i\,(|u_{i,ni}| + \tilde{c}_i)]\right),$$

and it follows the positivity of h_i^{n+1} under the CFL condition (4.19) (from (3.2) we have $w_M \geq 1$). □

In the particular case where the function χ is defined through the minimization problem of Proposition 3.1, the deduced kinetic scheme satisfies another stability property as shown in the following theorem.

Theorem 4.2 *Under the CFL condition (4.19) and with the particular choice where χ is defined by (3.12) the kinetic scheme (4.3), (4.7), (4.12), (4.14) satisfies a conservative in-cell entropy inequality.*

Proof. In this case χ is invariant by rotation and then is the same one whatever is the interface under consideration. It follows that we can establish the proof at the kinetic level. Integrating the microscopic scheme (4.5) in ξ against $\xi^2/2$ we obtain - with the notations of Proposition 3.1

$$\epsilon\left(f_i^{n+1}\right) - E_i^n + \frac{\Delta t}{|C_i|} \sum_{j\in K_i} L_{ij} \int_{\mathbb{R}^2} \frac{|\xi|^2}{2}\,\xi\cdot\mathbf{n}_{ij}\; M_{ij}^n(\xi)d\xi = 0.$$

It follows from Proposition 3.1 - with $Z = 0$ - that the particular choice of χ implies

$$E_i^{n+1} - E_i^n + \sum_{j\in K_i} \sigma_{ij}\,\eta(\mathbf{U}_i^n, \mathbf{U}_j^n, \mathbf{n}_{ij}) \leq 0,$$

where $\eta(\mathbf{U}_i^n, \mathbf{U}_j^n, \mathbf{n}_{ij})$ denotes the entropy flux defined by

$$\eta(\mathbf{U}_i^n, \mathbf{U}_j^n, \mathbf{n}_{ij}) = \int_{\mathbb{R}^2} \frac{|\xi|^2}{2}\,\xi\cdot\mathbf{n}_{ij}\; M_{ij}^n(\xi)d\xi.$$

This concludes the proof. □

5 WELL-BALANCED SCHEME

In order to be able to compute realistic flows we consider now the case $\nabla Z \neq 0$. As motivated in the introduction the *well-balanced* requirement is to preserve a discrete equivalent of the continuous lake at rest steady state (2.9). Then we modify the original scheme in order to satisfy the following discrete local lake at rest steady state

$$\left(\begin{array}{ll} \forall j \in K_i & h_j^n + Z_j = h_i^n + Z_i = H \\ & \mathbf{u}_j^n = \mathbf{u}_i^n = 0 \end{array}\right) \quad \Rightarrow \quad \begin{array}{l} h_i^{n+1} + Z_i = H \\ \mathbf{u}_i^{n+1} = 0 \end{array} . \tag{5.1}$$

The well-balanced scheme that we present in this section is an adaptation to the two-dimensional flows of the *interface hydrostatic reconstruction* method developed by Audusse et al. in [3] in the 1d framework.
We first construct a piecewise constant approximation of the bottom topography $Z(x)$

$$Z_i = \frac{1}{|C_i|}\int_{C_i} Z(x)dx.$$

We define an interface topography

$$Z_{ij}^* = Z_{ji}^* = max(Z_i, Z_j), \tag{5.2}$$

and then we define new interface values by $\mathbf{U}_{ij}^* = (h_{ij}^*, h_{ij}^* \mathbf{u}_i)^T$ where h_{ij}^* is the *hydrostatic reconstructed* water depth

$$h_{ij}^* = (h_i + Z_i - Z_{ij}^*)_+. \tag{5.3}$$

We write the well-balanced scheme with the form

$$\begin{aligned} \mathbf{U}_i^{n+1} = \mathbf{U}_i^n \quad & - \quad \sum_{j\in K_i} \sigma_{ij} \mathcal{F}(\mathbf{U}_{ij}^{*,n}, \mathbf{U}_{ji}^{*,n}, \mathbf{n}_{ij}) - \sigma_i \mathcal{F}(\mathbf{U}_i^n, \mathbf{U}_{e,i}^n, \mathbf{n}_i) \\ & + \quad \sum_{j\in K_i} \sigma_{ij} S(\mathbf{U}_i^n, \mathbf{U}_{ij}^{*,n}, \mathbf{n}_{ij}), \end{aligned} \tag{5.4}$$

where

$$S(\mathbf{U}_i, \mathbf{U}_{ij}^*, \mathbf{n}_{ij}) = \left(\begin{array}{c} 0 \\ \frac{g}{2}(h_{ij}^{*2} - h_i^2)\mathbf{n}_{ij} \end{array} \right). \tag{5.5}$$

Theorem 5.1 *The scheme defined by (5.4)–(5.5) with (4.7)– (4.9) satisfies the following properties*

(i) it preserves the water depth positivity under the CFL condition

$$\Delta t \leq \min \frac{|C_i|}{\sum_{j\in K_i} \left[L_{ij}\ (|\mathbf{u}_i^n| + w_M \tilde{c}_{ij}^{*,n})\right] + L_i\ (|\mathbf{u}_i^n| + w_M \tilde{c}_i^n)}, \tag{5.6}$$

a fortiori if Δt satisfies (4.19),

(ii) it preserves the steady state of a lake at rest.

Proof. For (i), first we prove as in Theorem 4.1 that the water depth positivity is preserved under the CFL condition (5.6). Then as the definition (5.3) implies that $h_{ij}^* \leq h_i$, we deduce that the CFL condition (4.19) is more restrictive than (5.6).

