Some approximate Godunov schemes to compute shallow-water equations with topography

Thierry Gallouët [a], Jean-Marc Hérard [a,b,*], Nicolas Seguin [a,b]

[a] *Laboratoire d'Analyse Topologie et Probabilités, UMR CNRS 6632, Centre de Mathématique et Informatique, Université de Provence, 39 rue Joliot Curie, 13453 Marseille Cedex 13, France*
[b] *Département Mécanique des Fluides et Transferts Thermiques, Électricité de France, Recherche et Développement, 6 quai Watier, 78401 Chatou Cedex, France*

Abstract

We study here the computation of shallow-water equations with topography by Finite Volume methods, in a one-dimensional framework (though all methods introduced may be naturally extended in two dimensions). All methods are based on a discretisation of the topography by a piecewise function constant on each cell of the mesh, from an original idea of Le Roux et al. Whereas the Well-Balanced scheme of Le Roux is based on the exact resolution of each Riemann problem, we consider here approximate Riemann solvers. Several single step methods are derived from this formalism, and numerical results are compared to a fractional step method. Some test cases are presented: convergence towards steady states in subcritical and supercritical configurations, occurrence of dry area by a drain over a bump and occurrence of vacuum by a double rarefaction wave over a step. Numerical schemes, combined with an appropriate high-order extension, provide accurate and convergent approximations.
© 2003 Elsevier Science Ltd. All rights reserved.

1. Introduction

We study in this paper some approximate Godunov schemes to compute shallow-water equations with a source term of topography, in a one-dimensional framework. All methods presented may be extended naturally to the two-dimensional model.

---

* Corresponding author. Tel.: +33-1-30-87-70-37; fax: +33-1-30-87-79-49.
E-mail addresses: gallouet@cmi.univ-mrs.fr (T. Gallouët), herard@cmi.univ-mrs.fr, herard@chi80bk.der.edf.fr (J.-M. Hérard), seguin@chi80bk.der.edf.fr (N. Seguin).

Shallow-water equations are based on conservation laws and provide a hyperbolic system. However, topography introduces some source term related to the unknown. Hence, analytic properties of the model of isentropic Euler equations are deeply modified, in comparison with the homogeneous case. For instance, a well-known problem is the occurrence of other equilibrium states (or steady states), due to the presence of the source term.

Several ways to compute conservation laws with source term have already been investigated. The main problem is the approximation of the source term and the numerical preservation of properties fulfiled by the continuous system. Some Finite Volume methods have been proposed, in particular the Well-Balanced schemes, which can maintain *all* steady states. These schemes have been initially introduced by Greenberg and co-workers [17,18] in the scalar case (see also [2,15]). Well-Balanced schemes have been recently extended to shallow-water equations with topography in [1,19] and friction in [6]. Since the Well-Balanced scheme is based on an exact Riemann solver as the Godunov scheme (see [14]), its main drawbacks are its calculation cost and the need to compute the "exact" solution of the Riemann problem. Other Finite Volume methods to deal with source terms exist too, for instance based on the Roe scheme (see [12,24]), or based on another approximation of the source term, like in [21].

Some properties of the continuous model (Riemann invariants, jump relations, etc.) are first exposed, and a study of the Riemann problem is briefly recalled. Thereafter, some approximate Gudunov methods are introduced to compute shallow-water equations, derived from the VFRoe-ncv formalism. VFRoe-ncv stands for the scheme VFRoe using a non-conservative variable. Finite Volume scheme VFRoe should not be confused with Roe scheme. VFRoe-ncv schemes are Finite Volume schemes, based on a linearised Riemann problem written with respect to a non-conservative variable. Some applications of VFRoe-ncv schemes are provided for the Euler equations (in [5,10,23]), for shallow-water equations with a flat bottom in [4] and for turbulent compressible flows [3]. The VFRoe-ncv schemes are based on an arbitrary change of variable, and on a linearisation of each interface Riemann problem. In the homogeneous case, the numerical flux is defined using the exact solution of the linearised Riemann problem and the conservative flux. However, the source term "breaks" the conservativity of the model. Thus, using a piecewise constant function to approximate the bottom, some approximate Riemann solvers are presented. The main advantages of this approach are the natural integration of the source term in the numerical methods and the use of a linearised Riemann problem, which minimizes the CPU time. Note that a scheme which exactly preserves a large class of steady states is obtained. In addition, a fractional step method (FSM) is performed, based on the VFRoe-ncv scheme introduced in [4]. This method enables to deal with vacuum and provides good results, too. To complete this presentation, a higher-order extension is provided, to increase the accuracy of the schemes when computing unsteady configurations or flows at rest.

Several numerical experiments are presented. All the test cases are one-dimensional, and are based on a non-trivial topography. Indeed, applications of shallow-water equations are one- or two-dimensional configurations. Hence, computational limitations are rather different from the gas dynamics frame where the main applications are two- and three-dimensional. Therefore, numerical experiments in an industrial context may be here performed on mesh containing several hundreds nodes. The tests include subcritical and transcritical flows over a bump [16] and a drain with a non-flat bottom. The convergence towards steady states is measured. A vacuum occurrence by a double rarefaction wave over a step is tested too. All the numerical

tests confirm the good behaviour of the numerical methods, including the fractional step method.

Eventually, some complementary tests with the Godunov and the VFRoe method are provided in Appendix A.

2. The shallow-water equations with topography

2.1. Governing equations

The shallow-water equations represent a free surface flow of incompressible water. The two-dimensional system may be written as follows:

$$h_{,t} + (hu)_{,x} + (hv)_{,y} = 0, \tag{1a}$$

$$(hu)_{,t} + (hu^2)_{,x} + (huv)_{,y} + g\left(\frac{h^2}{2}\right)_{,x} = -gh(Z_f)_{,x}, \tag{1b}$$

$$(hv)_{,t} + (huv)_{,x} + (hv^2)_{,y} + g\left(\frac{h^2}{2}\right)_{,y} = -gh(Z_f)_{,y}, \tag{1c}$$

where h denotes the water height, $\mathbf{u} = {}^t(u, v)$ the velocity, g the gravity constant and ∇Z_f the bed slope (g and $Z_f(x, y)$ are given, and Z_f must be at least $\mathscr{C}^0(\mathbb{R}^2)$) (see Fig. 1).

This study is restricted to the computation by Finite Volume schemes (see [9]). Since the hyperbolic system (1a)–(1c) remains unchanged under frame rotation, this two-dimensional problem may be solved considering on each interface of the mesh the following system:

$$h_{,t} + (hu_n)_{,n} = 0, \tag{2a}$$

$$(hu_n)_{,t} + \left(hu_n^2 + g\frac{h^2}{2}\right)_{,n} = -gh(Z_f)_{,n}, \tag{2b}$$

$$(hu_\tau)_{,t} + (hu_n u_\tau)_{,n} = 0, \tag{2c}$$

where $u_n = \mathbf{u} \cdot n$, $u_\tau = \mathbf{u} \cdot \tau$, n and τ the normal and the tangential vector to the interface ($\|n\| = \|\tau\| = 1$), and $( )_{,n}$ the derivate along the normal vector n.
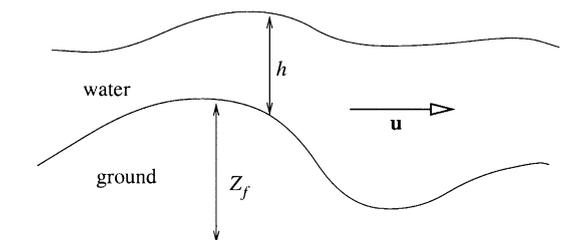

Fig. 1. Mean variables.

The pure one-dimensional shallow-water equations may be written as follows:

$$h_{,t} + (hu)_{,x} = 0, \tag{3a}$$

$$(hu)_{,t} + \left(hu^2 + g\frac{h^2}{2} \right)_{,x} + ghZ'_f(x) = 0. \tag{3b}$$

We focus in this paper on the numerical resolution of the one-dimensional systems (3a) and (3b).

Let us note that h and hu (also denoted Q in the following) are the conservative variables. So, vacuum (or dry bed) may be represented by $h = hu = 0$, which implies that u is not defined.

Remark 1. The change of variable from (h, Q) to (h, u) leads to the following equations for smooth solutions:

$$h_{,t} + Q_{,x} = 0,$$
$$u_{,t} + \psi_{,x} = 0,$$

where $\psi = (u^2/2 + g(h + Z_f))$.

These equations enable to define some stationary smooth solutions as follows:

$$Q_{,x} = 0 \quad \text{and} \quad \psi_{,x} = 0. \tag{4}$$

One may add to these equations Rankine Hugoniot relations (on smooth topography) for stationary shocks to complete the definition of stationary states.

2.2. The Riemann problem on a flat bottom

Assuming that the river bed is flat (i.e. $Z'_f(x) = 0$), the system (3a) and (3b) becomes homogeneous. Hence, we obtain a conservative system, which leads to the following Riemann problem:

$$\begin{cases} h_{,t} + Q_{,x} = 0, \\ Q_{,t} + \left(\dfrac{Q^2}{h} + g\dfrac{h^2}{2} \right)_{,x} = 0, \\ (h, Q)(x, 0) = \begin{cases} (h_L, Q_L) & \text{if } x < 0, \\ (h_R, Q_R) & \text{if } x > 0. \end{cases} \end{cases} \tag{5}$$

This problem, which is also the Riemann problem for isentropic Euler equations (for a particular state law) may be classically solved. Its solution is a similarity solution (i.e. a function of x/t) composed by three constant states, (h_L, Q_L), (h_1, Q_1) and (h_R, Q_R) separated by two genuinely non-linear (GNL) fields associated with eigenvalues $u - c$ and $u + c$ (where $c = \sqrt{gh}$). The intermediate state (h_1, Q_1) may be computed using through the 1-wave:

$$u_1 = \begin{cases} u_L - 2(\sqrt{gh_1} - \sqrt{gh_L}) & \text{if } h_1 < h_L, \\ u_L - (h_1 - h_L)\sqrt{g\dfrac{h_1 + h_L}{2h_1 h_L}} & \text{if } h_1 > h_L, \end{cases} \tag{6}$$

and through the 2-wave:

$$u_1 = \begin{cases} u_R + 2(\sqrt{gh_1} - \sqrt{gh_R}) & \text{if } h_1 < h_R, \\ u_R + (h_1 - h_R)\sqrt{g\dfrac{h_1 + h_R}{2h_1 h_R}} & \text{if } h_1 > h_R. \end{cases} \tag{7}$$

The latter two curves are derived from the Riemann invariants (when $h_1 < h_L$ and $h_1 < h_R$) for rarefaction waves and from the Rankine Hugoniot relations (when $h_1 > h_L$ and $h_1 > h_R$) for shock waves. Note that the intermediate velocity u_1 is defined only if

$$u_R - u_L < 2(\sqrt{gh_R} + \sqrt{gh_L}). \tag{8}$$

Otherwise, h_1 and Q_1 become null, and u_1 is undefined.

