SIAM J. SCI. COMPUT.
Vol. 30, No. 5, pp. 2587–2612

A POSITIVE PRESERVING HIGH ORDER VFROE SCHEME FOR SHALLOW WATER EQUATIONS: A CLASS OF RELAXATION SCHEMES*

CHRISTOPHE BERTHON† AND FABIEN MARCHE‡

Abstract. The VFRoe scheme has been recently introduced by Buffard, Gallouët, and Hérard [*Comput. Fluids*, 29 (2000), pp. 813–847] to approximate the solutions of the shallow water equations. One of the main interests of this method is to be easily implemented. As a consequence, such a scheme appears as an interesting alternative to other more sophisticated schemes. The VFRoe methods perform approximate solutions in good agreement with the expected ones. However, the robustness of this numerical procedure has not been proposed. Following the ideas introduced by Jin and Xin [*Comm. Pure Appl. Math.*, 45 (1995), pp. 235–276], a relevant relaxation method is derived. The interest of this relaxation scheme is twofold. In the first hand, the relaxation scheme is shown to coincide with the considered VFRoe scheme. In the second hand, the robustness of the relaxation scheme is established, and thus the nonnegativity of the water height obtained involving the VFRoe approach is ensured. Following the same idea, a family of relaxation schemes is exhibited. Next, robust high order slope limiter methods, known as MUSCL reconstructions, are proposed. The final scheme is obtained when considering the hydrostatic reconstruction to approximate the topography source terms. Numerical experiments are performed to attest the interest of the procedure.

Key words. positivity preserving, shallow water equations, VFRoe schemes, relaxation schemes, high order MUSCL extensions

AMS subject classifications. 75M12, 35L65, 65M12

DOI. 10.1137/070686147

1. Introduction.

The present work is devoted to the numerical approximations of the weak solutions of the shallow water equations, also known as the Saint–Venant equations. When specified in one space dimension, the system under consideration reads as follows:

$$\left\{ \begin{array}{l} \partial_t h + \partial_x (hu) = 0, \\ \partial_t (hu) + \partial_x \Big(hu^2 + gh^2/2 \Big) = -ghd'(x), \end{array} \right. \tag{1.1}$$

where h is the local water depth, u is the depth-averaged velocity, and $d : \mathbb{R} \to \mathbb{R}^+$ denotes the topography. For the sake of simplicity in the notations, it is convenient to rewrite the system (1.1) in the following condensed form:

$$\partial_t \mathbf{w} + \partial_x \mathbf{f}(\mathbf{w}) = \mathbf{S}(\mathbf{w}, d), \tag{1.2}$$

with

$$\mathbf{w} = \begin{pmatrix} h \\ hu \end{pmatrix}, \quad \mathbf{f}(\mathbf{w}) = \begin{pmatrix} hu \\ hu^2 + \dfrac{g}{2} h^2 \end{pmatrix}, \quad \text{and} \quad \mathbf{S}(\mathbf{w}) = \begin{pmatrix} 0 \\ -ghd'(x) \end{pmatrix}, \tag{1.3}$$

*Received by the editors March 23, 2007; accepted for publication (in revised form) February 25, 2008; published electronically August 1, 2008.
http://www.siam.org/journals/sisc/30-5/68614.html
†Université de Nantes, Laboratoire de Mathématiques Jean Leray, 2 rue de la Houssiniere, BP 92208, 44322 Nantes Cedex 3, France (christophe.berthon@math.univ-nantes.fr).
‡Institut de Mathématiques et Modélisation de Montepllier, Université Montpellier 2, Place Eugène Bataillon, CC 051, 34000 Montpellier, France (fabien.marche@math.univ-montp2.fr).

where $\mathbf{w} : \mathbb{R} \times \mathbb{R}^+ \to \Omega$ is the state vector in conservative variables and $\mathbf{f}(\mathbf{w}) : \Omega \to \mathbb{R}^2$ stands for the flux function. The convex set Ω of the admissible states is defined by the following:

$$\Omega = \left\{ \mathbf{w} \in \mathbb{R}^2;\ h \geq 0, u \in \mathbb{R} \right\}. \tag{1.4}$$

The homogeneous Saint–Venant system associated with (1.1), given by

$$\partial_t \mathbf{w} + \partial_t \mathbf{f}(\mathbf{w}) = 0, \tag{1.5}$$

is known to be hyperbolic over Ω. As a consequence, the solutions may develop shock discontinuities. In order to rule out the unphysical solutions, the system (1.5) must be supplemented by entropy inequalities [27, 28] (see also [7] and the references therein):

$$\partial_t \tilde{\eta}(\mathbf{w}, d) + \partial_x \tilde{G}(\mathbf{w}, d) \leq 0, \tag{1.6}$$

where

$$\eta(\mathbf{w}) = h\frac{u^2}{2} + \frac{g}{2}h^2, \quad G(\mathbf{w}) = \Big(h\frac{u^2}{2} + gh^2\Big)u,$$
$$\tilde{\eta}(\mathbf{w}, d) = \eta(\mathbf{w}) + hgd, \quad \tilde{G}(\mathbf{w}, d) = G(\mathbf{w}) + hgdu.$$

Let us emphasize that the steady-state solutions of (1.1) are of primary importance. These specific solutions are given by (for instance, see [18, 13]):

$$\begin{cases} hu = \text{cste}, \\ \dfrac{u^2}{2} + gh + gd = \text{cste}. \end{cases} \tag{1.7}$$

One of them, the steady state of a lake at rest, defined by

$$u = 0, \quad h + d = \text{cste},$$

plays a crucial role. From a numerical point of view, the considered schemes must satisfy such a property. After the pioneer work of Greenberg et al. [18, 19], schemes that preserve the lake at rest are called well-balanced schemes.

During the last ten years, several schemes have been proposed according to the well-balanced strategy (see [7, 13, 23, 24, 25]). More recently, in [1], a well-balanced hydrostatic reconstruction has been derived. This technique can be applied to any conservative finite volume scheme approximating the homogeneous Saint–Venant system (1.5) to obtain a suitable approximation of the topography source terms. In addition, as soon as the homogeneous system's scheme is robust, the hydrostatic reconstruction preserves the robustness: It preserves the nonnegativity of h, and it performs the relevant approximations of the dry states, with $h = 0$ (for instance, see also [33, 34] for relevant modelizations and approximations of dry areas).

Involving the framework of the hydrostatic reconstruction, we have to consider robust schemes to approximate the weak solutions of (1.5). This avoids some useful but sophisticated strategy (see [13] for a Godunov approach, or [7] for a relaxation scheme). Now, when approximating (1.5), several schemes come with the required robustness, the Suliciu relaxation scheme [2, 3, 7] (see also [12, 26] for pioneer works), for instance. However, another type of scheme is frequently used when considering the shallow water flows; namely, the VFRoe scheme [14, 15, 17]. Currently, this scheme is

known to be very easy to implement, it performs accurate numerical approximations, and it is able to deal with dry areas. Unfortunately, no robustness result is established for the VFRoe scheme. This is the aim of the present work. The reader is also referred to the work of Jin and Wen [24, 25], where another type of well-balanced scheme is proposed; namely, the interface-type numerical method, which also produces an easy and accurate numerical scheme for the shallow water equations.

The present paper is organized as follows. In the next section, we recall the derivation of the VFRoe scheme. The derivation is performed when involving the shallow water equations. Section 3 is devoted to relaxation schemes. A relevant relaxation model is proposed and analyzed. Following the ideas introduced by Jin and Xin [26] (see also [2, 3, 7, 12]), the resulting relaxation scheme is shown to coincide with the VFRoe scheme. In other words, the VFRoe scheme is written as a relaxation scheme (see LeVeque and Pelanti [31], where the Roe scheme is seen as a relaxation type method). As a consequence, an approximate Riemann solver, associated with the VFRoe scheme, is known and yields to establish the expected nonnegativity of the numerical water depth. Following the same idea, a class of robust relaxation schemes is exhibited. The next section is devoted to the high order extensions. Following an idea introduced by Perthame and Shu [37] (see also [4, 5, 6] for recent developments), we propose high order MUSCL reconstructions, which preserves the robustness property. At this level, the robust full discretization of the shallow water equations (1.1) with topography is obtained when considering the hydrostatic reconstruction. In the last section, numerical tests are performed and attest the interest of the proposed numerical procedure.

2. The VFRoe scheme. The VFRoe schemes are approximate Godunov-type schemes [14, 15]. They turn out to be useful for approximating the solutions of the shallow water equations since such schemes are able to deal with the dry areas. Moreover, their implementations are very easy. We briefly recall the basis of the VFRoe scheme, proposed in [14], involving nonconservative variables. For the sake of simplicity, we restrict ourselves to the regular meshes of size Δx such that $\Delta x = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}$, $i \in \mathbb{Z}$, and we note the time step by Δt, with $t^{n+1} = t^n + \Delta t$, $n \in \mathbb{N}$.

In a more general setting, we consider the numerical approximation of a hyperbolic system of conservation laws in the form (1.5). To access such an issue, the VFRoe scheme is adopted. This approximate Godunov-type scheme is based on the evaluation of the value at the interface between two neighboring cells when considering the following linearization. First, we adopt an admissible change of variable, $\mathbf{u} = \mathcal{U}(\mathbf{w})$. With some abuse in the notations, we set $\mathcal{W}(\mathbf{u}) = \mathbf{w}$ the inverse function of $\mathcal{U}$. Considering smooth-enough solutions, the system (1.5) writes as follows:

$$\partial_t \mathbf{u} + B(\mathbf{u})\partial_x \mathbf{u} = 0, \tag{2.1}$$

where $B(\mathbf{u}) = (\nabla_{\mathbf{u}}\mathcal{W}(\mathbf{u}))^{-1} A(\mathcal{W}(\mathbf{u}))\nabla_{\mathbf{u}}\mathcal{W}(\mathbf{u})$ and $A(\mathbf{w})$ is the Jacobian matrix associated with the flux function $\mathbf{f}$. Next, the following linearized Riemann problem is considered:

$$\partial_t \mathbf{u} + B(\tilde{\mathbf{u}})\partial_x \mathbf{u} = 0, \tag{2.2}$$

$$\mathbf{u}(x,0) = \begin{cases} \mathbf{u}_L = \mathcal{U}(\mathbf{w}_L) & \text{if} \quad x < 0, \\ \mathbf{u}_R = \mathcal{U}(\mathbf{w}_R) & \text{if} \quad x > 0, \end{cases} \tag{2.3}$$

where $\tilde{\mathbf{u}} := \tilde{\mathbf{u}}(\mathbf{u}_L, \mathbf{u}_R)$ represents any averaging of the variable $\mathbf{u}_L$ and $\mathbf{u}_R$. One of the aims of the present paper is to discuss these averagings and to propose relevant

choice to preserve physical assumptions. With some abuse in the notations, we will set $B(\tilde{\mathbf{u}}) := \tilde{B}(\mathbf{u}_L, \mathbf{u}_R)$.

Since the above problem is linear, the exact solution $\mathbf{u}^*(x/t; \mathbf{u}_L, \mathbf{u}_R)$ is easily obtained. A numerical flux function is then defined as follows:

$$\mathbf{f}^*(\mathbf{w}_L, \mathbf{w}_R) = \mathbf{f}\Big(\mathcal{W}(\mathbf{u}^*(0; \mathcal{U}(\mathbf{w}_L), \mathcal{U}(\mathbf{w}_R)))\Big). \tag{2.4}$$

Hence the explicit form of the finite volume VFRoe scheme is given by

$$\mathbf{w}_i^{n+1} = \mathbf{w}_i^n - \frac{\Delta t}{\Delta x}\left(\mathbf{f}^*\Big(\mathcal{W}(\mathbf{u}^*(0; \mathcal{U}(\mathbf{w}_i^n), \mathcal{U}(\mathbf{w}_{i+1}^n)))\Big) - \mathbf{f}^*\Big(\mathcal{W}(\mathbf{u}^*(0; \mathcal{U}(\mathbf{w}_{i-1}^n), \mathcal{U}(\mathbf{w}_i^n)))\Big)\right). \tag{2.5}$$

The numerical flux function is obviously consistent with the exact flux function since we have $\mathbf{u}^*(0, \mathbf{u}, \mathbf{u}) = \mathbf{u}$ for all $\mathbf{u}$, and then we get

$$\mathbf{f}^*(\mathbf{w}, \mathbf{w}) = \mathbf{f}(\mathcal{W}(\mathcal{U}(\mathbf{w}))) = \mathbf{f}(\mathbf{w}).$$

This leads to a conservation scheme (2.5) for any linearization matrix $\tilde{B}(\mathbf{u}_L, \mathbf{u}_R)$.