To prove (ii), we assume that the solution at time n satisfies (5.1), then we have

$$\sum_{j\in K_i} \sigma_{ij} \mathcal{F}(\mathbf{U}_{ij}^{*,n}, \mathbf{U}_{ji}^{*,n}, \mathbf{n}_{ij}) = \sum_{j\in K_i} \sigma_{ij} \begin{pmatrix} 0 \\ \frac{g}{2} h_{ij}^{*2} \mathbf{n}_{ij} \end{pmatrix}. \tag{5.7}$$

Concerning the boundary term, we assume also that the boundary conditions will preserve the steady state, so they can be either a slip condition, a given flux $\mathbf{q} = 0$ or a water depth given $h + Z = H$. Following the treatment of the boundary conditions developed in [9] the boundary term reduces to

$$\sigma_i \mathcal{F}(\mathbf{U}_i^n, \mathbf{U}_{e,i}^n, \mathbf{n}_i) = \begin{pmatrix} 0 \\ \frac{g}{2} h_i^2 \mathbf{n}_i \end{pmatrix}. \tag{5.8}$$

From (5.4)–(5.5), we obtain finally

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \sum_{j\in K_i} \sigma_{ij} \begin{pmatrix} 0 \\ \frac{g}{2} h_i^2 \mathbf{n}_{ij} \end{pmatrix} - \sigma_i \begin{pmatrix} 0 \\ \frac{g}{2} h_i^2 \mathbf{n}_i \end{pmatrix} \tag{5.9}$$

and using the property

$$\sum_{j\in K_i} L_{ij} \mathbf{n}_{ij} + L_i \mathbf{n}_i = 0,$$

this proves the preservation of the steady state. □

Remark 5.1 *From the definitions (5.2)–(5.5), it is obvious that for $Z = Cst$, we recover the original scheme.*

Remark 5.2 *It appears in the proof of Theorem 5.1 that to construct the scheme on the interface values $\mathbf{U}_{ij}^*$ instead of the cell values $\mathbf{U}_i$ allows to numerically preserve at each interface the balance between the hydrostatic pressure and the influence of the topographic source terms in a lake at rest. This explains the name interface hydrostatic reconstruction method. For further details refer to [3].*

Remark 5.3 *By constrast with the homogeneous case we can not prove that the fully discrete kinetic scheme (5.4)–(5.5) satisfies an entropy inequality, whatever is the time step. Nevertheless it is proved in [3] that the semi-discrete version of the scheme (5.4)–(5.5) satisfies a conservative in-cell entropy inequality.*

6 SECOND ORDER EXTENSION

The first order scheme defined in Sections 4-5 can be extended to a "formally" second order one using a MUSCL like extension (see [43]). In Section 6.1, we define limited reconstructed variables and in Section 6.2, we introduce a "second order" well-balanced scheme.

6.1 Second order reconstructions

In the definition of the flux (4.7), we replace the piecewise constant values $\mathbf{U}_i, \mathbf{U}_j$ by more accurate reconstructions deduced from piecewise linear approximations, namely the values $\mathbf{U}_{ij}, \mathbf{U}_{ji}$ reconstructed on both sides of the interface. More precisely, we are looking for piecewise linear approximation of the primitive variable $\hat{\mathbf{W}} = (h, u_n, u_\tau)^T$, actually the detailed expression of the flux given in (4.14) uses the primitive variables.

We divide each cell C_i in subtriangles obtained by joining each edge Γ_{ij} to the node P_i,

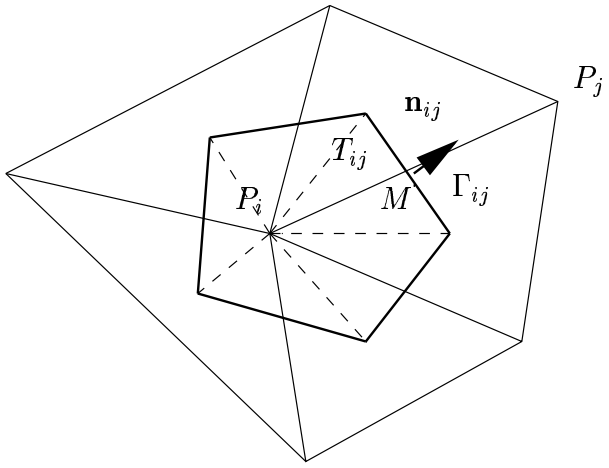


Figure 6.1: Subcells T_{ij}.

we denote T_{ij} the subtriangle related to Γ_{ij} (see Fig.(6.1)). We denote $|C_{ij}|$ the area of T_{ij}. Let M be the middle point of the interface Γ_{ij}, we define $\hat{\mathbf{W}}_{ij} = (h_{ij}, u_{ij,n}, u_{ij,\tau})^T$ as an approximation of $\hat{\mathbf{W}}$ at point M, deduced from a piecewise linear reconstruction on the subtriangle T_{ij}:

$$\hat{\mathbf{W}}_{ij} = \hat{\mathbf{W}}_i + \overrightarrow{P_i M} \cdot \nabla \hat{\mathbf{W}}_{ij} \tag{6.1}$$

with $\nabla \hat{\mathbf{W}}_{ij}$ defined here as follows (see [24]).

