HIGH ORDER FINITE VOLUME SCHEMES BASED ON RECONSTRUCTION OF STATES FOR SOLVING HYPERBOLIC SYSTEMS WITH NONCONSERVATIVE PRODUCTS. APPLICATIONS TO SHALLOW-WATER SYSTEMS

MANUEL CASTRO, JOSÉ M. GALLARDO, AND CARLOS PARÉS

ABSTRACT. This paper is concerned with the development of high order methods for the numerical approximation of one-dimensional nonconservative hyperbolic systems. In particular, we are interested in high order extensions of the generalized Roe methods introduced by I. Toumi in 1992, based on WENO reconstruction of states. We also investigate the well-balanced properties of the resulting schemes. Finally, we will focus on applications to shallow-water systems.

1. INTRODUCTION

The motivating question of this paper was the design of well-balanced high order numerical schemes for PDE systems that can be written under the form

$$(1.1) \qquad \frac{\partial w}{\partial t} + \frac{\partial F}{\partial x}(w) = \mathcal{B}(w)\frac{\partial w}{\partial x} + S(w)\frac{d\sigma}{dx},$$

where the unknown $w(x,t)$ takes values on an open convex subset D of $\mathbb{R}^N$, F is a regular function from D to $\mathbb{R}^N$, $\mathcal{B}$ is a regular matrix-valued function from D to $\mathcal{M}_{N \times N}(\mathbb{R})$, S is a function from D to $\mathbb{R}^N$, and $\sigma(x)$ is a known function from $\mathbb{R}$ to $\mathbb{R}$.

System (1.1) includes as particular cases: systems of conservation laws ($\mathcal{B} = 0$, $S = 0$), systems of conservation laws with source term or balance laws ($\mathcal{B} = 0$), and coupled systems of conservation laws.

More precisely, the discretization of the shallow-water systems that govern the flow of one layer or two superposed layers of immiscible homogeneous fluids was focused. The corresponding systems can be written respectively as a balance law or a coupled system of two conservation laws. Systems with similar characteristics also appear in other flow models, such as boiling flows and two-phase flows (see [13]).

It is well known that standard methods that correctly solve systems of conservation laws can fail in solving (1.1), especially when approaching equilibria or near to equilibria solutions. In the context of shallow-water equations, Bermúdez and Vázquez-Cendón introduced in [2] the condition called *conservation property* or *C-property*: a scheme is said to satisfy this condition if it correctly solves the steady state solutions corresponding to water at rest. This idea of constructing numerical schemes that preserve some equilibria, which are called in general *well-balanced* schemes, has been extended in different ways; see, e.g., [3], [6], [7], [8], [12], [14], [17], [18], [21], [22], [26], [28], [29], [30], [35].

Among the main techniques used in the derivation of well-balanced numerical schemes, one of them consists in first choosing a standard conservative scheme for the discretization of the flux terms and then discretizing the source and the coupling terms in order to obtain a consistent scheme which correctly solves a predetermined family of equilibria. This was the approach in [2] where the authors proved, in the context of shallow-water equations, that numerical schemes based on Roe methods for the discretization of the flux terms and *upwinding* the source term exacly solve equilibria corresponding to water at rest. In [12] it was shown that the technique of modified equations can be helpful in the deduction of well-balanced numerical schemes.

This procedure has been succesfully applied to obtain high order numerical schemes for some particular cases of (1.1) (see, for instance, [4], [38] and [39]). The main disadvantage of this first technique is its lack of generality: the calculation of the correct discretization of the source and the coupling terms depends on both the specific problem and the conservative numerical scheme chosen.

Another technique to obtain well-balanced first order schemes for solving (1.1) consists in considering piecewise constant approximations of the solutions that are updated by means of Approximate Riemann Solvers at the intercells. In particular, Godunov's methods, i.e., methods based on Exact Riemann Solvers, have been used in the context of shallow-water systems in [1], [9], [10], [15], [21], [22]. This approach was also used in [5], where the flux and the coupling terms of a bilayer shallow-water system were treated together by using a generalized Roe linearization.