2.3. The Riemann problem with a piecewise constant topography

Following the idea developed by Le Roux [19], the topography is described by a piecewise constant function. Therefore, adding the "partial" differential equation concerning Z_f, the following Riemann problem may be obtained:

$$\begin{cases} Z_{f,t} = 0, \\ h_{,t} + (hu)_{,x} = 0, \\ Q_{,t} + \left(\dfrac{Q^2}{h} + g\dfrac{h^2}{2}\right)_{,x} + gh(Z_f)_{,x} = 0, \\ (h, Q, Z_f)(x, 0) = \begin{cases} (h_L, Q_L, Z_{f_L}) & \text{if } x < 0, \\ (h_R, Q_R, Z_{f_R}) & \text{if } x > 0. \end{cases} \end{cases} \tag{9}$$

Note that this Riemann problem does not correspond to the Riemann problem associated with the system (3a) and (3b), since the topography is not smooth. The jump of topography along the curve $x/t = 0$ introduces a problem for the definition of the product of distributions, focusing on non-smooth solutions (see [7,8] for more details). So, the jump relations across the discontinuity $x/t = 0$ are not defined. Assuming that $h > 0$ and restricting to smooth solutions, the system (9) may be written:

$$Z_{f,t} = 0, \tag{10a}$$

$$h_{,t} + Q_{,x} = 0, \tag{10b}$$

$$u_{,t} + \left(\dfrac{u^2}{2} + g(h + Z_f)\right)_{,x} = 0. \tag{10c}$$

We note $\psi = (u^2/2 + g(h + Z_f))$ in the following. One may deduce the conservation law on entropy for non-viscous smooth solutions:

$$\eta_{,t} + (Q\psi)_{,x} = 0, \tag{11}$$

$$\eta = h\frac{u^2}{2} + g\frac{h^2}{2} + ghZ_f. \tag{12}$$

Moreover, system (10a)–(10c) provides the Riemann invariants through the stationary wave. Since the wave located at $x/t = 0$ is a contact discontinuity, we assume that the Rankine Hugoniot

relations identify with the Riemann invariants. Thus, the Riemann problem (9) admits a Linearly Degenerated field of speed 0 such that

$$[\![Q]\!] = 0, \tag{13a}$$

$$[\![\psi]\!] = 0, \tag{13b}$$

where $[\![\alpha]\!]$ represents the jump of α across the wave.

Two GNL fields also compose the solution of the Riemann problem (9), which are the same as in the flat bottom case. Hence, to connect a state W to a state W_a through the wave $u - c$, one may use the following relations (a rarefaction wave occurs when $h < h_a$, and a shock wave occurs when $h > h_a$):

$$Z_f = Z_{f_a}, \tag{14a}$$

$$u = \begin{cases} u_a - 2(\sqrt{gh} - \sqrt{gh_a}) & \text{if } h < h_a, \\ u_a - (h - h_a)\sqrt{g\dfrac{h + h_a}{2hh_a}} & \text{if } h > h_a. \end{cases} \tag{14b}$$

In the same way, to connect a state W to a state W_b through the wave $u + c$, one may use the following relations (a rarefaction wave occurs when $h < h_b$, and a shock wave occurs when $h > h_b$):

$$Z_f = Z_{f_b}, \tag{15a}$$

$$u = \begin{cases} u_b + 2(\sqrt{gh} - \sqrt{gh_b}) & \text{if } h < h_b, \\ u_b + (h - h_b)\sqrt{g\dfrac{h + h_b}{2hh_b}} & \text{if } h > h_b. \end{cases} \tag{15b}$$

Moreover, to connect a state W to a state W_c through the stationary wave, one uses the Riemann invariants:

$$Q = Q_c, \tag{16a}$$

$$\psi = \psi_c. \tag{16b}$$

Remark 2. Several remarks about the resolution of the Riemann problem (9) follow.

1. Since Eq. (3b) is not conservative when the bottom Z_f is discontinuous, the associated jump relation (13b) is just the sense we give to the third equation of (9) at the discontinuity of Z_f. Since this discontinuity is a contact discontinuity for (9), this choice is quite natural because it corresponds to the Riemann invariants in the associated field.
2. Even assuming (13b), the resolution of the Riemann problem (9) remains undetermined. Indeed, combining Eqs. (16a) and (16b), the connection through the stationary wave results in looking for solution (h, Q) of the couple of equations $Q = Q_c$ and $Q_c^2/(2h^2) + g(h + Z_f) = \psi_c$, where Z_f, Q_c and ψ_c are given. This equation may admit zero, one or two solutions. Therefore, without additionnal informations, the solution remains unknown.

3. Let us emphasize that the Riemann problem (9) may be resonant, that is, a GNL wave may be superposed with the stationary wave. Moreover, waves are not ordered which renders the resolution of the Riemann problem much more complex than in the flat bottom case.

   At this stage, it clearly appears that the classical method to solve the Riemann problem (described in [13] for instance) is not sufficient to construct the solution of (9), owing to items 1–3.

4. Le Roux and co-workers [6,19,25] have proposed an original method to solve the Riemann problem. It is based on a linear connection between Z_{f_L} and Z_{f_R} and the study of stationary solutions of (3a), (3b) inside the connection. Some relations naturally appear and allow the complete resolution of the Riemann problem (9). In particular, except when a stationary shock wave occurs, states on each side of the discontinuity $x/t = 0$ must be either both subcritical or both supercritical. Therefore, no ambiguity remains when the parametrisation (16a) and (16b) admits two solutions because one of the two solutions is subcritical while the other is supercritical.

5. Note that, in a simpler framework (see [26]), a similar method may provide existence and uniqueness of the entropy solution.

We will discuss below two families of schemes which are intended to provide a convergent approximation of the above-mentioned system. The first series is based on straightforward approximate Godunov schemes which account for topography. The second series is based on the fractional step method.

3. Single step methods

We present in this section several ways to solve the shallow-water equations with source term by Finite Volume schemes (see [9,27] for instance). The description of the methods computed herein is split in two steps: the Finite Volume scheme provided by integration of (3a) and (3b) and the solver at each interface.

3.1. An approximate Godunov-type scheme

We introduce herein a Finite Volume scheme following the idea proposed by Greenberg and co-workers [17,18].

Focusing on system (3a) and (3b), it consists in using a piecewise bottom, flat on each cell, in the "continuous" framework (see [6,19]). Thus, the source term $-ghZ'_f(x)$ is reduced to a sum of Dirac masses occuring on each interface [7]. Hence, since the Finite Volume formalism is based on the integration of the system (3a) and (3b) on a cell $]x_{i-1/2}; x_{i+1/2}[\times[t^n; t^{n+1}[$, the source term does not appear explicitly (contrary to the scheme investigated in [12] for instance). As mentioned above, such an approximation of the topography introduces a stationary wave at the interface of each local Riemann problem. Though the Well-Balanced scheme of Greenberg and Le Roux is based on the exact solution of (9), we focus here on approximate Riemann solvers. These Riemann solvers are based on an approximate solution of the problem (9), and the numerical flux is computed from the conservative flux and the approximate solution at each interface.

Let us note $W = {}^t(Z_f, h, Q)$ the conservative variable, $F(W) = {}^t(0, Q, hu^2 + gh^2/2)$ the associated conservative flux and Δx_i and Δt the space and time-steps. We denote W_i^n the approximation of

$$\frac{1}{\Delta x_i} \int_{x_{i-1/2}}^{x_{i+1/2}} W(x, t^n) \, dx.$$

So, the Finite Volume scheme may be written as follows:

$$W_i^{n+1} = W_i^n - \frac{\Delta t}{\Delta x_i} \left(F(W_{i+1/2}^*(0^-; W_i, W_{i+1})) - F(W_{i-1/2}^*(0^+; W_{i-1}, W_i)) \right), \tag{17}$$

where $W_{i+1/2}^*(x/t; W_i, W_{i+i})$ is the (exact or approximate) solution of the Riemann problem (9) with $L = i$ and $R = i + 1$. As mentioned above, the source term only contributes to the computation of the (exact or approximate) solutions $W_{i+1/2}^*(x/t; W_i, W_{i+i})$ but it does not appear explicitly in the expression of the scheme (17). However, the approximation of the topography by a piecewise constant function implies that the numerical flux is not continuous through each interface of the mesh, contrary to the homogeneous and conservative case. Whereas the numerical flux associated with Eq. (3a) has to be continuous (since this equation is homogeneous and conservative), the numerical flux associated with Eq. (3b) becomes discontinuous in the non-flat bottom case, according to relations (13a) and (13b). In order to obtain a constant numerical flux for Eq. (3a), we will have, in some cases, to modify the scheme (17) (see (21) for instance).

Note that the Finite Volume scheme (17) associated with the exact interface Riemann solver (i.e. the Well-Balanced scheme presented in [19]) is able to maintain all steady states. Moreover, let us emphasize that the scheme (17) may be easily extended to a multi-dimensional framework (indeed, the formalism presented is very similar to Finite Volume schemes).

3.2. The VFRoe-ncv formalism

Since the Well-Balanced scheme [19] is based on an exact Riemann solver as the Godunov scheme [14], its main drawbacks are its calculation cost and the need to compute the exact solution of the Riemann problem (9). Thus, we suggest to compute the state $W_{i+1/2}^*(x/t; W_i, W_{i+1})$ by approximate Riemann solvers.

All the Riemann solvers presented here may be derived from the VFRoe-ncv formalism [5,10]. The VFRoe-ncv schemes are based on the exact solution of a linearised Riemann problem. Their construction may be split in three steps. The first step consists in writing the initial system under a non-conservative form, by an arbitrary change of variable $Y(W)$ (we denote by $W(Y)$ the inverse change of variable). Afterwards, the Riemann problem (9) is linearised averaging the convection matrix:

$$\begin{cases} Y_{,t} + B(\widehat{Y}) Y_{,x} = 0, \\ Y(x, 0) = \begin{cases} Y_L = Y(W_L) & \text{if } x < 0, \\ Y_R = Y(W_R) & \text{if } x > 0, \end{cases} \end{cases} \tag{18}$$

where $B(Y) = (W_{,Y}(Y))^{-1} F_{,W}(W(Y)) W_{,Y}(Y)$ and $\widehat{Y} = ((Y_L + Y_R)/2)$.

As a result, the Riemann problem (9) becomes a linear Riemann problem, which is solved exactly. Denoting $(\tilde{l}_k)_{k=1,2,3}$ and $(\tilde{r}_k)_{k=1,2,3}$, respectively, left and right eigenvectors of $B(\widehat{Y})$, $(\tilde{\lambda}_k)_{k=1,2,3}$ eigenvalues of $B(\widehat{Y})$, the exact solution $Y^*(x/t; Y_L, Y_R)$ of (18) is defined by:

$$
Y^*((x/t)^-; Y_L, Y_R) = Y_L + \sum_{x/t > \tilde{\lambda}_k} \left({}^t\tilde{l}_k \cdot [\![Y]\!]_L^R \right)\tilde{r}_k,
$$

$$
Y^*((x/t)^+; Y_L, Y_R) = Y_R - \sum_{x/t < \tilde{\lambda}_k} \left({}^t\tilde{l}_k \cdot [\![Y]\!]_L^R \right)\tilde{r}_k,
\tag{19}
$$

where $[\![\alpha]\!]_L^R = \alpha_R - \alpha_L$. Both are equal when $x/t \neq \tilde{\lambda}_k$ and $x/t = \tilde{\lambda}_k$ corresponds to a discontinuity of Y^*, $k = 1, 2, 3$. Thus, the solution written in terms of the conservative variable is

$$
W^*(x/t; W_L, W_R) = W(Y^*(x/t; Y_L, Y_R)).
\tag{20}
$$

Therefore numerical fluxes in (17) are computed using (20). In a conservative and homogeneous framework, the numerical flux is defined by the conservative flux computed with the approximate solution at the interface $x/t = 0$. However, the Riemann problem (9) provides a stationary wave at the interface, which introduces a jump of the numerical flux across it (which appears even when the exact solution of (9) is computed).

We emphasize that the source term of topography $-gZ'_f(x)$ appears naturally and explicitly in the expression of intermediate states computed by the following schemes.

3.3. The VFRoe (Z_f, h, Q) scheme

We consider first the conservative variable $W = {}^t(Z_f, h, Q)$. Note that this solver corresponds to the initial VFRoe scheme [23]. The main interest of this interface Riemann solver is the discrete continuity of Q through the stationary wave, in agreement with the Riemann invariant (13a).