If the point M belongs to the triangle T_k, we denote $\nabla \hat{\mathbf{W}}_M = \nabla \hat{\mathbf{W}}|_{T_k}$ where $\nabla \hat{\mathbf{W}}|_{T_k}$ is the constant gradient of $\hat{\mathbf{W}}$ deduced from a P1 approximation on the triangle T_k. We denote by $\nabla \hat{\mathbf{W}}_i$ an approximate gradient at node P_i computed by a weighted average of the gradients on the surrounding triangles

$$\nabla \hat{\mathbf{W}}_i = \frac{\sum_{k \in \mathcal{T}_i} |C_k| \ \nabla \hat{\mathbf{W}}|_{T_k}}{\sum_{k \in \mathcal{T}_i} |C_k|} \tag{6.2}$$

and

$$\nabla \hat{\mathbf{W}}_{mi} = (1+\beta) \ \nabla \hat{\mathbf{W}}_i - \beta \ \nabla \hat{\mathbf{W}}_M, \quad 0 \leq \beta \leq 1, \tag{6.3}$$

where $\mathcal{T}_i$ is the set of triangles surrounding the node P_i.

Then we use an appropriate slope limiter to deduce $\nabla\hat{\mathbf{W}}_{ij}$

$$\nabla\hat{\mathbf{W}}_{ij} = \mathrm{Lim}(\nabla\hat{\mathbf{W}}_M, \nabla\hat{\mathbf{W}}_{mi}). \tag{6.4}$$

In the following computations we have used either the minmod limiter defined by

$$\mathrm{Lim}(a,b) = \begin{cases} 0 & \text{if } \mathrm{sign}(a) \neq \mathrm{sign}(b) \\ \mathrm{sign}(a)\min(|a|,|b|) & \text{otherwise} \end{cases}$$

or the Van Albada limiter defined by

$$\mathrm{Lim}(a,b) = \begin{cases} 0 & \text{if } \mathrm{sign}(a) \neq \mathrm{sign}(b) \\ \dfrac{a(b^2+\varepsilon)+b(a^2+\varepsilon)}{a^2+b^2+2\varepsilon} & \text{otherwise} \end{cases}$$

with $0 \leq \varepsilon \ll 1$.

Once the water depth is computed by (6.4), it is corrected (see [37]) to ensure the conservation of the reconstructed values, i.e.

$$\sum_{j\in K_i} |C_{ij}|h_{ij} = |C_i|h_i. \tag{6.5}$$

In the case where $\mathbf{B} = 0$, the "second order" scheme is obtained by replacing (4.3) by

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \sum_{j\in K_i} \sigma_{ij}\mathcal{F}(\mathbf{U}_{ij}^n, \mathbf{U}_{ji}^n, \mathbf{n}_{ij}) - \sigma_i\mathcal{F}(\mathbf{U}_i^n, \mathbf{U}_{e,i}^n, \mathbf{n}_i). \tag{6.6}$$

6.2 Second order well-balanced scheme

For the cases where $\mathbf{B} \neq 0$ we consider also piecewise linear approximation of the variable Z and we reconstruct values Z_{ij}, Z_{ji} on both sides of the interface as it is done before for the primitive variables. In fact, so that the second order scheme preserves the *well-balanced* property, it is necessary that the second order reconstruction preserves an interface equilibrium. It means that if

$$h_i + Z_i = h_j + Z_j = H, \qquad \mathbf{u}_i = \mathbf{u}_j = 0, \tag{6.7}$$

then the second order reconstructed values have to satisfy

$$h_{ij} + Z_{ij} = h_{ji} + Z_{ji} = H, \qquad \mathbf{u}_{ij} = \mathbf{u}_{ji} = 0. \tag{6.8}$$

The velocity part is obvious but - as we require also that the second order reconstruction preserves the positivity of the water depth - it is enounced in [3] that the right way to build a *well-balanced* second order scheme is to reconstruct and correct the variables $h + Z$ and h and then to deduce the interface values for Z - see [3] for further explanations, especially

for the case of dry/wet interface.
Then we adapt the first order strategy departing from the reconstructed interface values instead of the cell values. We define

$$Z_{ij}^* = Z_{ji}^* = \max(Z_{ij}, Z_{ji}) \tag{6.9}$$

then we define new interface values by $\mathbf{U}_{ij}^* = (h_{ij}^*,\ h_{ij}^* \mathbf{u}_{ij})^T$ with

$$h_{ij}^* = (h_{ij} + Z_{ij} - Z_{ij}^*)_+. \tag{6.10}$$

We write the second order well-balanced scheme with the form

$$\begin{aligned} \mathbf{U}_i^{n+1} = \mathbf{U}_i^n \quad &- \quad \sum_{j \in K_i} \sigma_{ij} \mathcal{F}(\mathbf{U}_{ij}^{*,n}, \mathbf{U}_{ji}^{*,n}, \mathbf{n}_{ij}) - \sigma_i \mathcal{F}(\mathbf{U}_i^n, \mathbf{U}_{e,i}^n, \mathbf{n}_i) \\ &+ \quad \sum_{j \in K_i} \sigma_{ij} [S(\mathbf{U}_{ij}^n, \mathbf{U}_{ij}^{*,n}, \mathbf{n}_{ij}) + S^c(\mathbf{U}_i^n, \mathbf{U}_{ij}^n, Z_i, Z_{ij}, \mathbf{n}_{ij})] \end{aligned} \tag{6.11}$$

where

$$S(\mathbf{U}_{ij}, \mathbf{U}_{ij}^*, \mathbf{n}_{ij}) = \begin{pmatrix} 0 \\ \dfrac{g}{2}(h_{ij}^{*2} - h_{ij}^2)\mathbf{n}_{ij} \end{pmatrix} \tag{6.12}$$

and

$$S^c(\mathbf{U}_i^n, \mathbf{U}_{ij}^n, Z_i, Z_{ij}, \mathbf{n}_{ij}) = \begin{pmatrix} 0 \\ -\dfrac{g}{2}(h_{ij} + h_i)(Z_{ij} - Z_i)\mathbf{n}_{ij} \end{pmatrix}. \tag{6.13}$$