The solution of the linearized Riemann problem (2.2) is defined everywhere except along $x/t = \tilde{\lambda}_k$ with the following relations:

$$\begin{aligned} \mathbf{u}^*(x/t, \mathbf{u}_L, \mathbf{u}_R) &= \mathbf{u}_L + \sum_{\frac{x}{t} > \tilde{\lambda}_k} ({}^t\tilde{\mathbf{l}}_k \cdot [\mathbf{u}]_L^R)\, \tilde{\mathbf{r}}_k, \\ &= \mathbf{u}_R - \sum_{\frac{x}{t} < \tilde{\lambda}_k} ({}^t\tilde{\mathbf{l}}_k \cdot [\mathbf{u}]_L^R)\, \tilde{\mathbf{r}}_k\,, \end{aligned} \tag{2.6}$$

with $[\mathbf{u}]_L^R = \mathbf{u}_R - \mathbf{u}_L$, where we have set $\tilde{\mathbf{l}}_k$, $\tilde{\lambda}_k$, and $\tilde{\mathbf{r}}_k$, $k = 1, \ldots, p$, respectively, as the left eigenvectors, eigenvalues, and right eigenvectors of $B(\tilde{\mathbf{u}})$.

We recall that the basic VFRoe scheme was first introduced in [35] with the more simple choice $\mathcal{U}(\mathbf{w}) = \mathbf{w}$, and thus $B(\tilde{\mathbf{u}}) = A(\tilde{\mathbf{w}})$. In particular, the stability of the VFRoe scheme in the scalar case is studied in [35]. The various changes of variables for the Euler equations were examined in [8, 9] and in [13, 14] for the shallow water equations.

Now, we specify the scheme when considering the shallow water equations (1.5). The solution of the associated Riemann problem, in the absence of dry areas, is composed of three distinct constant states separated by two genuinely nonlinear fields of speeds $u-c$ and $u+c$, respectively, where the celerity is usually defined by $c = \sqrt{gh}$. Motivated by the form of the Riemann invariant associated with the wave speeds, and far away from the dry areas ($h > 0$), we consider herein the change of variable $\mathcal{U}(\mathbf{w}) = {}^t(2c, u)$. The system (1.5) may be written related to $\mathbf{u}$ in the following nonconservation form:

$$\begin{cases} \partial_t(2c) + u\partial_x(2c) + c\partial_x u = 0, \\ \partial_t u + c\partial_x(2c) + u\partial_x u = 0. \end{cases} \tag{2.7}$$

The Jacobian matrix $B(\mathbf{u})$ becomes

$$B(\mathbf{u}) = \begin{pmatrix} u & c \\ c & u \end{pmatrix}. \tag{2.8}$$

In [14], the averaged state $(\mathbf{u}_L + \mathbf{u}_R)/2$ is proposed for the linearization involved in $B(\tilde{\mathbf{u}})$ and thus for the characterization of $\tilde{u}$ and $\tilde{c}$. In the present work, the choice

of the linearization stays free, and it will be specified later on. Independently on the definition of $\tilde{u}$ and $\tilde{c}$, we obtain the eigenvalues $\lambda^{\pm} = \tilde{u} \pm \tilde{c}$. The exact solution of the linearized Riemann problem (2.2) is easily computed to obtain

$$\mathcal{W}(\mathbf{u}^*(0;\mathcal{U}(\mathbf{w}_i^n),\mathcal{U}(\mathbf{w}_{i+1}^n))) = \begin{cases} \mathbf{w}_i^n & \text{if } (\lambda^-)_{i+\frac{1}{2}} > 0, \\ \mathcal{W}(\mathbf{u}_{i+\frac{1}{2}}) & \text{if } (\lambda^-)_{i+\frac{1}{2}} < 0 < (\lambda^+)_{i+\frac{1}{2}}, \\ \mathbf{w}_{i+1}^n & \text{if } (\lambda^+)_{i+\frac{1}{2}} < 0, \end{cases} \tag{2.9}$$

where $(\lambda^{\pm})_{i+\frac{1}{2}} = \tilde{u}_{i+\frac{1}{2}} \pm \tilde{c}_{i+\frac{1}{2}}$ according to the choice of the adopted linearization of $\tilde{u}$ and $\tilde{c}$. Here, we have set $\mathbf{u}_{i+\frac{1}{2}} = {}^t(2c_{i+\frac{1}{2}}, u_{i+\frac{1}{2}})$, where

$$c_{i+\frac{1}{2}} = \frac{1}{2}(c_i + c_{i+1}) - \frac{1}{4}(u_{i+1} - u_i), \qquad u_{i+\frac{1}{2}} = \frac{1}{2}(u_i + u_{i+1}) - (c_{i+1} - c_i). \tag{2.10}$$

This concludes the presentation of the considered VFRoe scheme. Lastly, we emphasize that, actually, the definition of $\tilde{u}$ and $\tilde{c}$ only influences the upwind definition to use.

3. A relaxation approach. Motivated by the pioneer work of Jin and Xin [26] (see also [11, 32]), we propose to introduce a relevant relaxation scheme to obtain a suitable interpretation of the VFRoe scheme (2.5)–(2.9). We approximate the weak solutions of the system (1.5) by the weak solutions of a suitable first order system with singular perturbations: *the relaxation model*. According to the work of Coquel and Perthame [12] or Bouchut [7] (see also [2, 3]), we propose a relaxation model which preserves most of the nonlinearities of the initial system. The resulting numerical scheme, based on the relaxation model, will be seen to coincide with the VFRoe method.

3.1. The relaxation model. Following the work of Suliciu [39, 40] (see also [2, 3, 7, 12]), we suggest substituting the celerity c and velocity u by the approximations Σ and U. These two new variables are governed by

$$\begin{cases} \partial_t \Sigma + \bar{u}\partial_x \Sigma + \dfrac{\bar{c}}{2}\partial_x U = \mu(c - \Sigma), \\ \partial_t U + 2\bar{c}\partial_x \Sigma + \bar{u}\partial_x U = \mu(u - U), \end{cases}$$

where $\bar{c}$ and $\bar{u}$ are relaxation parameters to be defined, while μ is a parameter intended to tend to infinity. The following first order system with singular perturbations:

$$\begin{cases} \partial_t h + \partial_x \left(\dfrac{\Sigma^2}{g} U \right) = 0, \quad t > 0,\ x \in \mathbb{R}, \\ \partial_t (hu) + \partial_x \left(\dfrac{\Sigma^2}{2g}(2U^2 + \Sigma^2) \right) = 0, \\ \partial_t \Sigma + \bar{u}\partial_x \Sigma + \dfrac{\bar{c}}{2}\partial_x U = \mu(c - \Sigma), \\ \partial_t U + 2\bar{c}\partial_x \Sigma + \bar{u}\partial_x U = \mu(u - U) \end{cases} \tag{3.1}$$

is considered to approximate the weak solutions of (1.5). Now, at least formally, in the limit of μ to infinity, we recover the initial system (1.5). This limit will be referred to as the equilibrium limit, defined by $\Sigma = c$ and $U = u$. The conservation of the water height h and the discharge hu in (3.1) then gives those of (1.5). In the following, we

establish that the scheme obtained from the relaxation model (3.1) is relevant, in a sense to be precised, to approximate solutions involving dry areas.

For the sake of simplicity in the notations, let us introduce the following abstract form of the relaxation system (3.1):

$$\partial_t \mathbf{W} + \partial_x \mathbf{F}(\mathbf{W}) = \mu \mathbf{R}(\mathbf{W}), \tag{3.2}$$

where we have set $\mathbf{W} = {}^t(h, hu, \Sigma, U)$ defined over the following convex set:

$$\mathcal{V} = \left\{ \mathbf{W} \in \mathbb{R}^4;\ h \geq 0 \right\}.$$

One interest of the relaxation model (3.1) stays in the linear degeneracy property satisfied by all of the fields. As a consequence, the Riemann problem turns out to be easy to solve. The next statement is devoted to solve the Riemann problem.

LEMMA 3.1. *Let $\bar{c} > 0$ and $\bar{u} \in \mathbb{R}$ such that $\bar{u} \pm \bar{c} \neq 0$. Assume $\mu = 0$. The first order system $(3.1)_{\mu=0}$ is hyperbolic for all $\mathbf{W} \in \mathcal{V}$. It admits $\lambda_1^0 = \lambda_2^0 = 0$ and $\lambda^{\pm} = \bar{u} \pm \bar{c}$ as eigenvalues, and the associated fields are linearly degenerated.*

Let $\mathbf{W}_L$ and $\mathbf{W}_R$ be constant states in $\mathcal{V}$ and define

$$\mathbf{W}_0(x) = \begin{cases} \mathbf{W}_L & \text{if } x < 0, \\ \mathbf{W}_R & \text{if } x > 0, \end{cases} \tag{3.3}$$

the initial data of the Riemann problem for the system $(3.1)_{\mu=0}$. Let us set

$$\Sigma^* = \frac{\Sigma_L + \Sigma_R}{2} - \frac{1}{4}(U_R - U_L), \tag{3.4}$$

$$U^* = \frac{U_L + U_R}{2} - (\Sigma_R - \Sigma_L). \tag{3.5}$$

Then the weak solution of the system $(3.1)_{\mu=0}$ and for the initial data (3.3) is given by

1. *If $\lambda^- < 0 < \lambda^+$:*

$$\mathbf{W}(x,t) = \begin{cases} \mathbf{W}_L & \text{if } \quad x/t < \lambda^-, \\ \mathbf{W}_L^* & \text{if } \lambda^- < x/t < 0, \\ \mathbf{W}_R^* & \text{if } 0 < x/t < \lambda^+, \\ \mathbf{W}_R & \text{if } \lambda^+ < x/t, \end{cases}$$

where

$$\mathbf{W}_L^* = {}^t(h_L^*, q_L^*, \Sigma^*, U^*), \quad \mathbf{W}_R^* = {}^t(h_R^*, q_R^*, \Sigma^*, U^*),$$

and the values of (h_L^*, q_L^*) and (h_R^*, q_R^*) are given by

$$h_L^* = h_L + \frac{\Sigma^{*2} U^* - \Sigma_L^2 U_L}{g\lambda^-}, \quad q_L^* = (hu)_L - \frac{\mathcal{I}(\Sigma^*, U^*) - \mathcal{I}(\Sigma_L, U_L)}{\lambda^-}, \tag{3.6}$$

$$h_R^* = h_R + \frac{\Sigma^{*2} U^* - \Sigma_R^2 U_R}{g\lambda^+}, \quad q_R^* = (hu)_R - \frac{\mathcal{I}(\Sigma^*, U^*) - \mathcal{I}(\Sigma_R, U_R)}{\lambda^+}, \tag{3.7}$$

where the function $\mathcal{I} : \mathbb{R}^2 \to \mathbb{R}$ is defined by

$$\mathcal{I}(\Sigma, U) = \frac{\Sigma^2}{2g}(\Sigma^2 + 2U^2). \tag{3.8}$$

2. *If* $0 < \lambda^- < \lambda^+$*:*

$$\mathbf{W}(x,t) = \begin{cases} \mathbf{W}_L & \text{if} \quad x/t < 0, \\ \mathbf{W}_L^* & \text{if} \quad 0 < x/t < \lambda^-, \\ \mathbf{W}_R^* & \text{if} \quad \lambda^- < x/t < \lambda^+, \\ \mathbf{W}_R & \text{if} \quad \lambda^+ < x/t, \end{cases}$$

where

$$\mathbf{W}_L^* = {}^t(h_L^*,\, q_L^*,\, \Sigma_L,\, U_L), \quad \mathbf{W}_R^* = {}^t(h_R^*,\, q_R^*,\, \Sigma^*,\, U^*),$$

and the values of (h_L^*, q_L^*) *and* (h_R^*, q_R^*) *are given by*

$$h_L^* = h_R^* + \frac{\Sigma_L^2 U_L - \Sigma^{*2} U^*}{g\lambda^-}, \quad q_L^* = q_R^* - \frac{\mathcal{I}(\Sigma_L, U_L) - \mathcal{I}(\Sigma^*, U^*)}{\lambda^-}, \tag{3.9}$$

$$h_R^* = h_R + \frac{\Sigma^{*2} U^* - \Sigma_R^2 U_R}{g\lambda^+}, \quad q_R^* = (hu)_R - \frac{\mathcal{I}(\Sigma^*, U^*) - \mathcal{I}(\Sigma_R, U_R)}{\lambda^+}, \tag{3.10}$$

where the function $\mathcal{I}$ *is defined by* (3.8).