If we develop the system (3a), (3b), we can write the convection matrix (which identifies with the jacobian matrix of the numerical flux $F_{,W}(W)$):

$$
B(Y) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ c^2 & c^2 - u^2 & 2u \end{pmatrix}.
$$

Eigenvalues of the matrix $B(Y)$ are

$$
\lambda_1 = 0, \quad \lambda_2 = u - c, \quad \lambda_3 = u + c.
$$

The associated matrix of right eigenvectors is

$$
\Omega = \begin{pmatrix} c^2 - u^2 & 0 & 0 \\ -c^2 & 1 & 1 \\ 0 & u - c & u + c \end{pmatrix}.
$$

If we refer to the exact solution (19) of the linearised Riemann problem (18), we can write

$$
W^*(0^+; W_L, W_R) = W^*(0^-; W_L, W_R) + \frac{[\![Z_f]\!]_L^R}{\tilde{c}^2 - \tilde{u}^2} \begin{pmatrix} \tilde{c}^2 - \tilde{u}^2 \\ -\tilde{c}^2 \\ 0 \end{pmatrix},
$$

where $\tilde{u} = u(\widehat{Y})$ and $\tilde{c} = c(\widehat{Y})$. This implies that the discharge Q is continuous through the stationary wave, according to relation (13a). So, the scheme associated to h is conservative. By the same way, one may write the relations to connect a state W to a state W_a through the $u - c$ wave:

$$W = W_a + \frac{1}{2} \left(\frac{\tilde{c}[\![Z_f]\!]_L^R}{\tilde{c} - \tilde{u}} + \frac{(\tilde{c} + \tilde{u})[\![h]\!]_L^R}{\tilde{c}} - \frac{[\![Q]\!]_L^R}{\tilde{c}} \right) \begin{pmatrix} 0 \\ 1 \\ \tilde{u} - \tilde{c} \end{pmatrix},$$

and the relations to connect a state W to a state W_b through the $u + c$ wave:

$$W = W_b + \frac{1}{2} \left(\frac{\tilde{c}[\![Z_f]\!]_L^R}{\tilde{c} + \tilde{u}} + \frac{(\tilde{c} - \tilde{u})[\![h]\!]_L^R}{\tilde{c}} + \frac{[\![Q]\!]_L^R}{\tilde{c}} \right) \begin{pmatrix} 0 \\ 1 \\ \tilde{u} + \tilde{c} \end{pmatrix}.$$

3.4. The VFRoe-ncv $(Z_f, 2c, u)$ scheme

We consider herein the change of variable $Y(W) = {}^t(Z_f, 2c, u)$. The choice of variable Y was motivated by the form of Riemann invariants associated with waves of speed $u - c$ and $u + c$ which are, respectively, $u + 2c$ and $u - 2c$ (see (6) and (7)). Moreover, in the flat bottom case (5), variable ${}^t(2c, u)$ provides a symmetrical convection matrix and the condition to maintain a positive intermediate sound speed is formally the same as the condition of vacuum occurrence (8) (see for more details [4,11]).

The system (3a), (3b) may be written related to Y as follows:

$$Z_{f,t} = 0,$$
$$(2c)_{,t} + u(2c)_{,x} + cu_{,x} = 0,$$
$$u_{,t} + c(2c)_{,x} + uu_{,x} + gZ_{f,x} = 0.$$

Note that this system is defined only if $h > 0$ and focusing on smooth solutions. The convection matrix $B(Y)$ is

$$B(Y) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & u & c \\ g & c & u \end{pmatrix}.$$

Eigenvalues of matrix $B(Y)$ read

$$\lambda_1 = 0, \quad \lambda_2 = u - c, \quad \lambda_3 = u + c.$$

If we denote by Ω the matrix of right eigenvectors, we may write

$$\Omega = \begin{pmatrix} u^2 - c^2 & 0 & 0 \\ gc & 1 & 1 \\ -gu & -1 & 1 \end{pmatrix}.$$

The solution provided by the linearised Riemann problem verify through the stationary wave

$$Y^*(0^+; Y_L, Y_R) = Y^*(0^-; Y_L, Y_R) + \frac{[\![Z_f]\!]_L^R}{\tilde{u}^2 - \tilde{c}^2} \begin{pmatrix} \tilde{c}^2 - \tilde{u}^2 \\ g\tilde{c} \\ -g\tilde{u} \end{pmatrix}.$$

The relation between a state Y and a state Y_a through the $u - c$ wave may be written:

$$Y = Y_a + \left(\frac{-g}{2(\tilde{u} - \tilde{c})} [\![Z_f]\!]_L^R + [\![c]\!]_L^R - \frac{[\![u]\!]_L^R}{2} \right) \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix},$$

and the relation to connect a state Y to a state Y_b through the $u + c$ wave is:

$$Y = Y_b + \left(\frac{g}{2(\tilde{u} + \tilde{c})} [\![Z_f]\!]_L^R + [\![c]\!]_L^R + \frac{[\![u]\!]_L^R}{2} \right) \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

One may easily note that the discharge Q computed by the VFRoe-ncv $(Z_f, 2c, u)$ solver is different on both sides of the interface. Hence, the scheme (17) is not conservative according to Eq. (3a). To avoid this problem, a new Finite Volume approximation of (3a) may be introduced:

$$h_i^{n+1} = h_i^n - \frac{\Delta t}{2\Delta x_i} \left((Q_{i+1/2}^- + Q_{i+1/2}^+) - (Q_{i-1/2}^- + Q_{i-1/2}^+) \right), \tag{21}$$

where $Q_{i+1/2}^-$ and $Q_{i+1/2}^+$ refer, respectively, to values at the left and the right side of the interface $x_{i+1/2}$. The scheme obtained from this approximate Riemann solver is able to deal with vacuum in the flat bottom case, according to tests provided in [4]. Moreover, some numerical results are provided in the last section with occurrence of dry area on a non-trivial topography.

3.5. The VFRoe-ncv (Z_f, Q, ψ) scheme

This approximate Riemann solver follows the same formalism as above. We consider herein the variable $Y(W) = {}^t(Z_f, Q, \psi)$ (with $Q = hu$ and $\psi = u^2/2 + g(h + Z_f)$). However, we may remark that this change of variable is not inversible, which may cause some problems to define the numerical flux. The choice of Y is related to the form of the Riemann invariants associated with the null velocity wave (13a) and (13b).

The system (3a) and (3b) written related to Y is

$$Z_{f,t} = 0,$$
$$Q_{,t} + uQ_{,x} + h\psi_{,x} = 0,$$
$$\psi_{,t} + gQ_{,x} + u\psi_{,x} = 0.$$

As a result, the convection matrix $B(Y)$ is

$$B(Y) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & u & h \\ 0 & g & u \end{pmatrix}.$$

As above, eigenvalues of matrix $B(Y)$ are

$$\lambda_1 = 0, \quad \lambda_2 = u - c, \quad \lambda_3 = u + c.$$

If Ω is the matrix of right eigenvectors, we may write

$$\Omega = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -c & c \\ 0 & g & g \end{pmatrix}.$$

The approximate Riemann problem to solve is the same as (18), whose solution $Y^*(x/t; Y_L, Y_R)$ is defined in (19). We have the following relation through the stationary wave:

$$Y^*(0^+; Y_L, Y_R) = Y^*(0^-; Y_L, Y_R) + \begin{pmatrix} [\![Z_F]\!]_L^R \\ 0 \\ 0 \end{pmatrix}.$$

Thus, the solution computed by this Riemann solver is in agreement with the Riemann invariants (13a) and (13b). Hence, this approximate Riemann solver associated with the scheme (17) is able to maintain a large class of steady states, i.e. those based on the Riemann invariants (13a) and (13b) (see Remark 5). A state Y may be connected to a state Y_a through the $u - c$ wave by

$$Y = Y_a + \left(\frac{-1}{2\tilde{c}} [\![Q]\!]_L^R + \tilde{c} [\![\psi]\!]_L^R \right) \begin{pmatrix} 0 \\ -\tilde{c} \\ g \end{pmatrix},$$

and a state Y is connected to a state Y_b through the $u + c$ wave by

$$Y = Y_b + \left(\frac{1}{2\tilde{c}} [\![Q]\!]_L^R + \tilde{c} [\![\psi]\!]_L^R \right) \begin{pmatrix} 0 \\ \tilde{c} \\ g \end{pmatrix}.$$

Remark 3. The convection matrix $B(Y)$ may be written in a symmetrical form, as follows:

$$B(Y) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & u & h \\ 0 & g & u \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & h/g \end{pmatrix}^{-1} \begin{pmatrix} 0 & 0 & 0 \\ 0 & u & h \\ 0 & h & hu/g \end{pmatrix}$$

Remark 4. Note that the system (10a)–(10c) provides a pseudo-conservative form for smooth solutions. Thus, one could use this form to define a Finite Volume scheme from it (with the VFRoe-ncv (Z_f, Q, ψ) solver for instance). However, one can easily verify that, even in the flat bottom case, the Rankine Hugoniot relations are not equivalent. Indeed, noting $v = u - \sigma$ (σ the shock speed) and $\bar{\sigma}$ the arithmetic average, the jump relations provided by the (real) system in the flat bottom case (5) are:

$$[\![hv]\!] = 0, \tag{22}$$

$$\overline{hv}[\![v]\!] + g\bar{h}[\![h]\!] = 0, \tag{23}$$

whereas the jump relations provided by the pseudo-conservative system (10a)–(10c) in the flat bottom case write

$$[\![hv]\!] = 0, \tag{24}$$

$$\bar{v}[\![v]\!] + g[\![h]\!] = 0, \tag{25}$$

which are not equivalent to the previous relations.

Remark 5. According to Remark 1 and relations (13a) and (13b) (assuming that the Riemann invariants and the Rankine Hugoniot relations identify through the LD field), one can define the following discrete steady states:

$$[\![Q]\!]_i^{i+1} = 0, \tag{26a}$$

$$[\![\psi]\!]_i^{i+1} = 0. \tag{26b}$$

Moreover, these states strictly include steady states with $u \equiv 0$

$$u_i = 0, \tag{27a}$$

$$[\![h + Z_f]\!]_i^{i+1} = 0. \tag{27b}$$

Remark 6. Steady states (26a) and (26b) are exactly preserved by the VFRoe-ncv (Z_f, Q, ψ). Moreover, all VFRoe-ncv schemes presented here preserve exactly steady states (27a) and (27b).

Remark 7. All VFRoe-ncv schemes presented above are conservative schemes when the bottom is flat, except when dealing with a stationary shock wave superposed with an interface of the mesh. To avoid this loss of conservativity, the numerical flux in $x_{i+1/2}$ becomes $F((W_{i+1/2}(0^-) + W_{i+1/2}(0^+))/2)$ to compute W_i^{n+1} and W_{i+1}^{n+1}. Hence, this correction provides a conservative scheme which guarantees the correct shock speeds for all shock waves.

We turn now to the second class of methods based on the splitting method.

4. Fractional step method

We present now a new scheme, based on a fractional step method (see [22,27,29]). The system (3a) and (3b) is split into two parts. The first one is the conservative and homogeneous system of P.D.E.:

$$h_{,t} + (hu)_{,x} = 0, \tag{28a}$$

$$(hu)_{,t} + \left(hu^2 + g\frac{h^2}{2}\right)_{,x} = 0. \tag{28b}$$

The second one is the system of O.D.E.:

$$h_{,t} = 0, \tag{29a}$$

$$(hu)_{,t} = -ghZ_f'(x). \tag{29b}$$

The effects of the source term are decoupled from the conservative system. So, a robust method may be applied to compute the system (28a) and (28b) (ensuring positivity of h), and a classical method is used to solve the O.D.E. (29a) and (29b).

4.1. The VFRoe-ncv $(2c, u)$ scheme

To compute the (strictly) hyperbolic, conservative and homogeneous system (28a) and (28b), we propose the VFRoe-ncv $(2c, u)$ scheme (see [4,11]). This system may be written in terms of non-conservative variable $Y(W) = {}^t(2c, u)$. Hence comes:

$$\frac{\partial Y}{\partial t} + B(Y)\frac{\partial Y}{\partial x} = 0$$

with

$$B(Y) = \begin{pmatrix} u & c \\ c & u \end{pmatrix}.$$

Matrix $B(Y)$ is symmetric. The intermediate state is given by (we set here $\widehat{Y} = \overline{Y}$):

$$u_s = \overline{u} - [\![c]\!]_L^R, \tag{30a}$$

$$c_s = \overline{c} - \frac{[\![u]\!]_L^R}{4}, \tag{30b}$$

where $[\![\alpha]\!]_L^R$ represents $\alpha_R - \alpha_L$, for each interface Riemann problem. Note that the linearisation has been made around the state $(2\overline{c}, \overline{u})$.