If this second procedure is followed, the main difficulty both from the mathematical and the numerical points of view comes from the presence of nonconservative products, which makes difficult even the definition of weak solutions: in general, the product $\mathcal{B}(w)w_x$ does not make sense as a distribution for discontinuous solutions. This is also the case for the product $S(w)\sigma_x$ when piecewise constant approximations of σ are considered.

A helpful strategy in solving these difficulties consists in considering system (1.1) as a particular case of a one-dimensional quasilinear hyperbolic system:

$$(1.2) \qquad \frac{\partial W}{\partial t} + \mathcal{A}(W)\frac{\partial W}{\partial x} = 0, \quad x \in \mathbb{R}, \ t > 0.$$

In effect, adding to (1.1) the trivial equation

$$\frac{\partial \sigma}{\partial t} = 0,$$

system (1.1) can be easily rewritten under this form (see [17], [18], [21], [22]).

In [11], Dal Maso, LeFloch, and Murat proposed an interpretation of nonconservative products as Borel measures, based on the choice of a family of paths in the phases space. After this theory it is possible to give a rigorous definition of

weak solutions of (1.2). Together with the definition of weak solutions, a notion of *entropy* has to be chosen as the usual Lax's concept or one related to an entropy pair. Once this choice has been done, the classical theory of simple waves of hyperbolic systems of conservation laws and the results concerning the solutions of Riemann problems can be extended to systems of the form (1.2).

The introduction of a family of paths does not only give a way to properly define the concept of weak solution for nonconservative systems, it also allows us to extend to this framework some basic concepts related to the numerical approximation of weak solutions of conservation laws. For instance, in [36] a general definition of Roe linearizations was introduced, also based on the use of a family of paths. In [28] a general definition of well-balanced schemes for solving (1.2) was introduced. It was shown that the well-balanced properties of these generalized Roe methods depend on the choice of the family of paths. Moreover, this general methodology was applied to some systems of the form (1.1) related to shallow-water flows, recovering some known well-balanced schemes, or resulting in new schemes.

The goal of this paper is to obtain the general expression of a well-balanced high order method for (1.2) based on the use of a first order Roe scheme and reconstruction of states. The interest of such a general expression is that, once obtained, particular schemes can be deduced for any system of the form (1.1), where the numerical treatment of source and coupling terms is automatically derived. To our knowledge, the present work is the first attempt to obtain well-balanced high order numerical schemes following this procedure.

The paper is organized as follows. In Section 2 we give some basic definitions and results about nonconservative systems, Roe linearizations and generalized Roe schemes, for which we will follow [28] closely. High order versions of the Roe schemes, based on reconstruction operators, are introduced in Section 3. Next, Section 4 is devoted to the analysis of the well-balanced properties of the high order schemes previously constructed. In Section 6, the WENO method is applied to build the reconstruction operators. Applications to a family of systems that generalize (1.1) are presented in Section 5, with particular interest in some shallow-water systems with one and two layers of fluid. Finally, Section 7 contains numerical results to test the performances of our high-order schemes. In particular, the high order well-balanced property is numerically verified.

2. Roe methods for nonconservative hyperbolic systems

Consider the system in nonconservative form

$$(2.1) \qquad\qquad W_t + \mathcal{A}(W)W_x = 0, \quad x \in \mathbb{R}, \ t > 0,$$

where we suppose that the range of $W(x,t)$ is contained inside an open convex subset Ω of $\mathbb{R}^N$, and $W \in \Omega \mapsto \mathcal{A}(W) \in \mathcal{M}_N(\mathbb{R})$ is a smooth locally bounded map. The system (2.1) is assumed to be strictly hyperbolic: for each $W \in \Omega$ the matrix $\mathcal{A}(W)$ has N real distinct eigenvalues $\lambda_1(W) < \cdots < \lambda_N(W)$. We also suppose that the jth characteristic field R_j is either genuinely nonlinear:

$$R_j(W) \cdot \nabla\lambda_j(W) \neq 0, \quad \forall W \in \Omega,$$

or linearly degenerate:

$$R_j(W) \cdot \nabla\lambda_j(W) = 0, \quad \forall W \in \Omega.$$

For discontinuous solutions W, the nonconservative product $\mathcal{A}(W)W_x$ does not make sense as a distribution. However, the theory developed by Dal Maso, LeFloch and Murat ([11]) allows us to give a rigorous definition of nonconservative products, associated to the choice of a family of paths in Ω.