3. *If* $\lambda^- < \lambda^+ < 0$*:*

$$\mathbf{W}(x,t) = \begin{cases} \mathbf{W}_L & \text{if} \quad x/t < \lambda^-, \\ \mathbf{W}_L^* & \text{if} \quad \lambda^- < x/t < \lambda^+, \\ \mathbf{W}_R^* & \text{if} \quad \lambda^+ < x/t < 0, \\ \mathbf{W}_R & \text{if} \quad 0 < x/t, \end{cases}$$

where

$$\mathbf{W}_L^* = {}^t(h_L^*,\, q_L^*,\, \Sigma^*,\, U^*), \quad \mathbf{W}_R^* = {}^t(h_R^*,\, q_R^*,\, \Sigma_R,\, U_R),$$

and the values of (h_L^*, q_L^*) *and* (h_R^*, q_R^*) *are given by*

$$h_L^* = h_L + \frac{\Sigma^{*2} U^* - \Sigma_L^2 U_L}{g\lambda^-}, \quad q_L^* = (hu)_L - \frac{\mathcal{I}(\Sigma^*, U^*) - \mathcal{I}(\Sigma_L, U_L)}{\lambda^-}, \tag{3.11}$$

$$h_R^* = h_L^* + \frac{\Sigma_R^2 U_R - \Sigma^{*2} U^*}{g\lambda^+}, \quad q_R^* = q_L^* - \frac{\mathcal{I}(\Sigma_R, U_R) - \mathcal{I}(\Sigma^*, U^*)}{\lambda^+}, \tag{3.12}$$

where the function $\mathcal{I}$ *is, once again, defined by* (3.8).

Let us note, from now on, that the intermediate values Σ^* and U^* remain free from the parameters $\bar{c}$ and $\bar{u}$. This remark will be crucial in what follows when establishing a relevant relationship between the relaxation scheme and the VFRoe scheme. Indeed, we will see that the relaxation numerical flux function solely depends on the pair (Σ^*, U^*), but for the definition (3.4)–(3.5) this function will coincide with the VFRoe numerical flux function given by (2.5)–(2.9). This remark will be obtained independently from the precise definition of the pair $(\bar{u}, \bar{c})$.

Proof. Considering the algebra of the system (3.2), the flux function Jacobian matrix is given by

$$\nabla_{\mathbf{W}}\mathbf{F}(\mathbf{W}) = \begin{pmatrix} 0 & 0 & 2\Sigma U/g & \Sigma^2/g \\ 0 & 0 & \Sigma(U^2+\Sigma^2)/g & 2\Sigma^2 U/g \\ 0 & 0 & \bar{u} & \bar{c}/2 \\ 0 & 0 & 2\bar{c} & \bar{u} \end{pmatrix}. \tag{3.13}$$

Easy calculations ensure that the vectors given by $\mathbf{r}_1^0 = {}^t(1,0,0,0)$, $\mathbf{r}_2^0 = {}^t(0,1,0,0)$, and

$$\mathbf{r}^{\pm} = \begin{pmatrix} 2\Sigma\, U/g \pm 2\Sigma^2/g \\ 2\Sigma(U^2+\Sigma^2)/g \pm 4\Sigma^2 U/g \\ \bar{u} \pm \bar{c} \\ \pm 2(\bar{u} \pm \bar{c}) \end{pmatrix}$$

are the right eigenvectors, respectively, associated to the eigenvalues $\lambda_1^0 = \lambda_2^0 = 0$ and $\lambda^{\pm} = \bar{u} \pm \bar{c}$. In addition, it is easy to see that all of the fields are linearly degenerated.

The Riemann solution is made up of four constant states, $\mathbf{W}_L$, $\mathbf{W}_L^*$, $\mathbf{W}_R^*$, and $\mathbf{W}_R$, separated by three contact discontinuities, a stationary and two propagating at velocity $\lambda^{\pm}$. The above third developed cases aim only at dealing with the location of the wave velocities λ^- and λ^+ with respect to the stationary discontinuity. We recall that, across the jth-contact discontinuity, the Riemann invariants associated with the jth-eigenvector are continuous. These Riemann invariants denoted ϕ_j are defined by $\nabla\phi_j \cdot \mathbf{r}_j = 0$, where $\mathbf{r}_j$ is the jth right eigenvector introduced above. After straightforward computations, we obtain the following Riemann invariants, defined field by field:

$$\begin{array}{lll} \phi_1^0 = \Sigma, & \phi_2^0 = U, & \\ \phi_1^{\pm} = 2\Sigma \pm U, & \phi_2^{\pm} = (\bar{u} \pm \bar{c})h - U\,\Sigma^2/g, & \phi_3^{\pm} = (\bar{u} \pm \bar{c})hu - \mathcal{I}(\Sigma, U). \end{array} \tag{3.14}$$

Now, considering the case $\lambda^- < 0 < \lambda^+$ and exploiting the continuity of these invariants across the associated contact discontinuity, we obtain the following system of equations:

$$\begin{cases} \Sigma_L^* = \Sigma_R^*, \\ U_L^* = U_R^*, \end{cases} \tag{3.15}$$

$$\begin{cases} U_L + 2\Sigma_L = U^* + 2\Sigma^*, \\ \lambda^- h_L - \frac{U_L \Sigma_L^2}{g} = \lambda^- h_L^* - \frac{U^*(\Sigma^*)^2}{g}, \\ \lambda^-(hu)_L - \mathcal{I}(\Sigma_L, U_L) = \lambda^- q_L^* - \mathcal{I}(\Sigma^*, U^*), \end{cases} \tag{3.16}$$

$$\begin{cases} U^* - 2\Sigma^* = U_R - 2\Sigma_R, \\ \lambda^+ h_R^* - \frac{U^* \Sigma_L^2}{g} = \lambda^+ h_R - \frac{U_R \Sigma_R^2}{g}, \\ \lambda^+(hu)_R - \mathcal{I}(\Sigma_R, U_R) = \lambda^+ q_R^* - \mathcal{I}(\Sigma^*, U^*). \end{cases} \tag{3.17}$$

The unique solution of this system is given by (3.4)–(3.7). The two remaining cases are developed following the same lines. The proof is thus completed. ∎

This analysis of the Riemann problem solutions can be supplemented by a study of the nonnegativity of the water depth h, establishing that $h_{L,R}^* \geq 0$ as soon as $h_{L,R} \geq 0$. At the moment, let us focus our attention on h_L^*; the same approach will be used considering h_R^*.

Since Σ^* and U^* do not depend on the eigenvalues $\lambda^{\pm}$, while $\lambda^{\pm}$ does not depend on the unknowns but just on the fixed parameters $\bar{u}$ and $\bar{c}$, with $h_L > 0$ and $h_R > 0$ it is clear that h_L^* remains nonnegative as soon as $|\lambda^{\pm}|$ is large enough. Now, assume that $h_L = 0$ and $h_R > 0$, to write with $U_L = 0$:

$$h_L^* = \begin{cases} \frac{1}{\lambda^- g}(\Sigma^*)^2 U^* & \text{if } \lambda^- < 0, \\ h_R + \frac{1}{\lambda^+ g}\left((\Sigma^*)^2 U^* - \Sigma_R^2 U_R\right) - \frac{1}{\lambda^- g}(\Sigma^*)^2 U^* & \text{if } \lambda^- > 0. \end{cases}$$

Two cases must be distinguished. If $U^* < 0$, we have just to set $\lambda^- < 0$. Reversely, if we have $U^* > 0$, the parameters $\bar{u}$ and $\bar{c}$ must be chosen to enforce $0 < \lambda^- < \lambda^+$ large enough to satisfy $h_L^* > 0$. Finally, in the case of $h_L = h_R = 0$, with the convention $U_L = U_R = 0$, we obtain $h_L^* = 0$. Involving the same analysis with h_R^*, we have easily established the following result.

LEMMA 3.2. *Assume that $h_L \geq 0$ and $h_R \geq 0$, with the convention $U_L = 0$ if $h_L = 0$ and $U_R = 0$ if $h_R = 0$. Then there exist suitable parameters $\bar{u}$ and $\bar{c}$ such that the functions h_L^* and h_R^*, defined in Lemma* 3.1*, are nonnegative.*

In order to help the reader in the full determination of the relaxation model, and thus in the characterization of the relaxation scheme, we underline that the relevant definition of the eigenvalues $\lambda^{\pm}$ coincides with the following choice of the relaxation parameters $(\bar{u}, \bar{c})$:

$$\bar{u} = \frac{\lambda^+ + \lambda^-}{2} \quad \text{and} \quad \bar{c} = \frac{\lambda^+ - \lambda^-}{2}.$$

From a practical point of view, one wants to define $\bar{u} = (u_L + u_R)/2$ and $\bar{c} = (c_L + c_R)/2$ (see [14]). As an example, such a choice can be considered and locally modified as soon as the nonnegativity of h_L^* or h_R^* is violated. Of course, this local correction of $\bar{u}$ and $\bar{c}$ is done according to Lemma 3.2. It is clear that many other choices can be suggested, and the numerical consequences resulting from the considered definition of $\bar{u}$ and $\bar{c}$ will be detailed in the next section. Under a more general framework, we can set $\bar{u} := \bar{u}(\mathbf{W}_L, \mathbf{W}_R)$ and $\bar{c} := \bar{c}(\mathbf{W}_L, \mathbf{W}_R)$, which reads $\bar{u} := \bar{u}(\mathbf{w}_L, \mathbf{w}_R)$ and $\bar{c} := \bar{c}(\mathbf{w}_L, \mathbf{w}_R)$ as soon as the states $\mathbf{W}_L$ and $\mathbf{W}_R$ are assumed to satisfy the equilibrium, defined by $U = u$ and $\Sigma = \sqrt{gh}$.

3.2. The relaxation solver. For the sake of completeness, we briefly recall the numerical relaxation procedure to approximate a weak solution of (1.5), which is usually in the framework of the relaxation scheme (see Jin and Xin [26], but also Coquel and Perthame [12], Baudin et al. [2], Bouchut [7] or Berthon [3]).

We consider the same mesh notations as introduced in section 2. To approximate the solution at time $t^{n+1} = t^n + \Delta t$, a splitting technique is adopted. In a first step, we solve the relaxation model (3.2) omitting the relaxation source terms which are considered in a second step. As usual, we assume that a piecewise constant approximate equilibrium solution $\mathbf{w}^h(x, t^n) \in \Omega$ is known at time t^n, defined by

$$\mathbf{w}^h(x, t^n) = \mathbf{w}_i^n = {}^t\left(h_i^n, (hu)_i^n\right), \quad x \in (x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}).$$

At the initial time, we set

$$\mathbf{w}_i^0 = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \mathbf{w}(x, 0) dx.$$

During the first step, we propose to evolve in time a relevant approximation of the relaxation model (3.2). To access such an issue, we introduce $\mathbf{W}^h \in \mathcal{V}$ such that, for all $0 < t < \Delta t$, the function $\mathbf{W}^h(x, t^n + t)$ is the weak solution of the Cauchy problem for the relaxation system $(3.2)_{\mu=0}$, namely,

$$\partial_t \mathbf{W} + \partial_x \mathbf{F}(\mathbf{W}) = 0, \tag{3.18}$$

supplemented by the following initial equilibrium data:

$$\begin{aligned} \mathbf{W}^h(x, t^n) &= \mathbf{W}_i^n \\ &= {}^t\left(h_i^n, (hu)_i^n, \Sigma_i^n, U_i^n\right), \quad x \in (x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}), \end{aligned}$$

where the equilibrium state is defined by $\Sigma_i^n = \sqrt{g\, h_i^n}$ and $U_i^n = (hu)_i^n / h_i^n$. Under the CFL-like condition

$$\frac{\Delta t}{\Delta x} \max_{i\in\mathbb{Z}} \left(|\lambda^-_{i+\frac{1}{2}}|, |\lambda^+_{i+\frac{1}{2}}| \right) \leq \frac{1}{2}, \tag{3.19}$$

the solution $\mathbf{W}^h$ at the time $t^n + \Delta t$ is made up of the juxtaposition of the noninteracting Riemann problem solution set at the cell interfaces $x_{i+\frac{1}{2}}$ for $i \in \mathbb{Z}$. Next, the projection of this solution on the piecewise constant functions reads

$$\mathbf{W}_i^{n+1,-} = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \mathbf{W}^h(x, t^n + \Delta t) dx.$$