Vacuum arises in the intermediate state of linearised Godunov solver if and only if initial data makes vacuum occur in the exact solution of the Riemann problem associated with the non-linear set of equations (see condition (8)). Actually, when focusing on the solution of the Riemann problem, vacuum may only occur when initial data are such that two rarefaction waves develop. Riemann invariants are preserved in that case, hence $u + 2c$ (respectively, $u - 2c$) is constant in the 1-rarefaction wave (respectively, the 2-rarefaction wave). Due to the specific form of the linearised system written in terms of non-conservative variable Y, one gets from a discrete point of view:

$$u_R - 2c_R = u_s - 2c_s, \tag{31a}$$
$$u_L + 2c_L = u_s + 2c_s. \tag{31b}$$

Thus, the linearised solver is well suited to handle double rarefaction waves in the solution of the exact Riemann problem. Hence, the discrete condition to ensure the positivity of c_s is

$$u_R - u_L < 2(\sqrt{gh_R} + \sqrt{gh_L}),$$

which exactly identifies with the continuous condition (8).

4.2. The fractional step method

The Finite Volume scheme which computes the homogeneous system (28a) and (28b) may be written as follows:

$$W_i^{n+1/2} = W_i^n - \frac{\Delta t}{\Delta x_i} \left(F(W_{i+1/2}^*(0; W_i, W_{i+1})) - F(W_{i-1/2}^*(0; W_{i-1}, W_i)) \right), \tag{32}$$

where $W_{i+1/2}^*(x/t; W_i, W_{i+1})$ is the solution of the Riemann problem at the interface $x_{i+1/2}$, approximated by the VFRoe-ncv $(2c, u)$ solver.

The system of O.D.E. (29a) and (29b) is approximated by an explicit Euler method for the time part, and by a centered discretisation for the space part:

$$h_i^{n+1} = h_i^{n+1/2},$$
$$Q_i^{n+1} = Q_i^{n+1/2} - \frac{\Delta t}{\Delta x_i} g h_i^{n+1/2} \left(\frac{Z_{f\,i+1} - Z_{f\,i-1}}{2} \right). \tag{33}$$

Note that the property of the VFRoe-ncv $(2c, u)$ scheme concerning the occurrence of vacuum is not modified by step (33). Some numerical results with dry area provided in the following confirm the good behaviour of the fractional step method over vacuum.

Note that neither steady states (26a) and (26b) nor steady states (27a) and (27b) are maintained by the whole algorithm. This phenomenon is well known and will be discussed in the following, based on some numerical experiments, to emphasize that the algorithm is able to converge towards steady states.

Remark 8. In the flat bottom case, the fractional step method (32) and (33) and the VFRoe-ncv $(Z_f, 2c, u)$ scheme presented before provide the same algorithm.

Remark 9. The two steps may be recast in one single step form, as follows:

$$h_i^{n+1} = h_i^n - \frac{\Delta t}{\Delta x_i} \left(Q_{i+1/2}^* - Q_{i-1/2}^* \right),$$
$$Q_i^{n+1} = Q_i^n - \frac{\Delta t}{\Delta x_i} \left(\left(hu^2 - gh^2/2 \right)_{i+1/2}^* - \left(hu^2 - gh^2/2 \right)_{i-1/2}^* \right) - \frac{\Delta t}{\Delta x_i} g h_i^{n+1} \left(\frac{Z_{f\,i+1} - Z_{f\,i-1}}{2} \right),$$

where $(\ )_{i+1/2}^*$ denotes the variable computed by the VFRoe-ncv $(2c, u)$ scheme at the interface $x_{i+1/2}$.

5. A higher-order extension

All schemes previously presented are derived from "first-order" methods. We introduce in this section an extension to obtain more accurate results and to increase rate of convergence (related to the mesh size). This method is based on a linear reconstruction on each cell by the method introduced by Van Leer [28], namely MUSCL (monotonic upwind schemes for conservation laws). This formalism is usually applied in a conservative and homogeneous framework (see [10] for numerical measures with some VFRoe-ncv schemes on Euler system, with non-smooth solutions). However, the source term of topography deeply modifies the structure of the solutions.

When applied to the shallow-water equations on a flat bottom, the MUSCL method would limit the slope of variables h and u for instance. However, the source term of topography must be taken into account. Indeed, refering to a steady state such that $h + Z_f \equiv C^{\text{ste}}$ and $u \equiv 0$, a classical

MUSCL reconstruction breaks the balance of the state. Since a general class of steady states are defined by Q and ψ constant (see Remark 5), one may require that the reconstruction does not modify these states. Moreover, the method must be able to deal with vacuum. We present here a slope limiter which verifies these requirements.

For sake of simplicity, all variables used in this section are supposed to be time-independent. Indeed, the MUSCL method is applied at each time-step, i.e. t is locally fixed to t^n at the nth time-step. Moreover, though this MUSCL method may be computed on irregular meshes, we restrict this presentation to constant space step Δx.

Some notations are first introduced. Let $\{\alpha_i\}_{i\in\mathbb{Z}}$ a variable, constant on each cell, where a cell is $I_i = [x_{i-1/2}; x_{i+1/2}]$. Let $x_i = (x_{i+1/2} + x_{i-1/2})/2$ and $\delta_i(\alpha)$ the (constant) slope associated to α_i, on the cell I_i. Let $\alpha_i^{\text{lin}}(x)$, $x \in I_i$, the function defined on I_i by

$$\alpha_i^{\text{lin}}(x) = \alpha_i - \delta_i(x - x_i), \quad x \in I_i.$$

Thus, to compute numerical flux at an interface $x_{i+1/2}$, the initial data become $\alpha_i^{\text{lin}}(x_{i+1/2})$ and $\alpha_{i+1}^{\text{lin}}(x_{i+1/2})$ of the local Riemann problem instead of α_i and α_{i+1}. This step is the same as in the classical framework.

The modification of the algorithm to take into account the topography is thus restricted to the choice of variables for which the MUSCL reconstruction is applied to and to the computation of the slope δ_i. The first variable is the momentum Q. A classical minmod slope limiter is used (see for instance [20]):

$$\delta_i(Q) = \begin{cases} s_{i+1/2}(Q) \min\left(|Q_{i+1} - Q_i|, |Q_i - Q_{i-1}|\right)/\Delta x & \text{if } s_{i-1/2}(Q) = s_{i+1/2}(Q), \\ 0 & \text{else,} \end{cases} \tag{34}$$

where

$$s_{i+1/2}(\alpha) = \text{sign}(\alpha_{i+1} - \alpha_i).$$

Such a slope limiter is TVD (total variation diminishing) in the following sense:

Property 1. *Let Ω an open subset $\mathbb{R}$ (here $\Omega = \mathbb{R}$). Let us define the total variation of a function $v \in L^1_{\text{loc}}(\Omega)$:*

$$\|v\| = \sup\left\{ \int_\Omega v \, \text{div} \, \phi \, dx, \ \phi \in \mathscr{C}^1_0(\Omega), \ \|\phi\|_{L^\infty(\Omega)} \leqslant 1 \right\}.$$

If v^{cst} and v^{lin} are the functions which, respectively, represent the constant and linear piecewise approximations of v:

$$v^{\text{cst}}(x) = v_i \quad i \in \mathbb{Z} \text{ such that } x \in I_i,$$
$$v^{\text{lin}}(x) = v_i^{\text{lin}} \quad i \in \mathbb{Z} \text{ such that } x \in I_i,$$

then v^{lin} defined by the minmod slope limiter verifies:

$$\|v^{\text{lin}}\| \leqslant \|v^{\text{cst}}\|. \tag{35}$$

The linear reconstruction on Q based on (34) verifies Property 1.

As mentioned above, stationary states must be preserved by the method, in order to permit convergence in time to steady states. To satisfy this requirement, one may choose to apply the

reconstruction on ψ and to verify the Property 1 for ψ. However, the change of variable from (h, Q) to (Q, ψ) is not inversible. Thus, the slope limitation is made on the water height, but the computation of the slope $\delta_i(h)$ is modified to take into account ψ. Let us first define:

$$\delta_i(h) = \begin{cases} s_{i+1/2}(h + Z_f) \min(h_i, |(h + Z_f)_{i+1} - (h + Z_f)_i|, |(h + Z_f)_i - (h + Z_f)_{i-1}|)/\Delta x \\ \quad \text{if } s_{i-1/2}(h + Z_f) = s_{i+1/2}(h + Z_f), \\ 0 \quad \text{else.} \end{cases} \tag{36}$$

The term h_i in the minimum enables the method to deal with vacuum. The profile of ψ does not appear in the computation of $\delta_i(h)$ (though ψ and $g(h + Z_f)$ identify when $u \equiv 0$). Hence, when the source term is locally non-null (i.e. $Z_{fi-1} \neq Z_{fi}$ or $Z_{fi} \neq Z_{fi+1}$), $\delta_i(h)$ must be modified, according to values of ψ_{i-1}, ψ_i and ψ_{i+1}. Since the slope limiters are based on a TVD requirement for the linear reconstruction, we impose a TVD-like condition on ψ, for the computation of $\delta_i(h)$. Let Ψ be the function

$$\Psi(Z_f, h, Q) = \frac{Q^2}{2h^2} + g(h + Z_f).$$

All methods presented in this paper use the following values, $\forall i \in \mathbb{Z}$:

$$\begin{aligned} \Psi_i^- &= \Psi\left(Z_{f\,i}, h_i - \delta_i(h)\frac{\Delta x}{2}, Q_i - \delta_i(Q)\frac{\Delta x}{2}\right), \\ \Psi_i &= \Psi\left(Z_{f\,i}, h_i, Q_i\right) \quad (= \psi_i), \\ \Psi_i^+ &= \Psi\left(Z_{f\,i}, h_i + \delta_i(h)\frac{\Delta x}{2}, Q_i + \delta_i(Q)\frac{\Delta x}{2}\right). \end{aligned}$$

Following these notations, Ψ_i is the value of Ψ at the center of each cell I_i, Ψ_i^- is the value of Ψ at the right of each interface $x_{i-1/2}$ and Ψ_i^+ is the value of Ψ at the left of each interface $x_{i+1/2}$, $i \in \mathbb{Z}$. The computation of numerical flux at an interface $x_{i+1/2}$ needs Ψ_i^+ and Ψ_{i+1}^-. Following notations previously introduced, let Z_f^{cst}, h^{cst} and Q^{cst} be the piecewise constant approximations and let Z_f^{lin}, h^{lin} and Q^{lin} be the piecewise linear approximations. Thus, one can easily verify that $\|\Psi(Z_f^{\text{lin}}, h^{\text{lin}}, Q^{\text{lin}})\|$ is less than or equal to $\|\Psi(Z_f^{\text{cst}}, h^{\text{cst}}, Q^{\text{cst}})\|$. Hence, the reconstructions (34) and (36) do not imply that Ψ Property 1. An idea to solve this problem should be limiting "strongly" h (i.e. computing $\delta_i(h) = 0$) if $\Psi(Z_f^{\text{lin}}, h^{\text{lin}}, Q^{\text{lin}})$ does not verify the TVD requirement. However, this condition may be considered too restrictive. Thus, we introduce the following condition:

$$\begin{aligned} 0 &\leqslant \left|\Psi_i - \Psi_i^-\right| \leqslant \left|\Psi_i - \Psi_{i-1}^-\right|/2, \\ 0 &\leqslant \left|\Psi_i^+ - \Psi_i\right| \leqslant \left|\Psi_{i+1} - \Psi_i\right|/2, \end{aligned} \tag{37}$$

illustrated by Fig. 2.