Remark 6.1 *By contrast to the first order scheme, we have to introduce here a centered term (6.13) to satisfy the consistency with the source terms.*

Theorem 6.1 *The formally second order scheme defined by (6.11)–(6.13) with (4.7)- (4.9) satisfies the following properties*

(i) it preserves the water depth positivity under the CFL condition

$$\begin{aligned} \Delta t \leq \quad \min \Bigg[&\min_{i \in \mathbf{S_i}} \min_{j \in K_i} \frac{|C_{ij}|}{L_{ij}\ (|\mathbf{u}_{ij}^n| + w_M \tilde{c}_{ij}^{*,n})}, \\ &\min_{i \in \mathbf{G_i}} \max_{0 \leq \alpha \leq 1} \left(\alpha \frac{|C_i|}{L_i\ (|\mathbf{u}_i^n| + w_M \tilde{c}_i^n)},\ (1-\alpha) \min_{j \in K_i} \frac{|C_{ij}|}{L_{ij}\ (|\mathbf{u}_{ij}^n| + w_M \tilde{c}_{ij}^{*,n})} \right) \Bigg], \end{aligned} \tag{6.14}$$

(ii) it preserves the steady state of a lake at rest.

Proof. We follow the idea developed in [7] for the 1D problem. First we assume that P_i is an interior node. Using (4.7) and the definition (4.4) of σ_{ij}, the scheme (6.11) defines h_i^{n+1} by

$$h_i^{n+1} = \frac{1}{|C_i|}[|C_i|h_i^n - \frac{1}{\Delta t}\sum_{j\in K_i} L_{ij}(\mathrm{F}_h^+(\mathbf{U}_{ij}^{*,n}, \mathbf{n}_{ij}) + \mathrm{F}_h^-(\mathbf{U}_{ji}^{*,n}, \mathbf{n}_{ij}))]. \tag{6.15}$$

Using (6.5), we have

$$h_i^{n+1} = \frac{1}{|C_i|}\sum_{j\in K_i}[\ |C_{ij}|h_{ij}^n - \frac{L_{ij}}{\Delta t}(\mathrm{F}_h^+(\mathbf{U}_{ij}^{*,n}, \mathbf{n}_{ij}) + \mathrm{F}_h^-(\mathbf{U}_{ji}^{*,n}, \mathbf{n}_{ij}))]. \tag{6.16}$$

To verify the positivity of h_i^{n+1}, it is sufficient to have

$$|C_{ij}|h_{ij}^n - \frac{L_{ij}}{\Delta t}\mathrm{F}_h^+(\mathbf{U}_{ij}^{*,n}, \mathbf{n}_{ij}) \geq 0 \quad \text{for j} \in \mathrm{K_i}, \tag{6.17}$$

now using the relation $h_{ij}^* \leq h_{ij}$ and following the proof of theorem (4.1), we obtain that the inequality (6.17) is satisfied under the condition (6.14).
If P_i is a boundary node, we have

$$\begin{aligned} h_i^{n+1} = \frac{1}{|C_i|}[|C_i|h_i^n \quad &- \quad \frac{1}{\Delta t}\sum_{j\in K_i} L_{ij}(\mathrm{F}_h^+(\mathbf{U}_{ij}^{*,n}, \mathbf{n}_{ij}) + \mathrm{F}_h^-(\mathbf{U}_{ji}^{*,n}, \mathbf{n}_{ij})) \\ &- \quad \frac{L_i}{\Delta t}(\mathrm{F}_h^+(\mathbf{U}_i^n, \mathbf{n}_i) + \mathrm{F}_h^-(\mathbf{U}_{e,i}^n, \mathbf{n}_i))]. \end{aligned} \tag{6.18}$$

Then using (6.5) we can write

$$|C_i|h_i = \alpha|C_i|h_i + (1-\alpha)\sum_{j\in K_i}|C_{ij}|h_{ij}), \quad ,0 \leq \alpha \leq 1, \tag{6.19}$$

and with the same arguments as previously, we obtain the positivity of the waterdepth under the condition (6.14). □

Remark 6.2 *To obtain the second order accuracy it is necessary to consider more sophisticated reconstruction technics as the second order ENO reconstruction (see [3] for a rigorous proof for the one-dimensional Saint-Venant system and [28] for error estimates in the scalar case). They are not used here because they are very tricky to apply on 2d unstructured meshes, nevertheless the numerical results presented in the next section show an obvious accuracy improvement obtained with the formally second order scheme developed in this paper (see [1] for more details about ENO method on unstructured meshes). Second order accuracy in time can be obtained as usual by a Runge-Kutta method and the CFL condition need not be modified.*