A local definition of the parameter $\bar{u}$ and $\bar{c}$ at each interface $x_{i+\frac{1}{2}}$ is considered (for instance, see [14]). At each interface $x_{i+\frac{1}{2}}$, we choose $\mathbf{W}_L = \mathbf{W}_i^n$ and $\mathbf{W}_R = \mathbf{W}_{i+1}^n$ to define the parameters $\bar{u}_{i+\frac{1}{2}} := \bar{u}(\mathbf{w}_i^n, \mathbf{w}_{i+1}^n)$ and $\bar{c}_{i+\frac{1}{2}} := \bar{c}(\mathbf{w}_i^n, \mathbf{w}_{i+1}^n)$ according to the nonnegativity condition stated Lemma 3.2. Assuming the CFL restriction (3.19), the relaxation parameters may vary from one interface to another. For convenience in what follows and to emphasize the admissible local choice of the parameters, we rewrite $\mathbf{W}_i^{n+1,-}$ arguing the well-known formalism introduced by Harten, Lax, and van Leer [21]:

$$\mathbf{W}_i^{n+1,-} = \frac{1}{2} \left(\bar{\mathbf{W}}_R(\mathbf{W}_{i-1}^n, \mathbf{W}_i^n) + \bar{\mathbf{W}}_L(\mathbf{W}_i^n, \mathbf{W}_{i+1}^n) \right), \tag{3.20}$$

where

$$\begin{aligned} \bar{\mathbf{W}}_L(\mathbf{W}_L, \mathbf{W}_R) &= \frac{2\Delta t}{\Delta x} \int_{-\frac{\Delta x}{2\Delta t}}^{0} \mathbf{W}_r(\xi;\ \mathbf{W}_L, \mathbf{W}_R)\, d\xi, \\ &= \mathbf{W}_L - \frac{2\Delta t}{\Delta x} \left(\mathbf{F}(\mathbf{W}_r(0;\ \mathbf{W}_L, \mathbf{W}_R)) - \mathbf{F}(\mathbf{W}_L) \right), \end{aligned} \tag{3.21}$$

and

$$\begin{aligned} \bar{\mathbf{W}}_R(\mathbf{W}_L, \mathbf{W}_R) &= \frac{2\Delta t}{\Delta x} \int_{0}^{\frac{\Delta x}{2\Delta t}} \mathbf{W}_r(\xi;\ \mathbf{W}_L, \mathbf{W}_R)\, d\xi, \\ &= \mathbf{W}_R - \frac{2\Delta t}{\Delta x} \left(\mathbf{F}(\mathbf{W}_R) - \mathbf{F}(\mathbf{W}_r(0;\ \mathbf{W}_L, \mathbf{W}_R)) \right). \end{aligned} \tag{3.22}$$

The function $\mathbf{W}_r(.;\ \mathbf{W}_L, \mathbf{W}_R)$ denotes the solution of the Riemann problem for (3.18), where the initial data is prescribed by (3.3).

As soon as $\mathbf{W}_L := \mathbf{W}(\mathbf{w}_L)$ and $\mathbf{W}_R := \mathbf{W}(\mathbf{w}_R)$ are defined from the equilibrium states $\mathbf{w}_L$ and $\mathbf{w}_R$, it is crucial to notice from now on the following identities:

$$\begin{aligned} \mathbf{F}(\mathbf{W}_L)|_{[h,\, hu]} &= \mathbf{f}(\mathbf{w}_L), \\ \mathbf{F}(\mathbf{W}_R)|_{[h,\, hu]} &= \mathbf{f}(\mathbf{w}_R), \end{aligned}$$

where the notation $\mathbf{F}(.)|_{[h,\, hu]}$ denotes the restriction of $\mathbf{F}$ to the component $(h,\, hu)$.

The second step of the scheme is devoted to the relaxation procedure. At time $t = t^n + \Delta t$, we define the updated approximate equilibrium solution $\mathbf{w}^{n+1}(x)$ as follows:

$$\mathbf{w}^{n+1}(x) = {}^t\Big(h_i^{n+1,-}, (hu)_i^{n+1,-} \Big), \quad x \in (x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}), \tag{3.23}$$

and we set $\Sigma_i^{n+1} = \sqrt{g h_i^{n+1}}$ and $U_i^{n+1} = (hu)_i^{n+1} / h_i^{n+1}$.

In fact, this second step amounts to solving the system

$$\partial_t \mathbf{W} = \mu \mathbf{R}(\mathbf{W}),$$

with the piecewise constant approximation $\mathbf{W}_i^{n+1,-}$ as initial data while μ tends to infinity. The derivation of the relaxation scheme is thus achieved.

In fact, the numerical relaxation flux function we have just derived exactly coincides with the VFRoe numerical flux function given by (2.5)–(2.9). Involving the classical framework of the finite volume method, the described relaxation scheme summarizes as follows:

$$\mathbf{w}_i^{n+1} = \mathbf{w}_i^n - \frac{\Delta t}{\Delta x}\left(\mathbf{f}_{i+\frac{1}{2}}^n - \mathbf{f}_{i-\frac{1}{2}}^n\right), \tag{3.24}$$

where the numerical flux function is defined by

$$\begin{aligned} \mathbf{f}_{i+\frac{1}{2}}^n &= \mathbf{f}(\mathbf{w}_i^n, \mathbf{w}_{i+1}^n), \\ &= \mathbf{F}\left(\mathbf{W}_r(0;\ \mathbf{W}(\mathbf{w}_i^n), \mathbf{W}(\mathbf{w}_{i+1}^n))\right)|_{[h,hu]}, \end{aligned} \tag{3.25}$$

with $\mathbf{W}_i^n = \mathbf{W}(\mathbf{w}_i^n)$ defined according to the equilibrium, i.e., $\Sigma_i^n = c_i^n$ and $U_i^n = u_i^n$.

Now, it is clear that both the relaxation scheme (3.24)–(3.25) and the VFRoe scheme (2.5)–(2.9) involve the same numerical flux function in the form:

$$\mathbf{f}(\mathbf{w}_L, \mathbf{w}_R) = \begin{pmatrix} (c^*)^2 u^*/g \\ (c^*)^2(2(u^*)^2 + (c^*)^2)/(2g) \end{pmatrix}, \tag{3.26}$$

where

$$c^* = \begin{cases} c_L & \text{if } \lambda^- > 0, \\ \dfrac{c_L + c_R}{2} - \dfrac{1}{4}(u_R - u_L) & \text{if } \lambda^- < 0 < \lambda^+, \\ c_R & \text{if } \lambda^+ < 0, \end{cases}$$

$$u^* = \begin{cases} u_L & \text{if } \lambda^- > 0, \\ \dfrac{u_L + u_R}{2} - (c_R - c_L) & \text{if } \lambda^- < 0 < \lambda^+, \\ u_R & \text{if } \lambda^+ < 0. \end{cases}$$

In fact, the two schemes may differ in the evaluation of the eigenvalues $\lambda^{\pm}$. In this sense, the VFRoe scheme (2.5)–(2.9) is closed when enforcing the linearization $\tilde{u} = \bar{u}$ and $\tilde{c} = \bar{c}$. We have thus proposed a relaxation interpretation of the well-known VFRoe scheme. We conclude by establishing the following expected robustness result.

THEOREM 3.3. *Assume that $\mathbf{w}_i^n \in \Omega$ for all $i \in \mathbb{Z}$, and assume that the eigenvalues $\lambda_{i+\frac{1}{2}}^{\pm}$ are evaluated according to the depth nonnegativity Lemma* 3.2. *Under the CFL condition* (3.19)*, the relaxation scheme* (3.24)–(3.25)*, or equivalently the VFRoe scheme* (2.5)–(2.9)*, preserves the nonnegativity of h: $h_i^{n+1} \geq 0$ for all $i \in \mathbb{Z}$. In addition, the scheme does not involve blow-up when dry areas are encountered.*

Proof. Since Lemma 3.2 is satisfied, the intermediate states, involved in the Riemann problem solution at each interface, preserve the nonnegativity of the depth. Then $h^h(x, t^n + \Delta t) \geq 0$ for all $x \in \mathbb{R}$, while we have

$$h_i^{n+1} = \frac{1}{\Delta x}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} h^h(x, t^n + \Delta t)dx.$$

We immediately deduce that $h_i^{n+1} \geq 0$. ∎

3.3. A class of relaxation model. The characterization of the relaxation model (3.1) was dictated by the definition of the VFRoe scheme. Following the same strategy, many other relaxation models can be considered. For instance, the reader is referred to [7, 3, 12] where Suliciu relaxation schemes are derived. These schemes are accurate and stable but turn out to be, sometimes, sophisticated in their implementation.

The aim of the present paper is to derive schemes which must be easy to implement, like the VFRoe method. Following the same idea, other relaxation methods involving easy linearization can be proposed. For instance, let us consider the following relaxation model:

$$\begin{cases} \partial_t h + \partial_x(\dfrac{\Sigma^2}{g}U) = 0, \quad t>0,\ x\in\mathbb{R}, \\ \partial_t(hu) + \partial_x\left(\dfrac{\Sigma^2}{2g}(2U^2+\Sigma^2)\right) = 0, \\ \partial_t\Sigma + \dfrac{\lambda^+-\lambda^-}{\alpha+\beta}\partial_x U + \dfrac{\alpha\lambda^- + \beta\lambda^+}{\alpha+\beta}\partial_x\Sigma = \mu(c-\Sigma), \\ \partial_t U + \dfrac{\alpha\lambda^+ + \beta\lambda^-}{\alpha+\beta}\partial_x U - \dfrac{\alpha\beta}{\alpha+\beta}(\lambda^+-\lambda^-)\partial_x\Sigma = \mu(u-U), \end{cases} \tag{3.27}$$

where α and β are positive parameters to be fixed. This model is nothing but an extension of (3.1). Indeed, (3.27) coincides with (3.1) as soon as $\alpha=\beta=2$. Considering physical applications, a suitable choice of the parameters α and β should give more accurate simulations.

We skip the algebra analysis of this system, which turns out to be easy (see also the above section). After straightforward computations, the resulting scheme reads in the form (3.24), where the numerical flux function is given by (3.26). Only the definition of (u^*, c^*) has changed and now are given by the following:

$$c^* = \begin{cases} c_L & \text{if } \lambda^- > 0, \\ \dfrac{\alpha c_L + \beta c_R}{\alpha+\beta} - \dfrac{1}{\alpha+\beta}(u_R-u_L) & \text{if } \lambda^- < 0 < \lambda^+, \\ c_R & \text{if } \lambda^+ < 0, \end{cases}$$

$$u^* = \begin{cases} u_L & \text{if } \lambda^- > 0, \\ \dfrac{\alpha u_L + \beta u_R}{\alpha+\beta} - \dfrac{\alpha\beta}{\alpha+\beta}(c_R-c_L) & \text{if } \lambda^- < 0 < \lambda^+, \\ u_R & \text{if } \lambda^+ < 0. \end{cases}$$

Once again, we can establish that the obtained updated depth h_i^{n+1} remains nonnegative as soon as the relaxation wave speeds λ^- and λ^+ are judiciously chosen.

4. High order extension. We next turn to considering an extension of the above first order schemes to increase the order of accuracy. To access such an issue, we adopt the celebrated MUSCL method introduced by van Leer [29]. This technique is based on a linear (or higher order) reconstruction instead of a piecewise constant approximation. This reconstruction allows us to evaluate states $\mathbf{w}_{i+1,l}$ and $\mathbf{w}_{i,r}$ at each side of the interface located at $x_{i+\frac{1}{2}}$. The MUSCL scheme thus reads as follows:

$$\mathbf{w}_i^{n+1} = \mathbf{w}_i^n - \frac{\Delta t}{\Delta x}(\mathbf{f}_{i+\frac{1}{2}} - \mathbf{f}_{i-\frac{1}{2}}), \tag{4.1}$$

but, at this time, the numerical flux function involves the high order evaluated states at interfaces:

$$\mathbf{f}_{i+\frac{1}{2}} = \mathbf{f}(\mathbf{w}_{i,r}, \mathbf{w}_{i+1,l}). \tag{4.2}$$

In order to enforce the required robustness of this modified scheme, several approaches can be considered. The most usual (see [1, 7, 16, 30]) implies *conservative* reconstruction in the following sense:

$$\frac{1}{2}(\mathbf{w}_{i,l} + \mathbf{w}_{i,r}) = \mathbf{w}_i^n. \tag{4.3}$$

As soon as the reconstructed states $\mathbf{w}_{i,l}$ and $\mathbf{w}_{i,r}$ stay in Ω, involving a relevant half-CFL restriction [7, 30], the updated states preserve the nonnegativity of the depth: $h_i^{n+1} \geq 0$. The assumption (4.3) appears as natural within the framework of a second-order MUSCL scheme, where linear reconstruction is proposed. Now, when high order reconstruction is suggested, the assumption (4.3) is too restrictive and cannot be assumed. After an idea introduced by Perthame and Shu [37] (see also some recent works [4, 5, 6]), relevant interface states can be evaluated in order to preserve the expected depth nonnegativity when (4.3) is violated.