Condition (37) imposes a TVD-like condition on Ψ. Indeed, assuming that Ψ_i^-, Ψ_i and Ψ_i^+ fulfil conditions (37) $\forall i \in \mathbb{Z}$, if Φ denotes the linear interpolation computed from Ψ_i^-, Ψ_i and Ψ_i^+ $\forall i \in \mathbb{Z}$, then Φ verifies the following TVD property:

$$\|\Phi\| \leqslant \|\Psi\left(Z_f^{\text{cst}}, h^{\text{cst}}, Q^{\text{cst}}\right)\|,$$

which may be seen as the counterpart of (35).
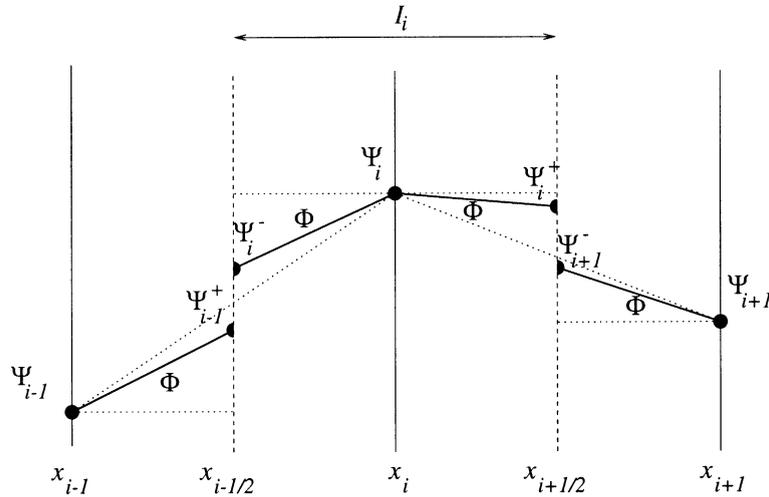
Fig. 2. A TVD-like reconstruction for Ψ.

We recall now all the steps of the algorithm used to compute slopes $\delta_i(h)$ and $\delta_i(Q)$ $\forall i \in \mathbb{Z}$:

1. _Computation of $\delta_i(Q)$:_

$$\delta_i(Q) = \begin{cases} s_{i+1/2}(Q) \min \left(|Q_{i+1} - Q_i|, |Q_i - Q_{i-1}| \right)/\Delta x & \text{if } s_{i-1/2}(Q) = s_{i+1/2}(Q), \\ 0 & \text{else.} \end{cases}$$

2. _Computation of $\delta_i(h)$:_

   - if $Z_{f\,i-1} = Z_{f\,i} = Z_{f\,i+1}$, then the minmod slope limiter is applied to compute $\delta_i(h)$,

   $$\delta_i(h) = \begin{cases} s_{i+1/2}(h) \min \left(|h_{i+1} - h_i|, |h_i - h_{i-1}| \right)/\Delta x & \text{if } s_{i-1/2}(h) = s_{i+1/2}(h), \\ 0 & \text{else,} \end{cases}$$

   - else, $\delta_i(h)$ is first computed by a classical minmod limiter on $h + Z_f$:

   $$\delta_i(h) = \begin{cases} s_{i+1/2}(h + Z_f) \min(h_i, |(h + Z_f)_{i+1} - (h + Z_f)_i|, |(h + Z_f)_i - (h + Z_f)_{i-1}|)/\Delta x \\ \qquad \text{if } s_{i-1/2}(h + Z_f) = s_{i+1/2}(h + Z_f) \\ 0 \quad \text{else,} \end{cases}$$

   - but if condition (37) is not fulfiled, then we reset $\delta_i(h)$ to
   $$\delta_i(h) = 0.$$

Let us emphasize that, when $\delta_i(h)$ is set to 0, conditions (37) may not be verified (because of the limitation on Q).

This slope limiter is combined with a second-order Runge–Kutta integration w.r.t. time (namely the Heun scheme). Some numerical results are described in the following and point out the good behaviour of the slope limiter obtained. A comparison with a classical reconstruction is presented in Section 6.3 (Fig. 13) which shows that the method may fail to converge towards a steady state when t tends to $+\infty$, if the modification described above is not computed.

6. Numerical results

Though several VFRoe-ncv schemes have been previously discussed, only numerical results performed by the VFRoe-ncv $(Z_f, 2c, u)$ scheme with the higher-order extension and by the fractional step method are presented here (some complementary tests are provided in Appendix A). Some experiments tested herein come from a workshop on dam-break wave simulation [16]. Most of them deal with steady states on non-trivial bottom. The ability of the methods to compute dry area is tested too. Let us emphasize that all the numerical results have been obtained without any ''clipping'' treatment (i.e. non-physical values such as negative *cell* water height are not artificially set to 0).

The first four tests are performed with the same topography. The channel length is $l = 25$ m. The bottom Z_f is defined as follows:

$$Z_f(x) = \begin{cases} 0.2 - 0.05(x - 10)^2 & \text{if 8 m} < x < 12 \text{ m}, \\ 0 & \text{else.} \end{cases}$$

Only initial and boundary conditions are modified.

All tests cases are computed with a CFL number set to 0, 4. Results of the flow at rest, the subcritical flow over a bump and the transcritical flow over a bump are plotted at $T_{\text{MAX}} = 200$ s.

6.1. Flow at rest

The initial condition of this test case is a flow at rest. Thus, numerically, it fulfils conditions (27a) and (27b), where $h > 0$. Since we compute a flow at rest, we impose $h + Z_f = \max(Z_f; 0.15)$ m and $Q = 0$ m^2/s all along the mesh, which contains 300 nodes. As expected, the VFRoe-ncv scheme exactly preserves the steady state (Figs. 3 and 4). Moreover, though it is not plotted here, we may emphasize that the behaviour of this scheme remains good in this case when the initial conditions are $h + Z_f = 0.5$ m (no dry cells) or $h = 0$ m (no water). The fractional step method does not maintain $h + Z_f$ and Q constant on the wet cells. The slope of topography introduces a
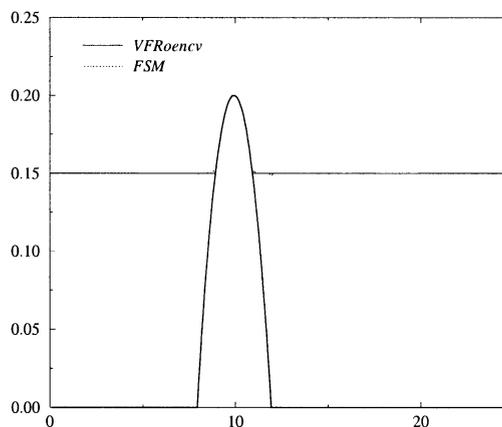
Fig. 3. Flow at rest: water height.
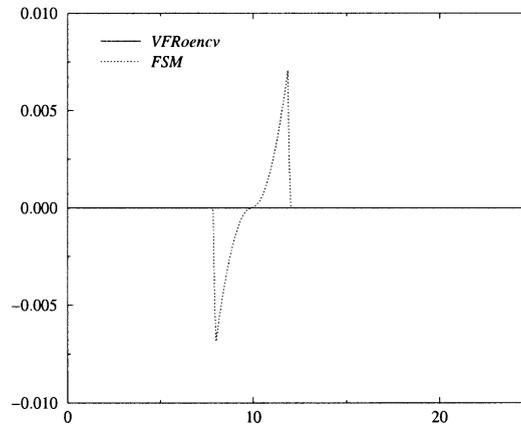
Fig. 4. Flow at rest: discharge.

convection of water. The fractional step method nonetheless converges towards the right solution when the mesh is refined.

The interest of the next three tests (extracted from [16]) is to study the convergence of this scheme towards a steady state. All these tests are performed on 300 cells. The boundary conditions are a positive imposed discharge Q_{in} on the left boundary, and a imposed height h_{out} on the right boundary (except in the case of a supercritical flow). The initial condition is set to $h = h_{out}$ and $Q = 0$. To discuss results, several profiles are plotted, namely h, Q and ψ vs space (in meters). Moreover, to illustrate the quantitative convergence of the methods, the normalised time variation in L^2-norm is plotted too (see Fig. 6 for instance): time t in seconds for x-axis and

$$\ln \frac{\|h^{n+1} - h^n\|_{L^2}}{\|h^3 - h^2\|_{L^2}}$$

for y-axis.

6.2. Subcritical flow over a bump

Here, the boundary conditions are $h_{out} = 2$ m and $Q_{in} = 4.42$ m^2/s. The two solutions provided by the VFRoe-ncv scheme and the fractional step method seem very close to each other, according to Fig. 5 (they are in agreement with the analytic solution). However, Figs. 7 and 8 focus on some differences between the two methods: whereas Q and ψ seem to be constant in the case of the VFRoe-ncv scheme, the fractional step method makes occur oscillations near variations of topography. The two profiles in Fig. 6 are superposed, and show that the two methods converge to steady state.

6.3. Transcritical flow over a bump

The boundary conditions are $Q_{in} = 1.53$ m^2/s and $h_{out} = 0.66$ m. The analytic solution of this test is smooth, with a decreasing part, beginning at the top of the bump and a critical (sonic) point
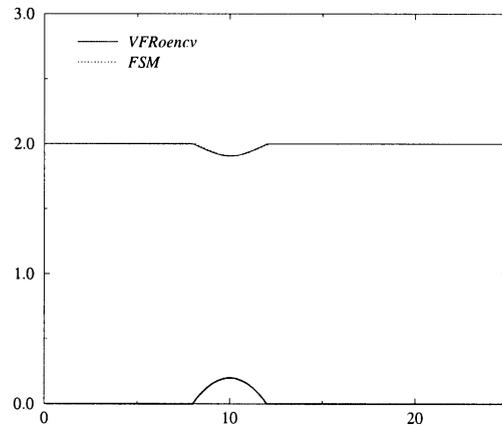
Fig. 5. Subcritical flow: water height.
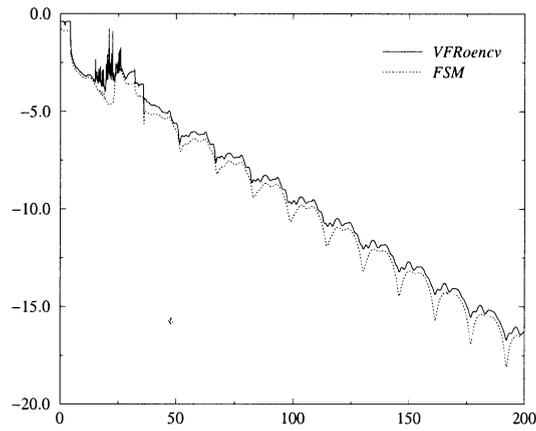
Fig. 6. Subcritical flow: normalised time variation in L^2-norm.
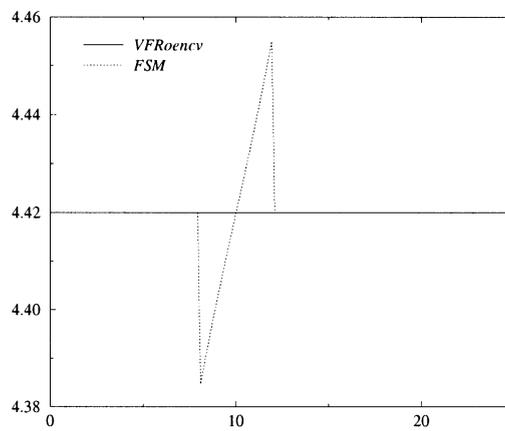
Fig. 7. Subcritical flow: discharge.
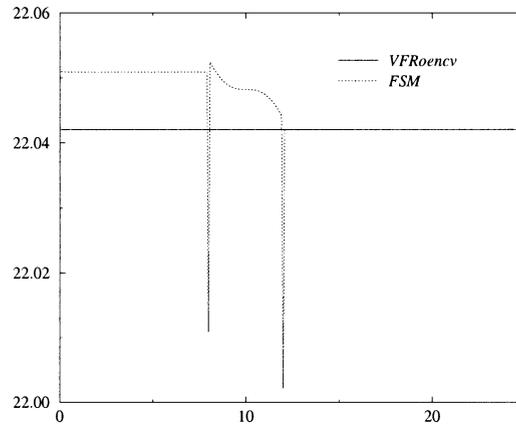
Fig. 8. Subcritical flow: ψ.

on the decreasing part of h. The solution at the right of the decreasing part is supercritical (the boundary condition h_{out} is only used when the flow is subcritical, during the transient part of the simulation). Fig. 9 shows that results provided by the VFRoe-ncv scheme and the fractional step method are similar and the critical point induces no problem (though methods are based on approximate Godunov schemes). According to Fig. 10, the time variation of the VFRoe-ncv scheme decreases slowler than the one of the FSM. In Figs. 11 and 12, one may notice that results performed by the VFRoe-ncv scheme are more accurate, since Q and ψ seem almost constant.