Definition 2.1. A *family of paths* in $\Omega \subset \mathbb{R}^N$ is a locally Lipschitz map

$$\Phi \colon [0,1] \times \Omega \times \Omega \to \Omega$$

that satisfies the following properties:

(1) $\Phi(0; W_L, W_R) = W_L$ and $\Phi(1; W_L, W_R) = W_R$, for any $W_L, W_R \in \Omega$.
(2) Given an arbitrary bounded set $\mathcal{B} \subset \Omega$, there exists a constant k such that

$$\left| \frac{\partial \Phi}{\partial s}(s; W_L, W_R) \right| \leq k|W_L - W_R|,$$

for any $W_L, W_R \in \mathcal{B}$ and for almost every $s \in [0,1]$.
(3) For every bounded set $\mathcal{B} \subset \Omega$, there exists a constant K such that

$$\left| \frac{\partial \Phi}{\partial s}(s; W_L^1, W_R^1) - \frac{\partial \Phi}{\partial s}(s; W_L^2, W_R^2) \right| \leq K(|W_L^1 - W_L^2| + |W_R^1 - W_R^2|),$$

for each $W_L^1, W_R^1, W_L^2, W_R^2 \in \mathcal{B}$ and for almost every $s \in [0,1]$.

Suppose that a family of paths Φ in Ω has been chosen. Then, for $W \in (L^\infty(\mathbb{R} \times \mathbb{R}^+) \cap BV(\mathbb{R} \times \mathbb{R}^+))^N$, the nonconservative product can be interpreted as a Borel measure denoted by $[\mathcal{A}(W)W_x]_\Phi$. When no confusion arises, we will drop the dependency on Φ.

A *weak solution* of system (2.1) is defined as a function $W \in (L^\infty(\mathbb{R} \times \mathbb{R}^+) \cap BV(\mathbb{R} \times \mathbb{R}^+))^N$ that satisfies the equality

$$W_t + [\mathcal{A}(W)W_x]_\Phi = 0.$$

In particular, a piecewise $\mathcal{C}^1$ function W is a weak solution of (2.1) if and only if the two following conditions are satisfied:

(i) W is a classical solution in the domains where it is $\mathcal{C}^1$.
(ii) Along a discontinuity W satisfies the jump condition

$$(2.2) \qquad \int_0^1 \left(\xi \mathcal{I} - \mathcal{A}(\Phi(s; W^-, W^+)) \right) \frac{\partial \Phi}{\partial s}(s; W^-, W^+) \, ds = 0,$$

where $\mathcal{I}$ is the identity matrix, ξ is the speed of propagation of the discontinuity, and W^-, W^+ are the left and right limits of the solution at the discontinuity.

Note that in the particular case of a system of conservation laws (that is, $\mathcal{A}(W)$ is the Jacobian matrix of some flux function $F(W)$) the jump condition (2.2) is independent of the family of paths, and it reduces to the usual Rankine-Hugoniot condition:

$$(2.3) \qquad\qquad F(W^+) - F(W^-) = \xi(W^+ - W^-).$$

In the general case, the selection of the family of paths has to be based on the physical background of the problem under consideration. Nevertheless, it is natural from the mathematical point of view to require this family to satisfy some hypotheses concerning the relation of the paths with the integral curves of the characteristic fields. For instance, if W_L and W_R are linked by an integral curve

of a linearly degenerate field, the natural choice of the path is a parametrization of that curve, as this choice assures that the contact discontinuity

$$(2.4) \qquad W(x,t) = \begin{cases} W_L & \text{if } x < \xi t, \\ W_R & \text{if } x > \xi t, \end{cases}$$

where ξ is the (constant) value of the corresponding eigenvalue through the integral curve, is a weak solution of the problem, as would be the case for a system of conservation laws.