We assume that the reconstructed states $\mathbf{w}_{i,l}$ and $\mathbf{w}_{i,r}$ belong to Ω and satisfy the required order of accuracy. Following [4], we introduce an intermediate state denoted $\mathbf{w}_i^*$ defined by

$$\frac{1}{3}(\mathbf{w}_{i,l} + \mathbf{w}_i^* + \mathbf{w}_{i,r}) = \mathbf{w}_i^n.$$

The reconstruction, or equivalently the characterization of $\mathbf{w}_{i,r}$ and $\mathbf{w}_{i,l}$, for all $i \in \mathbb{Z}$, must be done to enforce $\mathbf{w}_{i,l} \in \Omega$, $\mathbf{w}_{i,r} \in \Omega$, but also $\mathbf{w}_i^* \in \Omega$. Involving a relevant CFL condition, formally divided by three, the MUSCL scheme (4.1)–(4.2) preserves the nonnegativity of h_i^{n+1}. In other words, this procedure enforces the expected robustness property independently of the reconstruction process. The robustness result, stated in [4], can be applied in the present work. Thus we have the following.

THEOREM 4.1. *Assume that the reconstruction is such that* $\mathbf{w}_{i,l}$, $\mathbf{w}_i^*$, *and* $\mathbf{w}_{i,r}$ *belong to* Ω *for all* $i \in \mathbb{Z}$. *Assume the CFL-like condition:*

$$\frac{\Delta t}{\Delta x} \max\left(|\lambda^{\pm}(\mathbf{w}_{i,r}, \mathbf{w}_{i+1,l})|, |\lambda^{\pm}(\mathbf{w}_{i,l}, \mathbf{w}_i^*)|, |\lambda^{\pm}(\mathbf{w}_i^*, \mathbf{w}_{i,r})|\right) \leq \frac{1}{6}, \tag{4.4}$$

where $\lambda^{\pm}(\mathbf{w}_L, \mathbf{w}_R) := \bar{u}(\mathbf{w}_L, \mathbf{w}_R) \pm \bar{c}(\mathbf{w}_L, \mathbf{w}_R)$ *define the egeinvalues evaluated according to the depth nonnegativity Lemma* 3.2. *Then the updated states* $(\mathbf{w}_i^{n+1})_{i\in\mathbb{Z}}$, *defined by the MUSCL scheme* (4.1)–(4.2), *belong to* Ω.

Proof. Involving the robust first-order VFRoe scheme (2.5)–(2.9), the updated reconstructed states $\mathbf{w}_{i,l}$, $\mathbf{w}_{i,r}$, and $\mathbf{w}_i^*$ can be written as follows:

$$\begin{aligned}
\mathbf{w}_{i,l}^{n+1} &= \mathbf{w}_{i,l} - \frac{\Delta t}{\Delta x/3}\left(\mathbf{f}(\mathbf{w}_{i,l}, \mathbf{w}_i^*) - \mathbf{f}(\mathbf{w}_{i-1,r}, \mathbf{w}_{i,l})\right),\\
\mathbf{w}_i^{*,n+1} &= \mathbf{w}_i^* - \frac{\Delta t}{\Delta x/3}\left(\mathbf{f}(\mathbf{w}_i^*, \mathbf{w}_{i,r}) - \mathbf{f}(\mathbf{w}_{i,l}, \mathbf{w}_i^*)\right),\\
\mathbf{w}_{i,r}^{n+1} &= \mathbf{w}_{i,r} - \frac{\Delta t}{\Delta x/3}\left(\mathbf{f}(\mathbf{w}_{i,r}, \mathbf{w}_{i+1,l}) - \mathbf{f}(\mathbf{w}_i^*, \mathbf{w}_{i,r})\right).
\end{aligned}$$

Under the CFL restriction (4.4), we immediately deduce from Theorem 3.3 that the three updated states belong to Ω. We have just to note that the state $\mathbf{w}_i^{n+1}$, defined by (4.1)–(4.2), writes as follows:

$$\mathbf{w}_i^{n+1} = \frac{1}{3}(\mathbf{w}_{i,l}^{n+1} + \mathbf{w}_i^{*,n+1} + \mathbf{w}_{i,r}^{n+1}),$$

and the proof is completed. ∎

From a practical point of view, the reconstructed states at the interface read as follows:

$$\mathbf{w}_{i,l} = \mathbf{w}_i^n + \Delta_i^- \quad \text{and} \quad \mathbf{w}_{i,r} = \mathbf{w}_i^n + \Delta_i^+, \tag{4.5}$$

where $\Delta_i^{\pm}$ are the reconstruction increments. When considering conservative reconstructions, we have $\Delta_i^- + \Delta_i^+ = 0$. Now, as soon as $\Delta_i^- + \Delta_i^+ \neq 0$, we have to introduce the intermediate state $\mathbf{w}_i^*$, defined by:

$$\mathbf{w}_i^* = \mathbf{w}_i^n - (\Delta_i^- + \Delta_i^+).$$

Concerning the depth evaluation, the increments are modified to enforce the following restrictions:

$$h_i^n + (\Delta h)_i^- \geq 0, \quad h_i^n + (\Delta h)_i^+ \geq 0, \quad h_i^n - ((\Delta h)_i^- + (\Delta h)_i^+) \geq 0. \tag{4.6}$$

To illustrate this new restriction, we propose to introduce the following modified limitation:

$$(\Delta h)_i^{-,lim} = \theta \max\left((\Delta h)_i^-, -h_i^n\right), \quad (\Delta h)_i^{+,lim} = \theta \max\left((\Delta h)_i^+, -h_i^n\right), \tag{4.7}$$

where we have set

$$\theta = \begin{cases} 1 & \text{if } \max\left((\Delta h)_i^-, -h_i^n\right) + \max\left((\Delta h)_i^+, -h_i^n\right) \leq 0, \\ \min\left(1, \frac{h_i^n}{\max\left((\Delta h)_i^-, -h_i^n\right) + \max\left((\Delta h)_i^+, -h_i^n\right)}\right) & \text{otherwise.} \end{cases}$$

Once this step is performed, the high order reconstructed variables are defined by (4.5) but for the limited increments $\Delta^{-,lim}$ and $\Delta^{+,lim}$. This additional limitation process is more sophisticated than standard reconstruction, but it satisfies the condition (4.6). Of course other slope limitations can be suggested.

To conclude this brief robust MUSCL scheme presentation, let us note that the additional limitation preserves the order of the accuracy of the interface states as soon as the exact solution at time t^n, $\mathbf{w}(x, t^n)$ is smooth enough.

To illustrate our purpose, a fourth-order MUSCL reconstruction is proposed in the present work and used in the numerical investigations of section 6. Following [43], we set, for an arbitrary variable q,

$$(\Delta q)_i^- = -\frac{1}{6}(2\Delta^* \bar{q}_{i-\frac{1}{2}} + \Delta^* \tilde{q}_{i+\frac{1}{2}}), \qquad (\Delta q)_i^+ = \frac{1}{6}(\Delta^* \bar{q}_{i-\frac{1}{2}} + 2\Delta^* \tilde{q}_{i+\frac{1}{2}}), \tag{4.8}$$

where

$$\begin{aligned} \Delta^* \bar{q}_{i-\frac{1}{2}} &= \operatorname{minmod}\left(\Delta^* q_{i-\frac{1}{2}}, b\Delta^* q_{i+\frac{1}{2}}\right), \\ \Delta^* \tilde{q}_{i+\frac{1}{2}} &= \operatorname{minmod}\left(\Delta^* q_{i+\frac{1}{2}}, b\Delta^* q_{i-\frac{1}{2}}\right), \end{aligned} \tag{4.9}$$

and

$$
\begin{aligned}
&\Delta^* q_{i+\frac{1}{2}} = \Delta q_{i+\frac{1}{2}} - \tfrac{1}{6}\Delta^3 \bar{q}_{i+\frac{1}{2}},\\
&\Delta^3 \bar{q}_{i+\frac{1}{2}} = \Delta \bar{q}_{i-\frac{1}{2}} - 2\Delta \bar{q}_{i+\frac{1}{2}} + \Delta \bar{q}_{i+\frac{3}{2}},\\
&\Delta \bar{q}_{i-\frac{1}{2}} = \operatorname{minmod}\left(\Delta q_{i-\frac{1}{2}}, c\Delta q_{i+\frac{1}{2}}, c\Delta q_{i+\frac{3}{2}}\right),\\
&\Delta \bar{q}_{i+\frac{1}{2}} = \operatorname{minmod}\left(\Delta q_{i+\frac{1}{2}}, c\Delta q_{i+\frac{3}{2}}, c\Delta q_{i-\frac{1}{2}}\right),\\
&\Delta \bar{q}_{i+\frac{3}{2}} = \operatorname{minmod}\left(\Delta q_{i+\frac{3}{2}}, c\Delta q_{i-\frac{1}{2}}, c\Delta q_{i+\frac{1}{2}}\right),
\end{aligned}
\tag{4.10}
$$

with

$$\Delta q_{i+\frac{1}{2}} = q_{i+1} - q_i,$$

and

$$1 < b \leq 4, \qquad c = 2.$$

Concerning the numerical experiments proposed in the related section, the parameter b will be set to 4. It is worth mentioning that this fourth-order reconstruction reduces to a classical third order scheme when replacing $\Delta^* q_{i+\frac{1}{2}}$ by $\Delta q_{i+\frac{1}{2}}$ in (4.9), as developed in [43]. Considering the shallow water system (1.1), one has to reconstuct conservative variables h and hu following (4.8), (4.9), and (4.10) and to apply the modified limitation (4.7) only on h. Note that, when considering a well-balancing discretization of system (1.1), one has also to reconstruct and apply (4.7) on $H = h + d$ to preserve steady states at rest, as developed in the next section.

5. Well-balancing discretization for source terms. In this section we focus on the shallow water system with a bed slope source term (1.1).

5.1. Hydrostatic reconstruction. We recall here how to obtain a high order well-balanced scheme satisfying the preservation of steady states "at rest," following the lines of the method proposed first in [1] and extended to a higher order of accuracy in [36].

The first step is to build a high order reconstruction of the values at each side of the mesh interfaces. We find various ways to achieve such accuracy, including MUSCL reconstructions [43] (see also the above section) or ENO/WENO polynomial reconstructions [42, 36]. Note that the main limitation of all these methods concerns the ability to deal with the occurrence of dry areas. In the following, we use the fourth order MUSCL reconstruction (4.8), together with the modified limitation (4.7) to ensure the preservation of the water depth positivity within the reconstruction. Considering the cell i, we compute first reconstructions $\mathbf{w}_{i,r}$ and $\mathbf{w}_{i,l}$, respectively, located at $x_{i+\frac{1}{2}-}$ and $x_{i-\frac{1}{2}+}$. The values of $H_{i,l}$ and $H_{i,r}$ where $H = h + d$ are also reconstructed, and we deduce that the reconstructions of the topography $d_{i,l} = H_{i,l} - h_{i,l}$ and $d_{i,r} = H_{i,r} - h_{i,r}$. It ensures that if $u_i = 0$ and $H_i = H_{i+1}$ for all i, then $u_{i,l} = u_{i,r} = 0$ and $H_{i,l} = H_{i,r} = H_i$ for all i. Note that, even if the VFRoe solver introduced in section 1 is defined with nonconservative variables, numerical investigations have shown that the choice of reconstructing conservative variables gives better results in practice.