We present now the counterpart of Fig. 10, using the VFRoe-ncv scheme (with the non-conservative variable $(Z_f, 2c, u)$) associated to several reconstructions: the original three-points scheme without any reconstruction, the second-order scheme presented in Section 5 and the classical second-order scheme (i.e. the minmod limiter without modification). Results are plotted in Fig. 13. Whereas the first two schemes provide a similar profile, the second-order scheme with no modification does not converge very well on the coarse mesh. Nonetheless, oscillations remain bounded.
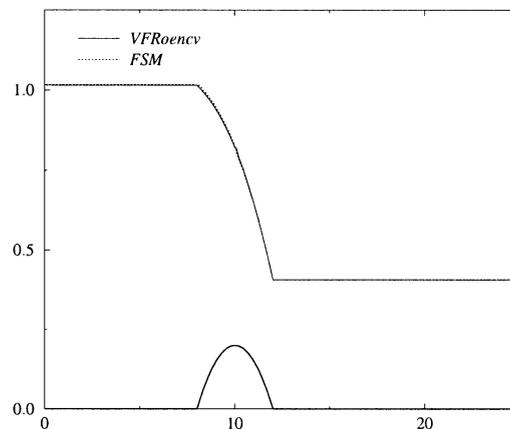
Fig. 9. Transcritical flow: water height.
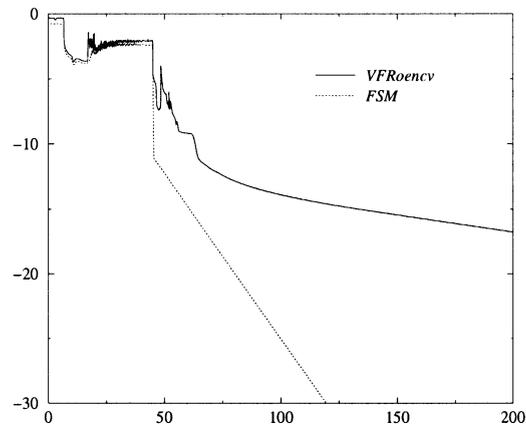
Fig. 10. Transcritical flow: normalised time variation in L^2-norm.
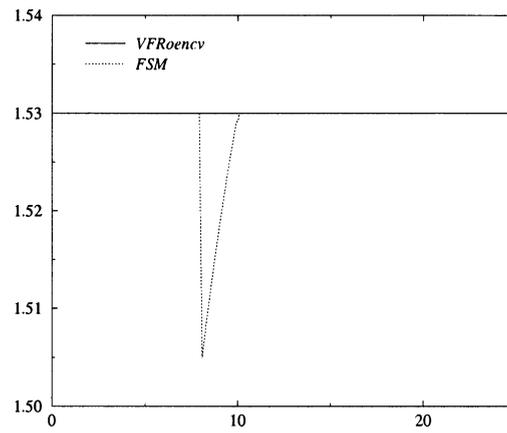
Fig. 11. Transcritical flow: discharge.
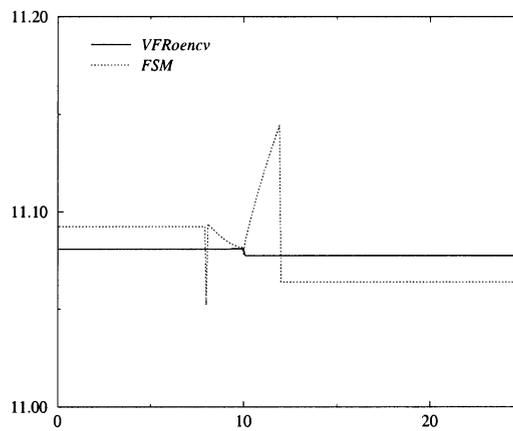
Fig. 12. Transcritical flow: ψ.
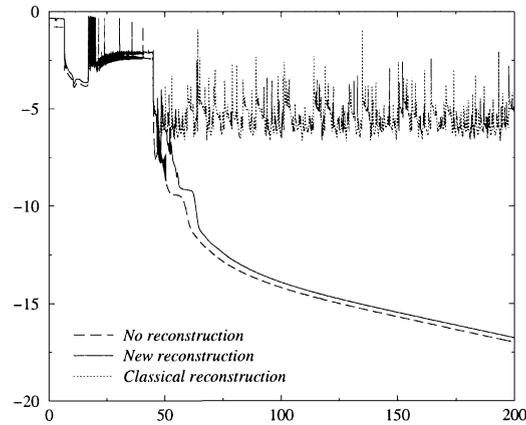
Fig. 13. Transcritical flow: normalised time variation in L^2-norm.

Note that this surprising behaviour can appear since the problem is not in conservative form. In a conservative framework, the second-order scheme with the classical reconstruction converges but the speed of convergence slows down compared with both other schemes (see [26]).

6.4. Drain on a non-flat bottom

The topography of this test case is the same as all cases previously presented. The left boundary condition is a "mirror state"-type condition, and the right boundary condition is an outlet condition on a dry bed [5]. The initial condition is set to $h + Z_f = 0.5$ m and $Q = 0$ m²/s. The solution of this test case at $t = +\infty$ is a state at rest on the left part of top of the bump with $h + Z_f = 0.2$ m and $Q = 0$ m²/s and a dry state (i.e. $h = 0$ m and $Q = 0$ m²/s) on the right side of the bump. Results are presented at several times: $t = 0, 10, 20, 100$ and 1000 s in Figs. 14–17. Note that, since a dry zone is expected at the downstream side of the bump, variable ψ is not defined in
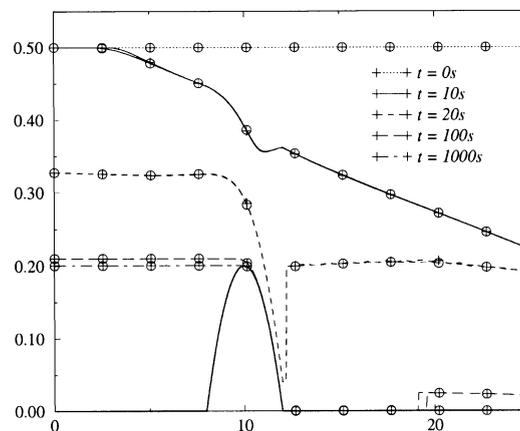
Fig. 14. Drain on a non-flat bottom: water height.
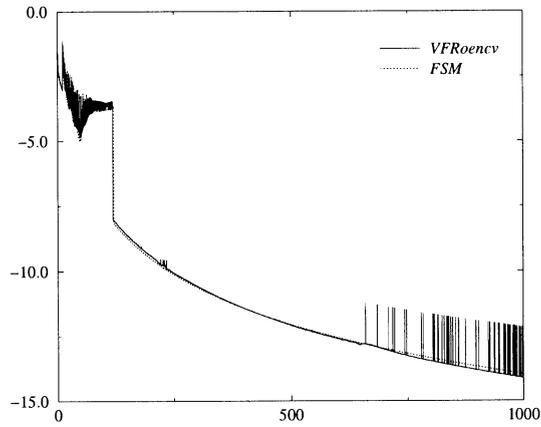
Fig. 15. Drain on a non-flat bottom: normalised time variation in L^2-norm.
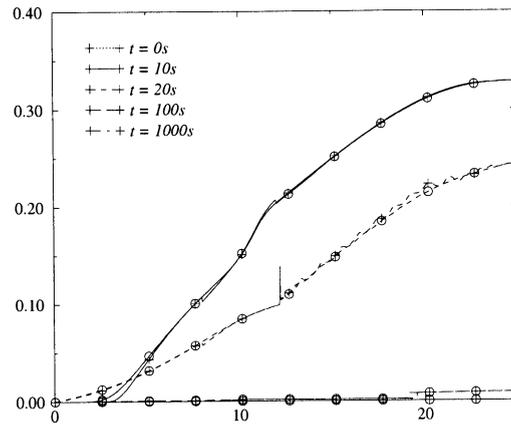
Fig. 16. Drain on a non-flat bottom: discharge.
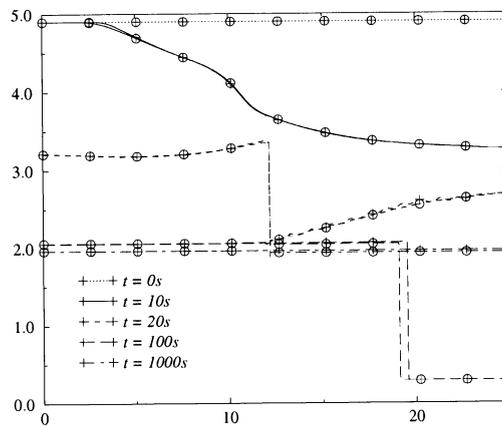
Fig. 17. Drain on a non-flat bottom: ψ.

this zone (thus, results plotted in Fig. 17 in this zone must not be taken into account). Fig. 14 represents the water height computed by the VFRoe-ncv scheme ("plus" symbols) and the fractional step method ("circle" symbols). Results at intermediate times are slightly different, but denote the same behaviour. However, if the final time T_{MAX} is increased, the fractional step method computes a level of water slightly lower than the level expected at the left of the bump, namely $h + Z_f = 0.2$ m. This numerical phenomenon has already been pointed out by Le Roux [19]. It is due to the non-preservation of discrete steady states (27a) and (27b) by the fractional step method. Note however that, when the mesh is refined, the level computed tends to $h + Z_f = 0.2$ m. Results performed by the VFRoe-ncv scheme are rather good, the expected steady state is well approximated, as shown in Figs. 14, 16 and 17. Furthermore, the time variation is decreasing for both methods.

6.5. Vacuum occurrence by a double rarefaction wave over a step

This numerical test is different from previous tests. Indeed, we do not study here the convergence towards a steady state but the ability of the numerical scheme to compute vacuum (i.e. dry bed). Moreover, the topography is not smooth (which indeed is not in agreement with initial assumptions). This test is based on a test proposed by Toro [16], but we introduce here a non-trivial topography: $Z_f = 1$ m if $25/3$ m $< x < 12.5$ m, and $Z_f = 0$ m otherwise (the total length is still 25 m). The initial water height is initialised to 10 m and the initial discharge is set to -350 m^2/s if $x < 50/3$ m and to 350 m^2/s otherwise. Results at several times are presented: 0, 0.05, 0.25, 0.45 and 0.65 s. In the case of a flat bottom, the solution would be composed of two rarefaction waves, with a dry zone occuring between the two waves. Here, since the topography is not flat, the two algorithms introduce waves, located on the jumps of topography (see Figs. 18 and 19, where sign "plus" represents the VFRoe-ncv scheme and the sign "circle" represents the fractional step method). Moreover, one may note that results computed by the two methods are close to each other, but more diffusive for the FSM method (since no MUSCL reconstruction has been performed for this algorithm).
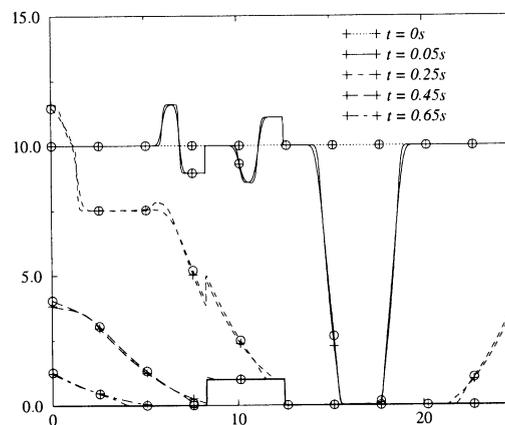

Fig. 18. Vacuum occurrence over a step: water height.
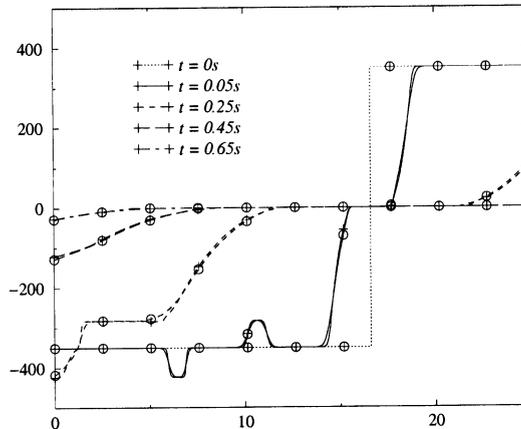
Fig. 19. Vacuum occurrence over a step: discharge.