7 NUMERICAL RESULTS

We present here the numerical results of different test problems. We begin with the two-dimensional version of a classical academic test problem extracted from [22] and commonly used (see, e.g. [12, 16]): a stationary flow over a parabolic bump for which an exact solution can be computed. We condider a rectangular channel with length 20. and width 2. (we assume an adimensionalized problem), the bottom is defined by

$$Z(x,y) = \begin{cases} 0.2 - 0.05(x-10.)^2 & \text{if } 8. \leq x \leq 12., \quad \forall y, \\ 0. \quad \text{else.} \end{cases}$$

Depending on the values of the boundary conditions, we compute three different flows defined as follows:

- fluvial flow
  inflow: $\mathbf{q}_g = (4.42, 0)^T$, outflow $h_g = 2.$,
- transcritical without shock (torrential outflow)
  inflow: $\mathbf{q}_g = (1.53, 0)^T$, initial water depth $h^0 = 0.66$,
- transcritical with shock
  inflow: $\mathbf{q}_g = (0.18, 0)^T$, outflow $h_g = 0.33$

The given discharge is prescribed for each node of the inflow boundary. The initial solution is given by $\mathbf{q}^0 = \mathbf{q}_g, h^0 = h_g$. The three flows are computed on a rather coarse unstructured mesh of 510 nodes and 886 triangles (60 edges on the length and 6 edges on the width). In Figures 7.1-(a), 7.1-(c) and 7.1-(e), the free surface profiles computed with the first order and second order schemes are compared to the exact solution. We have plotted only the points on the line $y = 0$ but we claim that the two-dimensional effects are negligible. Results are quite good for such a coarse mesh and the improvement due to the second order extension appears to be noticeable for all the cases, even the ones involving a discontinuity. Note also that the presence of a sonic point in the two last test cases does not need a special treatment.

Even if we have no error estimations for our problem, it appears interesting to look at the convergence rate of the error versus the space discretization for the three above problems. We have plotted in Figures 7.1-(b), 7.1-(d) and 7.1-(f), the Log (L^1-error) of the water depth versus Log (h_{a_0}/h_a) for the first and the second order scheme and they are compared to the theoretical order (we denote by h_a the average edge length and h_{a_0} the average edge length of the coarser mesh). These errors are computed on five meshes with respectively 10, 20, 30, 40 and 60 edges on the length of the channel. These meshes are very coarse, nevertheless, it appears that the computed convergence rate are not far from the theoretical ones, the formally second order scheme provides an effective convergence up to the second order when the flow is sufficiently smooth and according to the estimations, the second order scheme reduces to first order in the presence of a discontinuity.

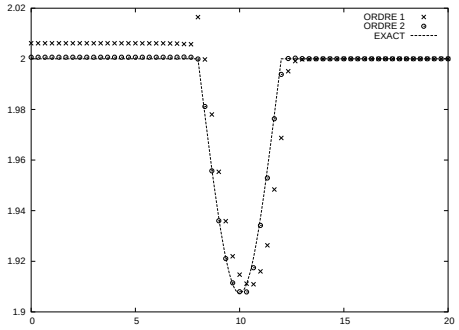
The second test problem is one of the tests of the Telemac code developed at EDF/LNHE [25], it concerns a water drop in a basin and we look at the solution after some reflections

on the walls. The basin is a 20.x20. square box with flat bottom, the initial solution shown in Fig. 7.2 (a), is defined by

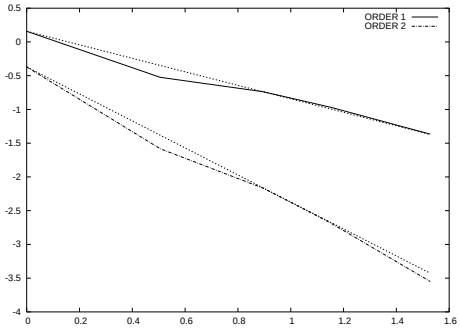
$$h = 2.4(1. + e^{-0.25[(x-10.05)^2+(y-10.05)^2]}), \quad \mathbf{u} = 0.$$

The solutions at $t = 1., 2., 3., 4.$s obtained with the second order approximation are given in Fig. 7.2 (b)-(e) while the solution at $t = 4.$s, damped by the first order scheme is shown in Fig. 7.2 (f). This result also shows the accuracy improvement due to the second order scheme.

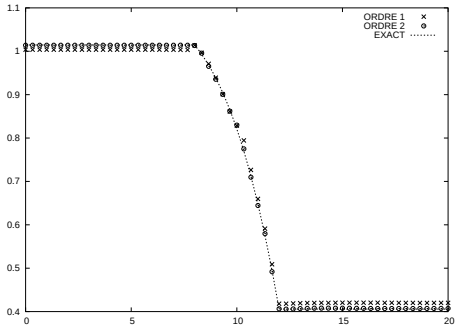
The third test problem is a real life application, it concerns the Malpasset dam break. All the details on the data and a reference solution computed with the Telemac code are given in [26]. We present here in Fig. 7.3, the initial solution and the second order solutions at $t = 1000$ s and $t = 2500$ s. These solutions are in good agreement with solutions obtained by other methods in [26]. The computation of this problem allows to test, among others, the ability of the method to treat the still water (the sea area before the wave reaches it) and the wet-dry interfaces.