In a second step, we perform the "hydrostatic reconstruction" of the values at each side of the mesh interfaces, taking into account the variations of the bottom and the balance obtained static flows. Interface topography values $d_{i+\frac{1}{2}}$ are defined as follows:

$$d_{i+\frac{1}{2}} = \max\left(d_{i,r}, d_{i+1,l}\right). \tag{5.1}$$

Then, the reconstruction of the water height on each side of the considered interface is defined as follows:

$$h_{i+\frac{1}{2}-} = \max(0, h_{i,r} + d_{i,r} - d_{i+\frac{1}{2}}), \quad h_{i+\frac{1}{2}+} = \max(0, h_{i+1,l} + d_{i+1,l} - d_{i+\frac{1}{2}}), \tag{5.2}$$

and we deduce from it the complete reconstructed values on each side of the interface as follows:

$$\mathbf{w}_{i+\frac{1}{2}-} = \begin{pmatrix} h_{i+\frac{1}{2}-} \\ h_{i+\frac{1}{2}-}\, u_{i,r} \end{pmatrix}, \quad \mathbf{w}_{i+\frac{1}{2}+} = \begin{pmatrix} h_{i+\frac{1}{2}+} \\ h_{i+\frac{1}{2}+}\, u_{i+1,l} \end{pmatrix}. \tag{5.3}$$

Then, using these new reconstructed values to compute the interface solution within a VFRoe formalism and with the notations introduced in section 2, the numerical flux $\mathbf{f}^*_{i+\frac{1}{2},j}$ is defined as follows:

$$\mathbf{f}^*_{i+\frac{1}{2}} = \mathbf{f}\Big(\mathcal{W}(\mathbf{u}^*(0; \mathcal{U}(\mathbf{w}_{i+\frac{1}{2}-}), \mathcal{U}(\mathbf{w}_{i+\frac{1}{2}+})))\Big). \tag{5.4}$$

Motivated by balancing requirements for static flows, the source term $\mathbf{S}_i$ is discretized and distributed to the cell interfaces using both high order and hydrostatic reconstructed values of the water height:

$$\mathbf{S}_i = \mathbf{S}_{i+\frac{1}{2}-} + \mathbf{S}_{i-\frac{1}{2}+} = \begin{pmatrix} 0 \\ \frac{g}{2} h^2_{i+\frac{1}{2}-} - \frac{g}{2} h^2_{i,r} \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{g}{2} h^2_{i,l} - \frac{g}{2} h^2_{i-\frac{1}{2}+} \end{pmatrix}. \tag{5.5}$$

A centered source term $S_{c,i}$ is added to preserve consistency and well-balancing discretization:

$$\mathbf{S}_{c,i} = \begin{pmatrix} 0 \\ g\dfrac{h_{i,l} + h_{i,r}}{2}(d_{i,l} - d_{i,r}) \end{pmatrix}. \tag{5.6}$$

It leads to the following semidiscrete formulation:

$$\frac{d}{dt}\mathbf{w}_i(t) + \frac{1}{\Delta x}\Big(\mathbf{f}^-_{i+\frac{1}{2}} - \mathbf{f}^+_{i-\frac{1}{2}}\Big) = \mathbf{S}_{c,i}, \tag{5.7}$$

with left and right numerical fluxes through the mesh interfaces defined as follows:

$$\begin{aligned} \mathbf{f}^-_{i+\frac{1}{2}} &= \mathbf{f}^*_{i+\frac{1}{2}} + \mathbf{S}_{i+\frac{1}{2}-} \\ &= \mathbf{f}\Big(\mathcal{W}(\mathbf{u}^*(0; \mathcal{U}(\mathbf{w}_{i+\frac{1}{2}-}), \mathcal{U}(\mathbf{w}_{i+\frac{1}{2}+})))\Big) + \begin{pmatrix} 0 \\ \frac{g}{2} h^2_{i,r} - \frac{g}{2} h^2_{i+\frac{1}{2}-} \end{pmatrix}, \end{aligned} \tag{5.8}$$

$$\begin{aligned} \mathbf{f}^+_{i+\frac{1}{2}} &= \mathbf{f}^*_{i+\frac{1}{2}} + \mathbf{S}_{i+\frac{1}{2}+} \\ &= \mathbf{f}\Big(\mathcal{W}(\mathbf{u}^*(0; \mathcal{U}(\mathbf{w}_{i+\frac{1}{2}-}), \mathcal{U}(\mathbf{w}_{i+\frac{1}{2}+})))\Big) + \begin{pmatrix} 0 \\ \frac{g}{2} h^2_{i+1,l} - \frac{g}{2} h^2_{i+\frac{1}{2}+} \end{pmatrix}. \end{aligned} \tag{5.9}$$

5.2. Well-balancing property for steady states. We briefly recall the proof for the well-balanced property. Assuming that $H = h + d$ is constant at time t and that $u = 0$, we have $H_{i,r} = H_{i,l}$ and

$$h_{i+\frac{1}{2}-} = \max(0, H_{i,r} - d_{i+\frac{1}{2}}) = \max(0, H_{i,l} - d_{i+\frac{1}{2}}) = h_{i+\frac{1}{2}+}, \tag{5.10}$$

and by construction,

$$\mathbf{w}_{i+\frac{1}{2}-} = \begin{pmatrix} h_{i+\frac{1}{2}-} \\ 0 \end{pmatrix} = \begin{pmatrix} h_{i+\frac{1}{2}+} \\ 0 \end{pmatrix} = \mathbf{w}_{i+\frac{1}{2}+}. \tag{5.11}$$

Considering (5.4) and (5.8), we obtain

$$\mathbf{f}^-_{i+\frac{1}{2}} = \begin{pmatrix} 0 \\ \frac{g}{2} h^2_{i,r} \end{pmatrix} \quad \text{and} \quad \mathbf{f}^+_{i-\frac{1}{2}} = \begin{pmatrix} 0 \\ \frac{g}{2} h^2_{i,l} \end{pmatrix}, \tag{5.12}$$

and the semidiscrete formulation (5.7) leads to the following identity:

$$\Delta x \frac{d}{dt} \begin{pmatrix} h_i(t) \\ (h_i u_i)(t) \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{g}{2} h^2_{i,r} \end{pmatrix} - \begin{pmatrix} 0 \\ \frac{g}{2} h^2_{i,l} \end{pmatrix} = \begin{pmatrix} 0 \\ g \frac{h_{i,l} + h_{i,r}}{2} (d_{i,l} - d_{i,r}) \end{pmatrix}. \tag{5.13}$$

Therefore, we have

$$\Delta x \frac{d}{dt}(h_i u_i)(t) = g \frac{h_{i,l} + h_{i,r}}{2}(h_{i,l} - h_{i,r}) + g \frac{h_{i,l} + h_{i,r}}{2}\Big((H_{i,l} - h_{i,l}) - (H_{i,r} - h_{i,r})\Big), \tag{5.14}$$

and

$$\Delta x \frac{d}{dt}(h_i u_i)(t) = g \frac{h_{i,l} + h_{i,r}}{2}(H_{i,r} - H_{i,l}) = 0. \tag{5.15}$$

The expected well-balanced property is thus satisfied.

6. Numerical results. In the following test cases, we use a fourth order accurate scheme, based on the stabilized VFRoe scheme introduced in section 2. The fourth order MUSCL reconstruction (4.8), together with the new modified limitation procedure (4.7), are considered. The hydrostatic reconstruction recalled in section 5 is used to account for topography variations in a well-balanced way. The semidiscrete discretization (5.7) is replaced by a third order TVD Runge–Kutta time discretization [38]. To assess the improvements of accuracy obtained with this high order reconstruction, we begin in the first two cases to perform a comparison, with the present scheme and classical first and second order VFRoe schemes. In all of the following cases, we show only numerical results obtained with the present fourth order scheme in situations involving steady states and occurrences of dry states. The CFL are, respectively, set to 0.5, 0.25, and 0.15 for first, second, and fourth order reconstructions, in agreement with the stability analysis performed in the above sections. Let us emphasize that, according to the work of Masella, Faille, and Gallouët, [35] (see also [20]), the present scheme is modified in order to handle the sonic rarefaction wave invoking the standard entropy correction.

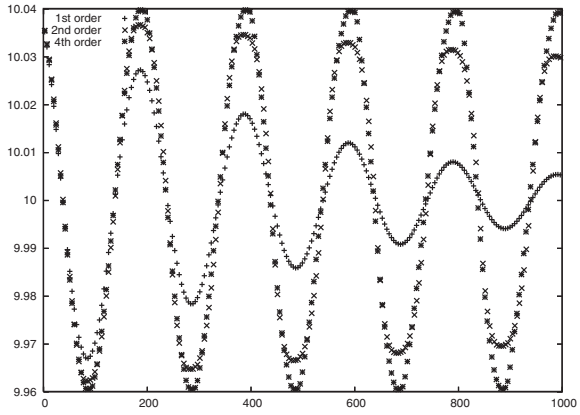

FIG. 6.1. *Sinusoidal wave propagation in shallow water. Free surface evolution: Comparison between first, second, and fouth order schemes for* 250 *cells.*

6.1. Sinusoidal wave propagation in shallow water. This first test describes a sinusoidal wave propagating in a channel with a horizontal bottom, taken from [22]. The channel is $A = 1000\,m$ long, and the still water depth is $H_0 = 10\,m$. The period and amplitude of the sinusoidal wave are $20.193\,s$ and $0.04\,m$, respectively. In this case, the ratio $A/H_0 \approx 0.016$ is close to the limit of shallow water conditions. Note that, this test differs from the original one introduced in [22], since we used an absorbing boundary [41] at the right boundary, in order to let the incoming wave exit the computational domain without reflexions. The channel is divided into 250 cells, and the propagation is performed, respectively, with schemes of first, second, and fourth order of accuracy to clearly exhibit the numerical diffusion induced by lower order methods. We obtain a steady-wave solution with a wavelength of $200\,m$, which is in agreement with the expected wavelength. As can be seen in Figure 6.1, the fourth order stabilized MUSCL reconstruction largely reduces the numerical diffusion observed with first and second order schemes, even with a small number of cells.

6.2. Subcritical flow over a bump. In this test, the bottom is defined as follows.

$$d(x) = \begin{cases} 0.2 - 0.05(x-10)^2 & \text{if } 8\,m < x < 12\,m, \\ 0 & \text{else,} \end{cases} \tag{6.1}$$

and we consider a channel of $25\,m$ long. The boundary conditions are $h_{out} = 2\,m$ and $Q_{in} = 4.42\,m^2/s$, corresponding to a subcritical flow. We obtain a stationary solution, shown in Figure 6.2. We observe in this figure the numerical results obtained with the fourth order scheme at time $T = 200\,s$ with 100 cells, superposed with the analytical solution. Since the results obtained with first, second, and fourth order schemes are similar once the convergence is achieved, we also show in Figure 6.3 a comparison between L^2-error time series for each order of accuracy. We can clearly observe the convergence toward the steady state at rest in the three cases. For scaling reasons, the transient part of the computation is partially truncated, and we show only in Figure 6.3 a time series of the L^2-error for values of time from $t = 50\,s$ and $t = 200\,s$.

Lastly, we show the L^2-errors table (see Table 6.1), computed for the water height, for various numbers of cells, at $t = 200\,s$. We can observe some improvements obtained with the fourth order reconstruction proposed. However, it is worth mentioning

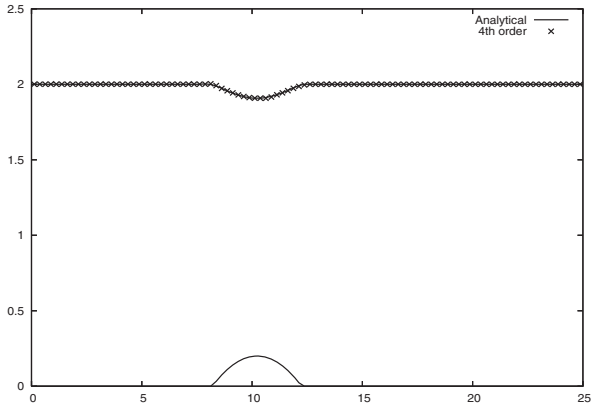

FIG. 6.2. *Subcritical flow over a bump: Water height at steady state.*

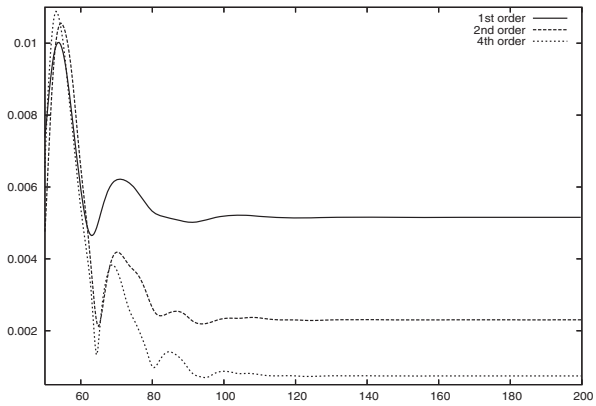

FIG. 6.3. *Subcritical flow over a bump: L^2-error time series, comparison between first, second, and fourth order.*

TABLE 6.1

Cells	First order scheme	Second order scheme	Fourth order scheme
50	5.155E-3	1.700E-3	7.422E-4
100	2.704E-3	6.062E-4	1.961E-4
200	1.383E-3	1.759E-4	6.155E-5
400	6.993E-4	5.730E-5	2.103E-5
800	3.516E-4	1.950E-5	7.242E-6
1600	1.763E-4	6.787E-6	1.562E-6

that we have only performed here a second order discretization of the source term, which lowered the overall accuracy. A better accuracy, can of course, be obtained using a stabilized high order reconstruction together with a high order quadrature discretization rule for the source term, like the one introduced in the recent work [36].