7. Conclusion

Some Finite Volume schemes have been studied in this paper to compute shallow-water equations with topography. Some relations of the system have been recalled, in the case of a piecewise constant function to approximate the topography. So, according to this approximation, several Finite Volume schemes have been introduced, based on the VFRoe-ncv formalism [5,10], namely the VFRoe-ncv schemes, in variable (Z_f, h, Q), $(Z_f, 2c, u)$ and (Z_f, Q, ψ). All the previous schemes are able to maintain steady states with $u \equiv 0$ and the latter can preserve a larger class of steady states. Moreover, a fractional step method based on the VFRoe-ncv $(2c, u)$ scheme (initially proposed in [4]) is presented. A higher-order extension is also presented, based on the minmod slope limiter, which takes into account steady states.

Refering to numerical results included in [16], one may conclude that the VFRoe-ncv $(Z_f, 2c, u)$ scheme (with the higher-order extension in space and a second-order Runge–Kutta time integration) provides accurate and convergent results. Moreover, the robustness of the method has been emphasised too, dealing with two tests with occurrence of dry area on non-trivial topography, though no clipping treatment has been introduced (i.e. no non-conservative treatment of negative water heights has been computed). The "first-order" fractional step method behaves well (in particular over vacuum), but does not approximate steady states as accurately as the VFRoe-ncv scheme.

Considering results performed by the Well-Balanced scheme, the expected accuracy is shown on some tests. This scheme has been compared with the VFRoe-ncv $(Z_f, 2c, u)$ scheme and numerical results confirm the good behaviour of the latter scheme. However, the Well-Balanced scheme is (several times) more expensive than a usual Godunov method, since the resolution of the Riemann problem is not obvious and many configurations must be considered (this essential difficulty is due to the stationary wave). Indeed, the CPU time required by the "higher"-order VFRoe-ncv $(Z_f, 2c, u)$ scheme is between 10 and 100 times lower than the CPU time required by the "first"-order Well-Balanced scheme.

We have also presented the basic VFRoe scheme (in variable (Z_f, h, Q)), with some results provided in Appendix A. The behaviour of this scheme is as good as the VFRoe-ncv $(Z_f, 2c, u)$ scheme. However, unlike the VFRoe-ncv $(Z_f, 2c, u)$ scheme, this method fails to deal with occurrence of a critical point, provided by an upstream boundary condition. Such a drawback has been emphasised too with the LeVeque scheme [21].

An interesting potential extension of the method presented here is to take into account a variable section $S(x, h)$ in the one-dimensional framework. In this case, the same technique may be used to approximate the corresponding source term.

Acknowledgements

Appendix A. Comparison with the Well-Balanced scheme

This section is devoted to the numerical comparison of the VFRoe-ncv scheme $(Z_f, 2c, u)$ with the Well-Balanced scheme presented in [19]. Note that the VFRoe-ncv scheme is computed with the higher-order extension and a second-order Runge–Kutta method whereas the Well-Balanced scheme tested is the original "first"-order scheme. Two tests are presented: a subcritical flow over a bump and a transcritical flow over a bump. The same topography is used for both tests:

$$Z_f(x) = \begin{cases} 0.2 - 0.05(x - 10)^2 & \text{if } 8 \text{ m} < x < 12 \text{ m}, \\ 0 & \text{else.} \end{cases}$$

Moreover, all results are plotted at $T_{\text{MAX}} = 200$ s. The CFL number is set to 0.4. Computations are performed on a mesh with 300 nodes. Only initial and boundary conditions differ between the two following tests.

A.1. Subcritical flow over a bump

This test computes a transient flow, which tends to become a steady subcritical flow (see test 6.2). The imposed boundary conditions are $Q_{\text{in}} = 4.42$ m²/s and $h_{\text{out}} = 2$ m. The initial conditions are $Q(t = 0, x) = 0$ m²/s and $h(t = 0, x) = h_{\text{out}}$ m. Fig. 20 represents the water height. Results performed by the two schemes are very close to each other. The normalised variation is plotted in Fig. 21. The x-axis is the time and the y-axis is

$$\ln \frac{\|h^{n+1} - h^n\|_{L^2}}{\|h^3 - h^2\|_{L^2}}.$$

One may remark that the two profiles are similar and both methods provide a stationary result. This confirms the good behaviour of the VFRoe-ncv scheme. Figs. 22 and 23 present Q and ψ. Whereas Fig. 22 shows that the two methods provide almost the same results, one can note that the two profiles are slightly different. The analytic solution is $\psi = 22.04205$. The slightly different values provided by the Well-Balanced scheme is due to iterative methods (Newton,
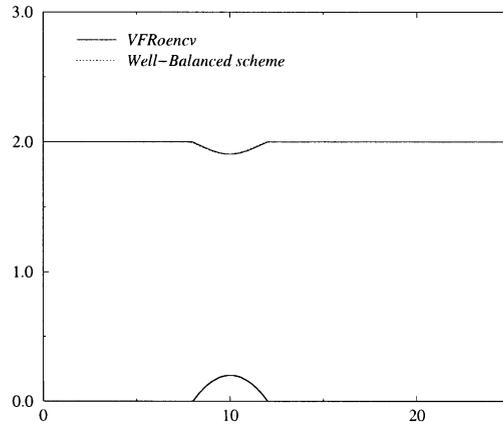
Fig. 20. Subcritical flow: water height.
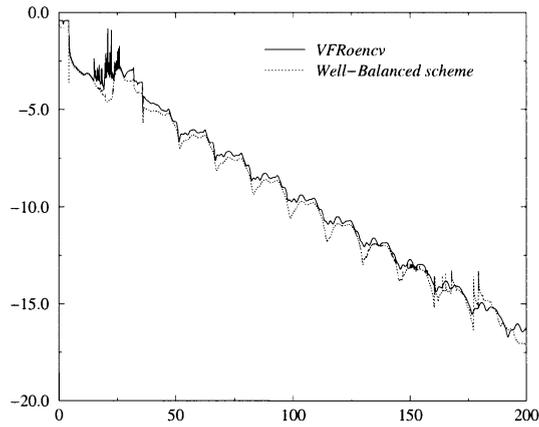
Fig. 21. Subcritical flow: normalised time variation in L^2-norm.
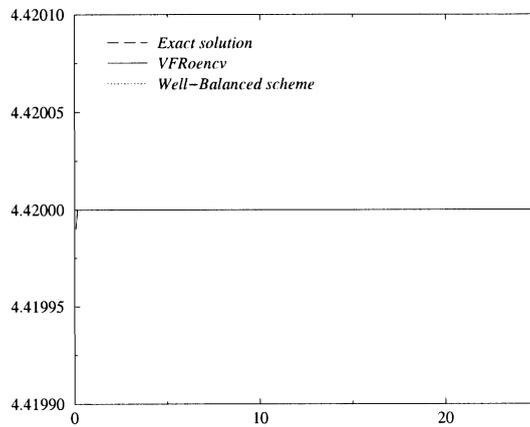

Fig. 22. Subcritical flow: discharge.
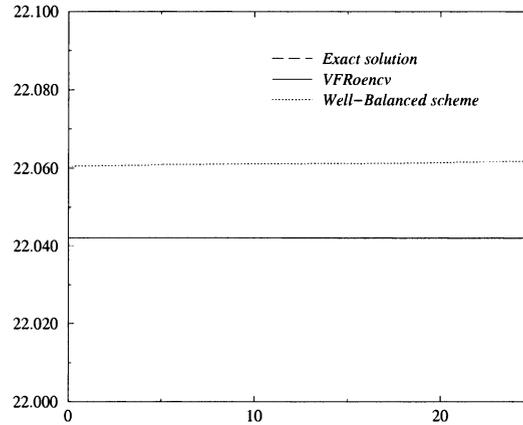
Fig. 23. Subcritical flow: ψ.

dichotomy, etc.) used to compute the exact solution of each interface Riemann problem. Indeed, these methods stop when the relative error is 10^{-5} or when the number of iterations is larger than 500.

A.2. Transcritical flow over a bump

The solution of this test case is a regular profile for the water height, with a subcritical flow upstream of the bump and a supercritical flow downstream of the bump (see test 6.3). The boundary conditions are $Q_{in} = 1.53$ m^2/s and $h_{out} = 0.66$ m. The initial conditions are $Q(t = 0, x) = 0$ m^2/s and $h(t = 0, x) = h_{out}$ m. Both profiles plotted in Fig. 24 provide a good approximation of the expected steady solution. Moreover, Fig. 25 shows that the two schemes compute almost stationary solutions at $t = T_{MAX}$. Fig. 26 shows that variable Q is accurately computed by both methods. Moreover, Fig. 27, which represents variable ψ, denotes a slight difference between the two methods, as it has already been noticed in the previous test case.
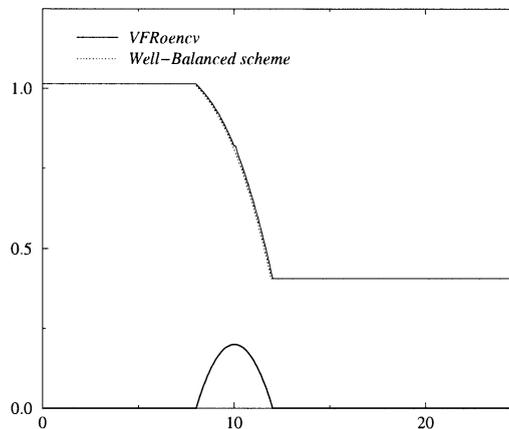

Fig. 24. Transcritical flow: water height.
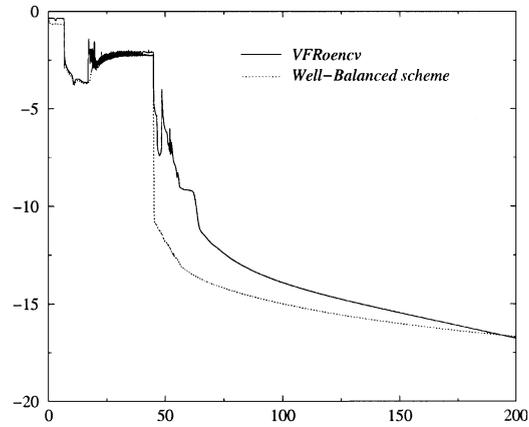
Fig. 25. Transcritical flow: normalised time variation in L^2-norm.
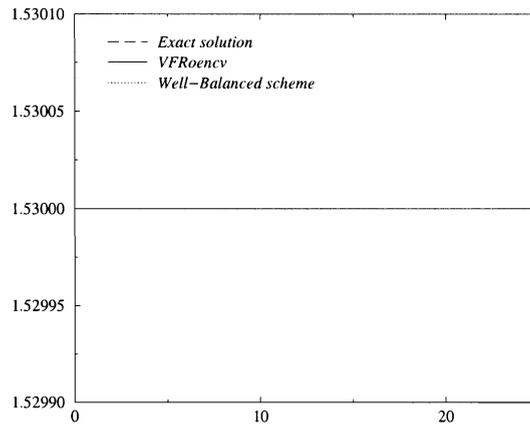

Fig. 26. Transcritical flow: discharge.
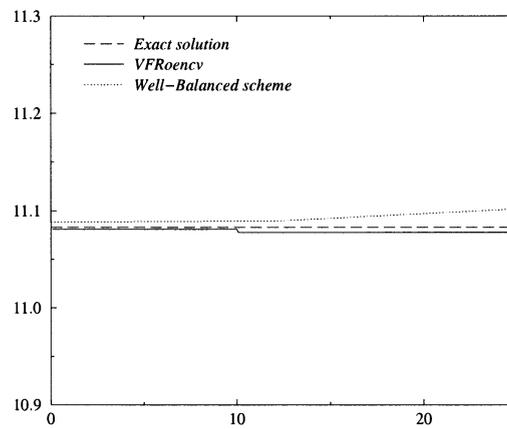
Fig. 27. Transcritical flow: ψ.