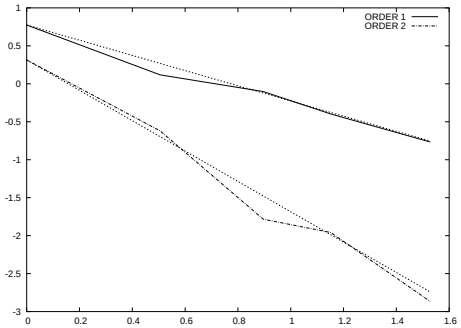
(a) Fluvial flow - Free Surface



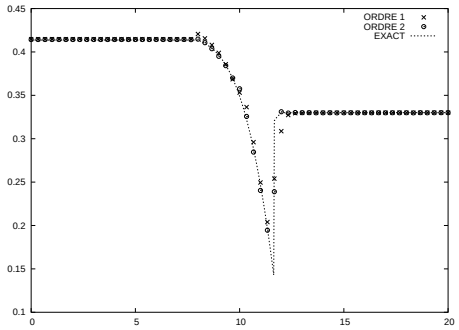
(b) Fluvial flow - Convergence rate



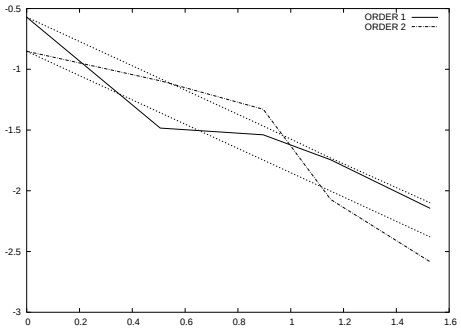
(c) Transcritical flow - Free Surface



(d) Transcritical flow - Convergence rate

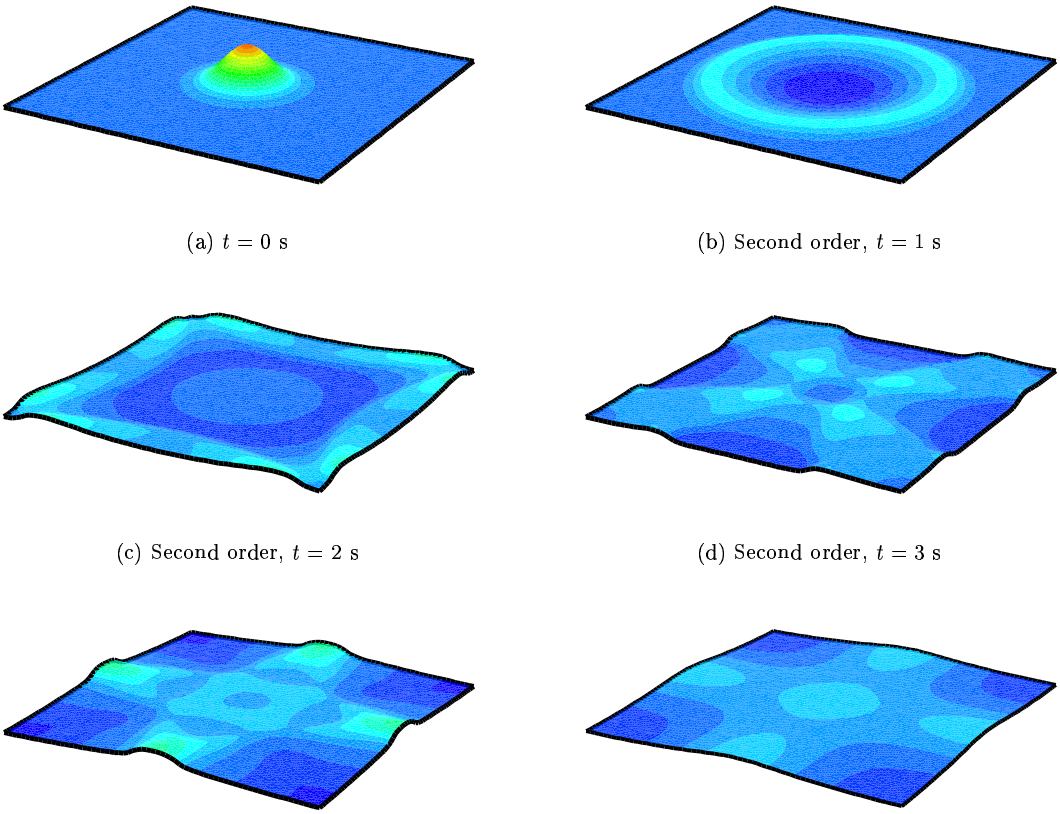


(e) Transcritical flow with shock - Free Surface



(f) Transcritical flow with shock - Convergence rate

Figure 7.1: Stationary flows over a bump.