We conclude the analysis of the present numerical experiment with a comparison of the efficiency of both second and fourth order methods. To consider a similar L^2-error at time $t = 200\,s$, we adopt the second order scheme with 1600 cells, while the fourth order method is used with 800 cells. Invoking the CPU time for these two tests, we obtain that the fourth scheme is twice as quick than the second order scheme.

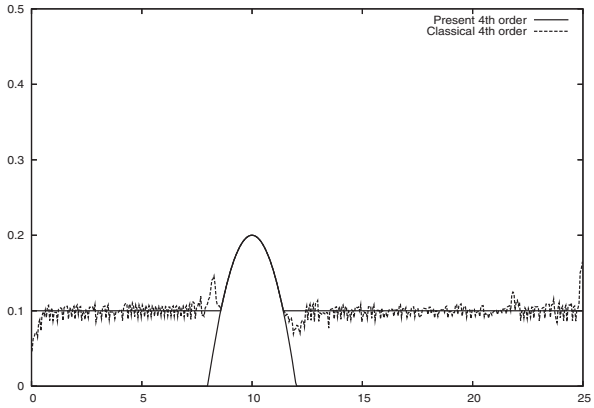

FIG. 6.4. *Flow at rest with a dry state: Water height. Comparison between stabilized fourth order and classical fourth order.*

6.3. Flow at rest with dry state. The initial condition of this test is a flow at rest, involving a dry area. We impose $h + d = \max(d, 0.15)\, m$ and $Q = 0\, m^2/s$ over the whole channel. The bottom variations are defined with (6.1). Numerical results, obtained with 100 cells at time $T = 200\, s$, are plotted in Figure 6.4 with and without the modified limitation (4.7) (respectively, in solid and dashed lines). We can clearly show that the stabilized fourth order scheme proposed here is well balanced and can handle dry areas, whereas large spurious oscillations are rapidly generated when stabilization (4.7) is not applied.

6.4. Drain on a nonflat bottom. The length of the channel and the topography of this test are the same as in the two previous cases. The initial condition here is a flow at rest, with a water height given by $h + d = 0.5\, m$ such that the bump is entirely submerged, and the initial discharge is set to $Q = 0\, m^2/s$ over the whole domain. For this more complex test, the whole domain is discretized into 500 cells. The left boundary is a solid-wall boundary (defined within a classical "mirror-state" procedure) and an outlet condition on a dry bed [13] at the right boundary. With such a configuration, the water progressively exits the computational domain, asymptotically converging towards a steady state composed of a state at rest on the left part of the bump, with $h + d = 0.2\, m$ and $Q = 0\, m^2/s$ and a dry state on the right side of the bump. Results obtained with the fourth order stabilized scheme are shown in Figure 6.5 for the water height and in Figure 6.6 for the discharge. These results are qualitatively similar to those obtained with the second order scheme introduced in [13]. We emphasize that the use of our stabilized fourth-order reconstruction (4.7) allows accurate results even during the draining part of the simulation, and that we accurately converge toward the expected steady state. We can observe the generation of tiny oscillations near the discontinuities.

6.5. Vacuum occurrence by a double rarefaction wave over a step. In this last test, introduced in [13], we do not study the convergence toward a steady state, but we rather focus on the ability of the proposed stabilized fourth order well-balanced scheme to deal with dry areas over a discontinuous bottom. The topography is different from previous tests and is defined as follows:

$$d(x) = \begin{cases} 1 & \text{if } 8.33\, m < x < 12.5\, m, \\ 0 & \text{else,} \end{cases} \tag{6.2}$$

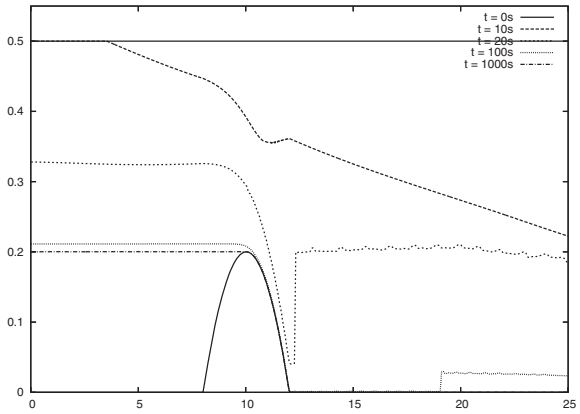

FIG. 6.5. *Drain on a nonflat bottom: Water height.*

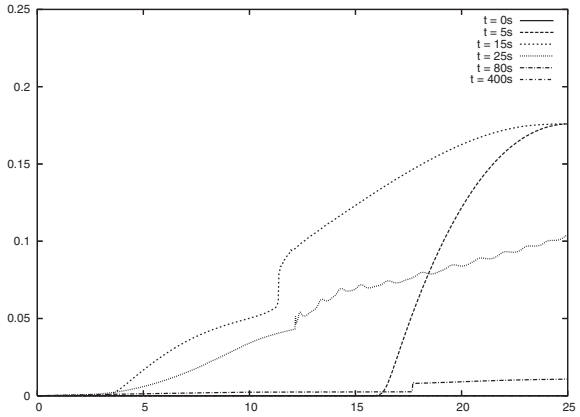

FIG. 6.6. *Drain on a nonflat bottom: Discharge.*

whereas the channel length is still 25 m, divided into 500 cells. The initial condition is set to $h + d = 10\ m$ and the initial discharge is defined as follows:

$$Q(x) = \begin{cases} -350\ m^2/s & \text{if } x < 50/3\ m, \\ 350\ m^2/s & \text{otherwise.} \end{cases} \tag{6.3}$$

Results at several times are shown in Figure 6.7 and Figure 6.8, respectively, for the free surface and the discharge, at time 0, 0.10, 0.30, 0.40, and 0.60 s. For the sake of clarity, we do not show the results obtained with first and second order schemes, but we stress that the results obtained here with our stabilized fourth order scheme are qualitatively very similar. The main difference lies in the occurrence of tiny oscillations near discontinuities for the water height, as observable in Figure 6.7.

6.6. The Carrier and Greenspan transient solution. We propose here to study one of the test cases developped in [34] to highlight the improvement obtained with the stabilized fourth order approach even in situations involving complex drying and flooding processes. The reader is referred to [34] or to the original paper [10] for a detailed description. Note that an analytical solution is provided in [10]. In this test, the initial water surface elevation is assumed to be depressed near the shoreline,

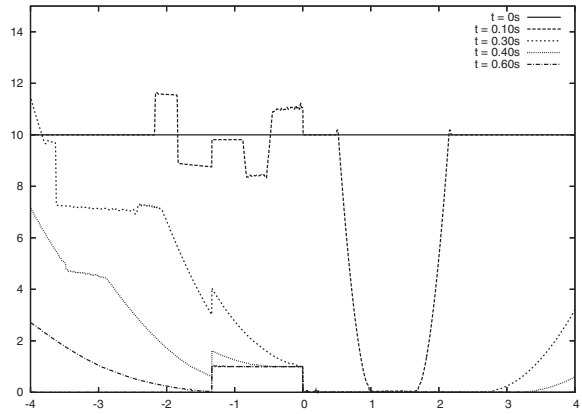

FIG. 6.7. *Vacuum occurrence by a double rarefaction wave over a step: Water height.*

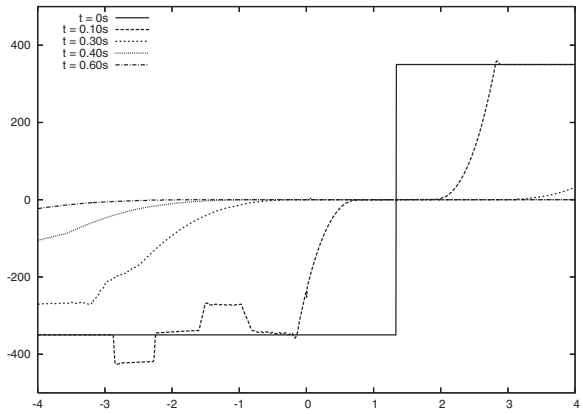

FIG. 6.8. *Vacuum occurrence by a double rarefaction wave over a step: Discharge.*

defined as the frontier between the fluid domain and the dry area, and the fluid held motionless. Then, the fluid is released at $t = 0$, and we mainly focus on the shoreline evolution. During the evolution, the shoreline rises above the mean sea level of value e, and then the water surface elevation asymptotically settles back to it. This initial condition is the lower curve on Figure 6.9. Let l be the typical length scale of this specific problem, and α be the beach slope. Nondimensional variables are defined as follows:

$$x^* = x/l, \quad h^* = h/(\alpha\, l), \quad u^* = u/\sqrt{g\alpha l}, \quad t^* = t/\sqrt{l/\alpha g}. \tag{6.4}$$

The bottom slope α is taken to be $1/50$; the results are presented here for $e = 0.1$, and the initial surface profile is imposed in the dimensional case, with the length scale $l = 20$ m. Note that the CFL is set to 0.7 for the first and second order schemes and to 0.3 for the stabilized fourth order scheme. Figure 6.9 shows comparisons between numerical results obtained with the stabilized fourth order scheme and the analytical solution for the surface elevation, at various values of time. These surface elevation profiles have been scaled with the parameter e. It is worth mentioning that the results introduced in [34] for this case and obtained with a second order well-balanced scheme were already very good. Thus a table of the comparison of L^2-errors obtained for various orders and several numbers of cells is provided at $t^* = 10$ (see Table 6.2).

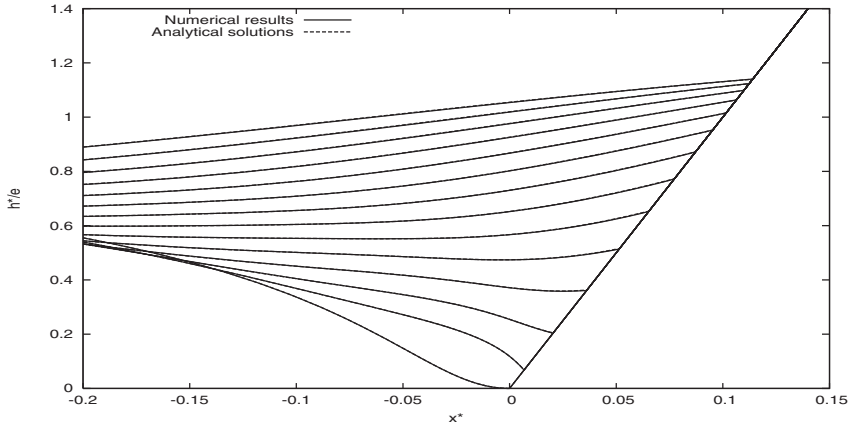

FIG. 6.9. *The Carrier and Greenspan transient solution. Comparison between numerical results (in solid lines) obtained with the fourth order scheme and analytical solutions (in dotted lines) for the surface elevation. Profiles of water height h^*/e are plotted versus x^*, for different values of time t^*, increasing from $t^* = 0$ (bottom curves) to $t^* = 1.4$ (top curves).*

TABLE 6.2

Cells	Second order scheme	Fourth order scheme
250	3.095E-3	5.912E-4
500	2.249E-3	2.151E-4
1000	2.058E-3	1.516E-4
2000	8.690E-4	1.223E-5

From Table 6.2 we observe that both schemes provide stability and good accuracy in the computation of the slow convergence toward the mean water level. The L^2-error computed for large values of time with the second order scheme seems to decrease very slowly with respect to the number of cells. This feature is slightly improved with the fourth order scheme. It appears from this test that the stabilized fourth order scheme provides more accuracy for large values of time.