This section confirms the good behaviour of the VFRoe-ncv $(Z_f, 2c, u)$ scheme. Indeed, results provided by this method with the higher-order extension are very close to those provided by the Well-Balanced scheme for the two presented test cases. Moreover, the CPU time required by the VFRoe-ncv scheme (with a second-order Runge–Kutta time integration and the higher-order extension) is between 10 and 100 times lower than the CPU time required by the "first"-order Well-Balanced scheme (no accurate CPU measurement might be done, because different computers and different languages have been used to program the methods; no optimisation has been searched for the Well-Balanced scheme; the accuracy and the CPU time of the Well-Balanced scheme deeply depend on the convergence of iterative methods in the exact interface Riemann solver).

Appendix B. Comparison with the VFRoe (Z_f, h, Q) scheme

We present here a numerical test performed with the VFRoe (Z_f, h, Q) scheme, with the higher-order extension previously presented and a second-order Runge–Kutta time approximation. The test case performed is the subcritical flow over a bump (see test 6.2). Let us recall the configuration of this test. The topography is

$$Z_f(x) = \begin{cases} 0.2 - 0.05(x - 10)^2 & \text{if } 8 \text{ m} < x < 12 \text{ m}, \\ 0 & \text{else}. \end{cases}$$

The boundary conditions are $Q_{\text{in}} = 4.42$ m^2/s and $h_{\text{out}} = 2$ m. The initial conditions are $Q(t = 0, x) = 0$ m^2/s and $h(t = 0, x) = 2$ m. The mesh contains 300 cells and the CFL number is 0.4. Both methods provide profiles of water height which are very close to each other in Fig. 28. Moreover, the time variation decreases with the same slope in Fig. 29. Figs. 30 and 31 show that the VFRoe scheme and the VFRoe-ncv scheme both compute almost constant values of Q and ψ, in agreement with the analytic solution (the relative error in L^∞-norm is around 10^{-5}). Thus, the VFRoe scheme also provides accurate results on this test case. As the high-resolution Godunov
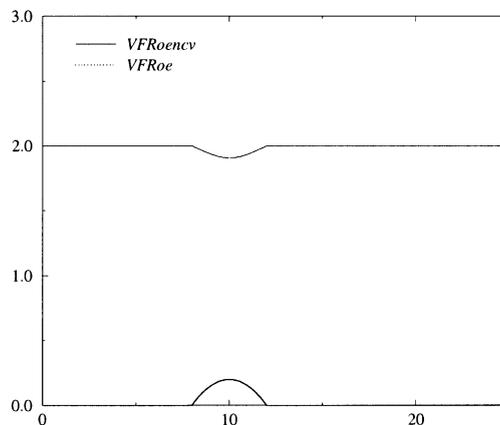

Fig. 28. Subcritical flow: water height.
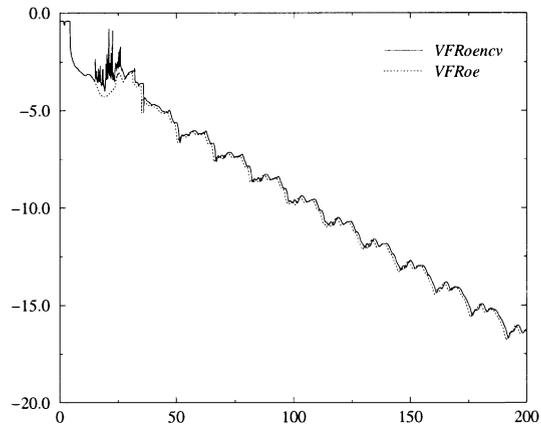
Fig. 29. Subcritical flow: normalised time variation in L^2-norm.
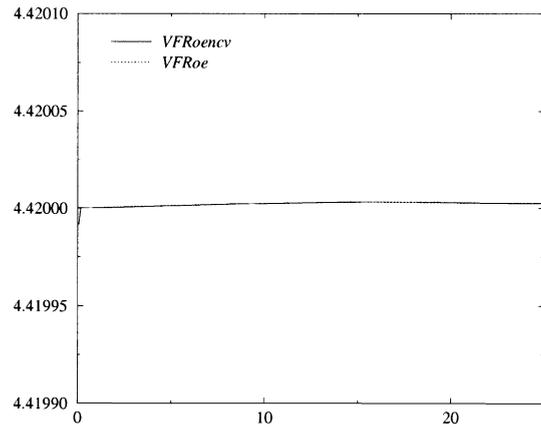

Fig. 30. Subcritical flow: discharge.
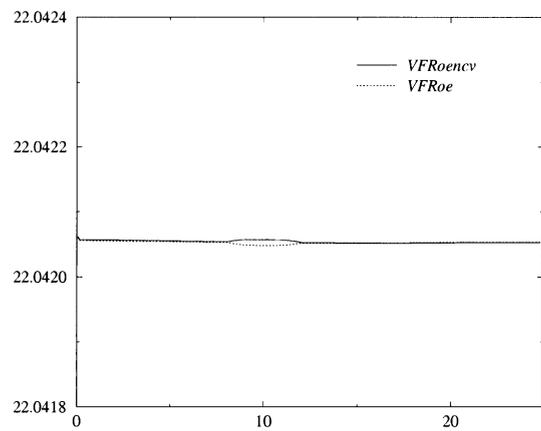
Fig. 31. Subcritical flow: ψ.

method proposed by LeVeque [21], the VFRoe scheme (but not the VFRoe-ncv scheme) fails to deal with occurrence of transcritical flow by an inlet condition (see test 6.3), even with the higher-order extension.

References

[1] Bon C. Modélisation et simulations numériques d'écoulements hydrauliques et de ruissellement en topographic quelconque, PhD thesis, Université Bordeaux I, France, 1997.

[2] Botchorishvili R, Perthame B, Vasseur A. Schémas d'équilibre pour les lois de conservation scalaires avec des termes sources raides, INRIA Report RR-3891, 2000. To appear in Mathematical Methods and Numerical Analysis.

[3] Buffard T, Gallouët T, Hérard JM. A naive Godunov scheme to compute a non-conservative hyperbolic system. Int Ser Numer Math 1998;129:129–38.

[4] Buffard T, Gallouët T, Hérard JM. A naive Godunov scheme to solve shallow-water equations. C R Acad Sci Paris 1998;I-326:885–90.

[5] Buffard T, Gallouët T, Hérard JM. A sequel to a rough Godunov scheme. Application to real gas flows. Comput Fluids 2000;29(7):813–47.

[6] Chinnayya A, Le Roux AY. A new general Riemann solver for the shallow-water equations with friction and topography, 1999. Available from the conservation law preprint server: Available from: http://www.math.ntnu.no/conservation/.

[7] Colombeau JF. Multiplication of distributions. Berlin: Springer; 1992.

[8] Dal Maso G, Le Floch PG, Murat F. Definition and weak stability of nonconservative products. J Math Pures Appl 1995;74:483–548.

[9] Eymard R, Gallouët T, Herbin R. Finite volume methods. In: Ciarlet PG, Lions JL, editors. Handbook of numerical analysis, vol. VII. Amsterdam: North-Holland; 2000. p. 729–1020.

[10] Gallouët T, Hérard JM, Seguin N. Some recent Finite Volume schemes to compute Euler equations using real gas EOS. LATP Report 00-021, Université de Provence, France, 2000. To appear in International Journal for Numerical Methods in Fluids.

[11] Gallouët T, Hérard JM, Seguin N. On the use of some symetrizing variables to deal with vacuum, 2001 [submitted in revised form].

[12] Garcia-Navarro P, Vazquez-Cendon ME. On numerical treatment of the source terms in the shallow water equations. Comput Fluids 2000;29(8):951–79.

[13] Godlewski E, Raviart PA. Numerical approximation of hyperbolic systems of conservation laws. Berlin: Springer; 1996.

[14] Godunov SK. A difference method for numerical calculation of discontinuous equations of hydrodynamics. Mat Sb 1959:271–300 [in Russian].

[15] Gosse L, Le Roux AY. A well balanced scheme designed for inhomogeneous scalar conservation laws. C R Acad Sci Paris 1996;1(323):543–6.

[16] Goutal N, Maurel F. In: Proceedings of the 2nd Workshop on Dam-break Wave Simulation, EDF-DER Report HE-43/97/016/B, 1997.

[17] Greenberg JM, Le Roux AY. A well balanced scheme for the numerical processing of source terms in hyperbolic equation. SIAM J Numer Anal 1996;33(1):1–16.

[18] Greenberg JM, Le Roux AY, Baraille R, Noussair A. Analysis and approximation of conservation laws with source terms. SIAM J Numer Anal 1997;34(5):1980–2007.

[19] Le Roux AY. Discrétisation des termes sources raides dans les problèmes hyperboliques. In: Systèmes hyperboliques: Nouveaux schémas et nouvelles applications. Écoles CEA-EDF-INRIA 'problèmes non linéaires appliqués', INRIA Rocquencourt (France), March 1998 [in French]. Available from: http://www-gm3.univ-mrs.fr/~leroux/publications/ay.le_roux.html.

[20] LeVeque RJ. Numerical methods for conservation laws. Basel: Birkhäusër; 1990.

[21] LeVeque RJ. Balancing source terms and flux gradients in high-resolution Godunov methods. J Comput Phys 1998;146:346–65.

[22] Lions PL, Mercier B. Splitting algorithms for the sum of two nonlinear operators. SIAM J Numer Anal 1979;16(6):964–79.

[23] Masella JM, Faille I, Gallouët T. On an approximate Godunov scheme. Int J Comp Fluid Dyn 1999;12:133–49.

[24] Roe PL. Approximate Riemann solvers, parameter vectors and difference schemes. J Comput Phys 1981;43:357–72.

[25] Seguin N. Génération et validation de Rozavel, un code équilibre en hydraulique 2D.—Mémoire de D.E.A., Université de Bordeaux I, 1999. Available from: http://www-gm3.univ-mrs.fr/~leroux/publications/n.seguin.html.

[26] Seguin N, Vovelle J. Analysis and approximation of a scalar conservation law with a flux function with discontinuous coefficients LATP report, 2002, centre de Mathematique et d'Informatique, 39 rue Joliot Curie, 13453 Marseille cedex 13, France.

[27] Toro EF. Riemann solvers and numerical methods for fluid dynamics. Berlin: Springer; 1997.

[28] Van Leer B. Toward the ultimate conservative difference scheme V. A second order sequel to Godunov's method. J Comp Phys 1979;32:101–36.

[29] Yanenko NN. Méthode à pas fractionnaires. Résolution de problèmes polydimensionnels de physique mathématique, Librairie Armand Colin, 1968. Traduit du russe par P.A. Nepomiastchy.