(a) $t = 0$ s

(b) Second order, $t = 1$ s

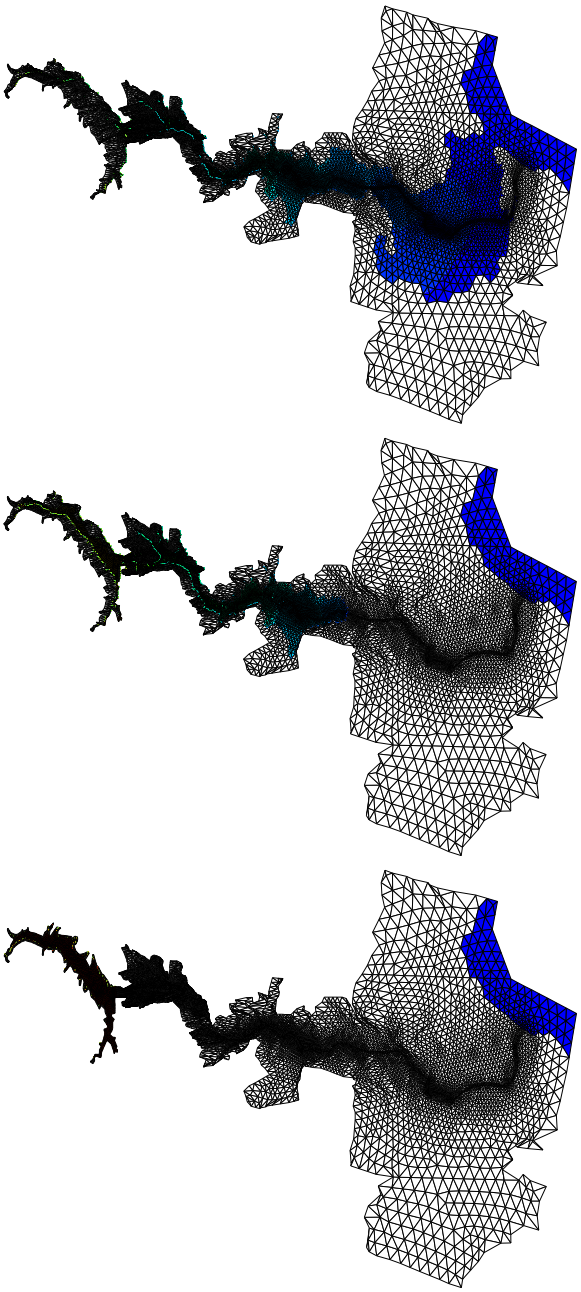
(c) Second order, $t = 2$ s

(d) Second order, $t = 3$ s

(e) Second order, $t = 4$ s

(f) First order, $t = 4$ s

Figure 7.2: Water drop in a basin.



(a) Initial solution (b) $t = 1000$s (c) $t = 2500$s

Figure 7.3: Malpasset dam break

8 Conclusion and outlook

In this article we have introduced on one hand a stable homogeneous kinetic solver and on the other hand a hydrostatic reconstruction method to compute the source term while preserving the stability properties of the homogeneous solver. We have also presented a second order compatible extension. Thanks to these three ingredients we finally derived a positivity preserving well-balanced second order scheme. According to the simplicity of the presented algorithms, we emphasize that this solution method seems to be a good compromise between efficiency, stability and accuracy. These properties are experimentally verified by using the algorithm to reproduce complex physical phenomena.

Moreover let us notice some extensions that can be derived. The stability properties of the kinetic solver can be used to derive stable schemes for avalanche flows [35] and can be extended to a multilayer Saint-Venant model [2]. An extension of the hydrostatic reconstruction that preserves all the subsonic steady states is under investigation and the same idea can be adapted to take into account the relation between the Darcy equation and the Saint-Venant system with strong friction coefficient.

Acknowledgements

The authors thank F. Bouchut, J.M. Hervouet, R. Klein and B. Perthame for fruitful discussions and helpful comments. This work was partially supported by EDF/LNHE and by HYKE European programme HPRN-CT-2002-00282 (http://www.hyke.org).

References

[1] Abgrall R., On essentially non-oscillatory schemes on unstructured meshes: analysis and implementation, *J. Comput. Phys.*, 114 (1994), no. 1, 45–58.

[2] Audusse E., A multilayer Saint-Venant model, *Discrete Cont. Dyn. Syst. Ser. B,* to appear.

[3] Audusse E., Bouchut F., Bristeau M.O., Klein R. and Perthame B., A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows, *SIAM J. Sc. Comp.*, to appear.

[4] Audusse E., Bristeau M. O., Transport of pollutant in shallow water, a two time steps kinetic method, *M2AN*, **37**, no.2, 389-416, 2003.

[5] Bermudez A., Dervieux A., Desideri J.A., Vazquez M.E., Upwind schemes for the two-dimensional shallow water equations with variable depth using unstructured meshes, *Comput. Methods Appl. Mech. Engrg.*, **155**, no. 1-2, 49–72, 1998.

[6] Bermudez A., Vazquez M.E., Upwind methods for hyperbolic conservation laws with source terms, *Comput. Fluids*, **23**, no. 8, 1049–1071, 1994.

[7] Bouchut F., *Nonlinear stability of finite volume methods for hyperbolic conservation laws, and well-balanced schemes for sources*, Frontiers in Mathematics, Birkhauser, 2004.

[8] Bouchut F., Mangeney-Castelnau A., Perthame B., Vilotte J.P., A new model of Saint-Venant and Savage-Hutter type for gravity driven shallow water flows, *C. R. Math. Acad. Sci.*, Paris, **336**, no. 6, 531-536, 2003.

[9] Bristeau M.O. and Coussin, B., Boundary Conditions for the Shallow Water Equations solved by Kinetic Schemes, *INRIA Report*, **4282**, 2001,http://www.inria.fr/RRRT/RR-4282.html

[10] Bristeau M. O. and Perthame B., Transport of Pollutant in Shallow Water using Kinetic schemes, *ESAIM Proceedings*, 10, CEMRACS 1999, 9-21, http://www.emath.fr/Maths/Proc/Vol.10.

[11] Botchorishvili R., Perthame B. and Vasseur A., Equilibrium Schemes for Scalar Conservation Laws with Stiff Sources, *Inria report*, **3891**, 2000.

[12] Chacon Rebollo T., Delgado A.D., Nieto E.D.F., An entropy-correction free solver for non homogeneous shallow water equations, *M2AN*, **37**, 363-390, 2003.