6.7. Oblique 2D dam break on a dry bed. We study in this last test the evolution of a mound of water over a flat bottom, which is suddenly released from an initial position, generating a propagating front such that the line of water propagates with an inclination of 45° with respect to the boundaries of the computational domain. The base of this domain is a $[-0.5, 0.5] \times [-0.5, 0.5]$ square. The initial condition, shown in Figure 6.10, is defined as follows:

$$h(0, x, y) = \begin{cases} h_L \text{ for } x + y \leq 0, \\ 0 \text{ otherwise,} \end{cases} \tag{6.5}$$

and the initial discharge is set to zero. The purpose of this last test is to show the possibility of an extension of the stabilized fourth order scheme towards two-dimensional (2D) simulations. The results shown here have been obtained using 100 cells in each direction, and the CFL is set to 0.2. We focus only on the surface elevation computed on the central cross section orthogonal to the propagating front (the $x = y$ plane) which is compared to the one-dimensional (1D) analytical solution, assuming that the effects induced by the boundaries can be neglected for this section and for a small time of evolution. A comparison with the analytical solution in the direction

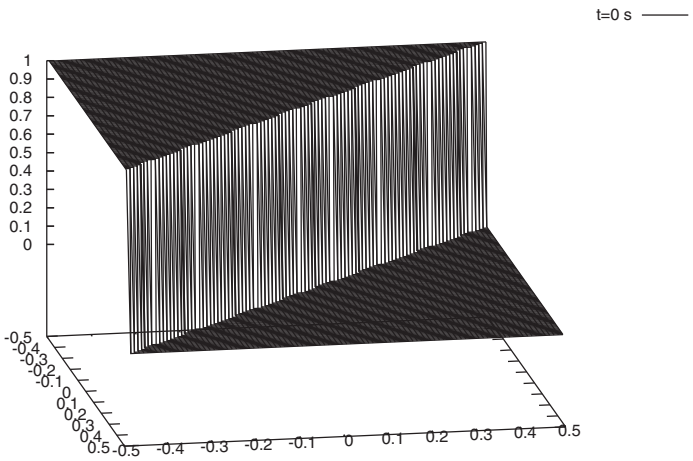

FIG. 6.10. *Oblique 2D dam break on a dry bed: Initial surface elevation.*

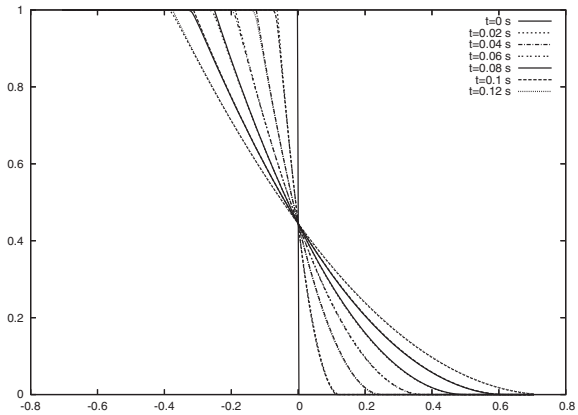

FIG. 6.11. *Oblique 2D dam break on a dry bed: Time evolution of the surface elevation in the central cross section and comparison with the analytical 1D solution.*

of propagation is reported on Figure 6.11. The exact and approximated solutions are superposed. To conclude this 2D test, after [5, 16], let us note that unstructured mesh extensions can be performed.

7. Conclusion. In the present work, we have derived a numerical relaxation scheme to approximate the weak solutions of the shallow water equation. The analysis of the obtained approximate Riemann solver allows us to prove the required robustness results. Indeed, involving a suitable CFL-like restriction, we establish the nonnegativity of the approximated depth. In addition, the considered relaxation scheme is able to deal with dry areas and no blow-up is developed. In fact, the relaxation scheme under consideration exactly coincides with one variant of the well-known VFRoe scheme. As a consequence, we have established that the VFRoe scheme does not produce a negative value of the water depth. This statement is obtained when involving a relevant upwind technique. Next, we have proposed high order extensions of this VFRoe scheme (second and fourth order) based on the van Leer slope reconstructions. Enforcing some suitable restriction in the gradient reconstruction

procedure, we have proved that the considered high order extensions, once again, preserve the nonnegativity of the depth. The final well-balanced scheme is obtained when assuming the hydrostatic reconstruction to approximate the geometrical source terms. Several 1D numerical experiments have shown the interest of the method. The extension towards 2D simulations is introduced and currently under investigation for realistic problems.

REFERENCES

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

[2] M. Baudin, C. Berthon, F. Coquel, R. Masson, and Q. H. Tran, *A relaxation method for two-phase flow models with hydrodynamic closure law*, Numer. Math., 99 (2005), pp. 411–440.

[3] C. Berthon, *Numerical approximations of the* 10*-moment Gaussian closure*, Math. Comp., 75 (2006), pp. 1809–1831.

[4] C. Berthon, *Stability of the MUSCL schemes for the Euler equations*, Commun. Math. Sci., 3 (2005), pp. 133–158.

[5] C. Berthon, *Robustness of MUSCL schemes for* 2*D unstructured meshes*, J. Comput. Phys., 218 (2006), pp. 495–509.

[6] C. Berthon, *Why the MUSCL-Hancock scheme is* L^1*-stable*, Numer. Math., 104 (2006), pp. 27–46.

[7] F. Bouchut, *Non-linear Stability of Finite Volume Methods for Hyperbolic Conservation Laws and Well-balanced Schemes for Sources*, Front. Math., Birkhauser, Basel, Switzerland, 2004.

[8] T. Buffard, T. Gallouët, and J. M. Hérard, *A naive Godunov scheme to solve shallow water equations*, C.R. Math. Acad. Sci. Paris, 326 (1998), pp. 385–390.

[9] T. Buffard, T. Gallouët, and J. M. Hérard, *A sequel to a rough Godunov scheme: Application to real gases*, Comput. Fluids, 29 (2000), pp. 813–847.

[10] G. F. Carrier and H. P. Greenspan, *Water waves of finite amplitude on a sloping beach*, J. Fluid Mech., 4 (1958), pp. 97–109.

[11] G. Q. Chen, C. D. Levermore, and T. P. Liu, *Hyperbolic conservation laws with stiff relaxation terms and entropy*, Comm. Pure Appl. Math., 47 (1995), pp. 787–830.

[12] F. Coquel and B. Perthame, *Relaxation of energy and approximate Riemann solvers for general pressure laws in fluid dynamics*, SIAM J. Numer. Anal., 35 (1998), pp. 2223–2249.

[13] T. Gallouët, J. M. Hérard, and N. Seguin, *Some approximate Godunov schemes to compute shallow-water equations with topography*, Computers and Fluids, 32 (2003), pp. 479–513.

[14] T. Gallouët, J. M. Hérard, and N. Seguin, *On the use of some symetrizing variables to deal with vacuum*, Calcolo, 40 (2003), pp. 163–194.

[15] T. Gallouët and T. Masella, *A rough Godunov scheme*, C. R. Math. Acad. Sci. Paris, 323 (1996), pp. 77–83.

[16] E. Godlewski and P. A. Raviart, *Numerical Approximation of Hyperbolic System of Conservation Laws*, Appl. Math. Sci. 118, Springer, New York, 1996.

[17] T. Gallouët, J. M. Hérard, and N. Seguin, *Some recent finite volume schemes to compute Euler equations using real gas EOS*, Int. J. Numer. Methods Fluids, 39 (2002), pp. 1073-1138.

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

[19] J. M. Greenberg, A. Y. Leroux, R. Baraille, and A. Noussair, *Analysis and approximation of conservation laws with source terms*, SIAM J. Numer. Anal., 34 (1997), pp. 1980–2007.

[20] A. Harten, J. M. Hyman, and P. D. Lax, *On finite difference approximations and entropy conditions*, Comm. Pure Appl. Math., 29 (1976), pp. 297–322.

[21] A. Harten, P. D. Lax, and B. Van Leer, *On upstream differencing and Godunov-type schemes for hyperbolic conservation laws*, SIAM Rev., 25 (1983), pp. 35–61.

[22] K. Hu, C. G. Mingham, and D. M. Causon, *Numerical simulation of wave overtopping of coastal structures using the non-linear shallow water equations*, Coastal Eng., 41 (2000), pp. 433–465.

[23] S. Jin, *A steady-state capturing method for hyperbolic systems with geometrical source terms*, M2AN Math. Model. Numer. Anal., 35 (2001), pp. 631–645.

[24] S. Jin and X. Wen, *Two interface-type numerical methods for computing hyperbolic systems with geometrical source terms having concentrations*, SIAM J. Sci. Comput., 26 (2005), pp. 2079–2101.
[25] S. Jin and X. Wen, *An efficient method for computing hyperbolic systems with geometrical source terms having concentrations*, Special issue dedicated to the 70th birthday of Professor Zhong-Ci Shi, J. Comput. Math., 22 (2004), pp. 230–249.
[26] S. Jin and Z. Xin, *The relaxation scheme for systems of conservation laws in arbitrary space dimension*, Comm. Pure Appl. Math., 45 (1995), pp. 235–276.
[27] P. D. Lax, *Hyperbolic systems of conservation laws and the mathematical theory of shock waves*, Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, 11, SIAM, Philadelphia, 1973.
[28] P. D. Lax, *Shock waves and entropy*, in Contributions to Nonlinear Functional Analysis, E. H. Zarantonello, ed., Academic Press, New York, 1971, pp. 603–634.
[29] B. van Leer, *Towards the ultimate conservative difference scheme. V. A second-order sequel to Godunov's method*, J. Comput. Phys., 32 (1979), pp. 101–136.
[30] R. J. LeVeque, *Finite Volume Methods for Hyperbolic Problems*, Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge, 2002.
[31] R. J. LeVeque and M. Pelanti, *A class of approximate Riemann solvers and their relation to relaxation schemes*, J. Comp. Phys., 172 (2001), pp. 572–591.
[32] T. P. Liu, *Hyperbolic conservation laws with relaxation*, Comm. Math. Phys., 108 (1987), pp. 153–175.
[33] F. Marche, *Derivation of a new two-dimensional viscous shallow water model with varying topography, bottom friction and capillary effects*, Eur. J. Mech. B Fluids, 26 (2007), pp. 49–63.
[34] F. Marche, P. Bonneton, P. Fabrie, and N. Seguin, *Evaluation of well-balanced bore-capturing schemes for 2D wetting and drying processes*, Int. J. Numer. Methods Fluids, 53 (2007), pp. 867–894.
[35] J. M. Masella, I. Faille, and T. Gallouët, *On an approximate Godunov scheme*, Int. J. Comp. Fluid. Dyn., 12 (1999), pp. 133–149.
[36] S. Noelle, N. Pankratz, G. Puppo, and J. R. Natvig, *Well-balanced finite volume schemes of arbitrary order of accuracy for shallow water flows*, J. Comp. Phys., 213 (2006), pp. 474–499.
[37] B. Perthame and C. W. Shu, *On positivity preserving finite volume schemes for Euler equations*, Numer. Math., 73 (1996), pp. 119–130.
[38] C. W. Shu, *Total-variations-diminishing time discretization*, SIAM J. Sci. Statist. Comput., 9 (1988), pp. 1073–1084.
[39] I. Suliciu, *On modelling phase transitions by means of rate-type constitutive equations, shock wave structure*, Int. J. Engrg. Sci., 28 (1990), pp. 829–841.
[40] I. Suliciu, *Some stability-instability problems in phase transitions modelled by piecewise linear elastic or viscoelastic constitutive equations*, Int. J. Engrg. Sci., 30 (1992), pp. 483–494.
[41] G. Wei and J. T. Kirby, *Time-dependant numerical code for extended Boussinesq equations*, J. Water. Port. Coast. Ocean Eng., 121 (1995), pp. 251–261.
[42] Y. Xing and C. W. Shu, *High order finite difference WENO schemes with the exact conservation property for the shallow water equations*, J. Comp. Phys., 208 (2005), pp. 206–227.
[43] S. Yamamoto and H. Daiguji, *Higher order accurate upwind schemes for solving the compressible Euler and Navier-Stokes equations*, Comput. Fluids, 22 (1993), pp. 259–270.