[13] Chacon Rebollo T., Delgado A.D., Nieto E.D.F., A family of stable numerical solvers for the shallow water equations with source terms, *Comp. Meth. Appl. Math. Engin.* 192 (2003) 203-225.

[14] Dafermos C.M., *Hyperbolic conservation laws in continuum physics*, Springer Verlag, GM 325, 1999.

[15] Ferrari S., Saleri F., A new two-dimensional Shallow Water model including pressure effects and slow varying bottom topography, *M2AN*, to appear.

[16] Gallouët T., Hérard J.M. and Seguin N., Some approximate Godunov schemes to compute shallow-water equations with topography, *Comput. and Fluids*, **32**, 479-513, 2003.

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

[18] Gosse L., A well-balanced scheme using nonconservative products designed for hyperbolic systems of conservation laws with source terms, *Math. Mod. Meth. Appl. Sci.*, **11**, no. 2, 339–365, 2001.

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

[20] Gerbeau J.-F. and Perthame B., Derivation of Viscous Saint-Venant System for Laminar Shallow Water; Numerical Validation, *Discrete Cont. Dyn. Syst. Ser. B,* **1**, No 1, 89-102, 2001.

[21] Godlewski E. and Raviart P.-A., *Numerical approximations of hyperbolic systems of conservation laws*, Applied Mathematical Sciences 118, Springer-Verlag, New York, 1996.

[22] Goutal N. and Maurel F. Proceedings of the 2nd workshop on dam-break simulation, *Note technique EDF*, HE-43/97/016/B, 1997.

[23] Gray J.M.N.T., Tai Y.C., Noelle S., Shock waves, dead zones and particle-free regions in rapid granular free-surface flows. *J. Fluid Mech.* 491 (2003), 161–181.

[24] Guillard H. and Abgrall R., *Modélisation numérique des fluides compressibles*, Series in Applied Mathematics, Gauthier-Villars, 2001.

[25] Hervouet J.M., *Hydrodynamique des écoulements à surface libre. Modélisation numérique avec la méthode des éléments finis* (in french), Presses des Ponts et Chaussées, 2003.

[26] Hervouet J.M., A high resolution 2-D dam-break model using parallelization, *Hydrol. Process.*, **14**, 2211-2230, 2000.

[27] Jin S., A steady state capturing method for hyperbolic systems with geometrical source terms, *M2AN*, **35**, 631-645, 2001.

[28] Katsaounis T. and Simeoni C., First and second order error estimates for the Upwind Interface Source method, *Math. Comp.*, to appear.

[29] Katsaounis T. and Makridakis C., Relaxation Models and Finite Element Schemes for the Shallow Water Equations, *Hyperbolic Problems : Theory, Numerics, Applications*, Springer, 621-631, 2003.

[30] Kurganov A., Levy D., Central-upwind schemes for the Saint-Venant system, *M2AN*, **36**, 397-425, 2002.

[31] LeVeque R.J., *Numerical Methods for Conservation Laws*, Lectures in Mathematics, ETH Zurich, Birkhauser, 1992.

[32] LeVeque R.J., Balancing source terms and flux gradients in high-resolution Godunov methods : the quasi-steady wave-propagation algorithm, *J. Comput. Phys.*, **146**, no. 1, 346–365, 1998.

[33] Lions P.L., Perthame B. and Souganidis P.E., Existence of entropy solutions for the hyperbolic systems of isentropic gas dynamics in Eulerian and Lagrangian coordinates, *Comm. Pure Appl. Math.*, **49**, no. 6, 599–638, 1996.

[34] Monthe L.A., Benkhaldoun F., Elmahi I., Positivity preserving finite volume Roe schemes for transport-diffusion equations, *Comput. Methods Appl. Mech. Engrg.*, **178**, 215-232, 1999.

[35] Mangeney-Castelnau A., Vilotte J.P., Bristeau M.O., Bouchut F., Perthame B., Simeoni C., Yernini S., A new kinetic scheme for Saint-Venant equations applied to debris avalanches, *INRIA Report* **4646**, 2002, http://www.inria.fr/RRRT/RR-4646.html

[36] Perthame B., *Kinetic formulations of conservation laws*, Oxford University Press, 2002.

[37] Perthame B. and Qiu Y., A variant of Van Leer's method for multidimensional systems of conservation laws, *J. Comp. Phys.*, **112(2)**, 370-381, 1994.

[38] Perthame, B., Simeoni, C., A kinetic scheme for the Saint-Venant system with a source term, *Calcolo*, **38(4)**, 201-231, 2001.

[39] Roe P. L., Approximate Riemann solvers, parameter vectors and difference schemes, *J. Comp. Phys.*, **43**, 357–372, 1981.

[40] Russo G., Central schemes for balance laws, *Hyperbolic problems : theory, numerics, applications*, Vol. I, II (Magdebourg, 2000), Internat. Ser. Numer. Math., 140, 141, 821-829, Birkhauser, Basel, 2001.

[41] de Saint-Venant A.J.C., Théorie du mouvement non-permanent des eaux, avec application aux crues des rivières et à l'introduction des marées dans leur lit (in french), *C. R. Acad. Sc.*, Paris, **73**, 147-154, 1871.

[42] Serre D., *Systèmes hyperboliques de lois de conservation* , Diderot, Paris, 1996.

[43] Van Leer B., Towards the Ultimate Conservative Difference Schemes. V. A Second Order Sequel to the Godunov's Method, *J. Comp. Phys.*, **32**, 1979.

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399