Due to these requirements, the explicit calculation of the path linking two given states W_L and W_R can be difficult: in most cases, the explicit expression of the solution of the Riemann problem related to the states is needed (see [28]).

Together with this definition of weak solutions, a notion of *entropy* has to be chosen, either as the usual Lax's concept or one related to an entropy pair (see [16] for details). Once this choice has been done, the theory of simple waves of hyperbolic systems of conservation laws and the results concerning the solutions of Riemann problems can be naturally extended to systems of the form (2.1) (see [11]).

Some of the usual numerical schemes designed for conservation laws can be adapted to the discretization of the more general system (2.1). This is the case of Roe schemes (see [31]): in [36] a general definition of Roe linearization was introduced, based again on the use of a family of paths.

Definition 2.2. Given a family of paths Ψ, a function $\mathcal{A}_\Psi : \Omega \times \Omega \to \mathcal{M}_N(\mathbb{R})$ is called a *Roe linearization* of system (2.1) if it verifies the following properties:

(1) For each $W_L, W_R \in \Omega$, $\mathcal{A}_\Psi(W_L, W_R)$ has N distinct real eigenvalues.
(2) $\mathcal{A}_\Psi(W, W) = \mathcal{A}(W)$, for every $W \in \Omega$.
(3) For any $W_L, W_R \in \Omega$,

$$(2.5) \qquad \mathcal{A}_\Psi(W_L, W_R)(W_R - W_L) = \int_0^1 \mathcal{A}(\Psi(s; W_L, W_R)) \frac{\partial \Psi}{\partial s}(s; W_L, W_R) \, ds.$$

Note again that if $\mathcal{A}(W)$ is the Jacobian matrix of a smooth flux function $F(W)$, (2.5) is independent of the family of paths and reduces to the usual Roe property:

$$(2.6) \qquad \mathcal{A}_\Psi(W_R - W_L) = F(W_R) - F(W_L).$$

Once a Roe linearization $\mathcal{A}_\Psi$ has been chosen, in order to construct numerical schemes for solving (2.1), computing cells $I_i = [x_{i-1/2}, x_{i+1/2}]$ are considered; let us suppose for simplicity that the cells have constant size Δx and that $x_{i+\frac{1}{2}} = i\Delta x$. Define $x_i = (i - 1/2)\Delta x$, the center of the cell I_i. Let Δt be the constant time step and define $t^n = n\Delta t$. Denote by W_i^n the approximation of the cell averages of the exact solution provided by the numerical scheme, that is,

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

Then, the numerical scheme advances in time by solving linear Riemann problems at the intercells at time t^n and taking the averages of their solutions on the cells at time t^{n+1}. Under usual CFL conditions, it can be written as follows (see [28]):

$$(2.7) \qquad W_i^{n+1} = W_i^n - \frac{\Delta t}{\Delta x}\left(\mathcal{A}_{i-1/2}^+(W_i^n - W_{i-1}^n) + \mathcal{A}_{i+1/2}^-(W_{i+1}^n - W_i^n)\right).$$

Here, the intermediate matrices are defined by

$$\mathcal{A}_{i+1/2} = \mathcal{A}_\Psi(W_i^n, W_{i+1}^n),$$

and, as usual,

$$\mathcal{L}_{i+1/2}^\pm = \begin{bmatrix} (\lambda_1^{i+1/2})^\pm & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & (\lambda_N^{i+1/2})^\pm \end{bmatrix}, \quad \mathcal{A}_{i+1/2}^\pm = \mathcal{K}_{i+1/2} \mathcal{L}_{i+1/2}^\pm \mathcal{K}_{i+1/2}^{-1},$$

where $\mathcal{L}_{i+1/2}$ is the diagonal matrix whose coefficients are the eigenvalues of $\mathcal{A}_{i+1/2}$:

$$\lambda_1^{i+1/2} < \lambda_2^{i+1/2} < \cdots < \lambda_N^{i+1/2},$$

and $\mathcal{K}_{i+1/2}$ is an $N \times N$ matrix whose columns are associated eigenvectors.

In the particular case of a system of conservation laws, (2.7) can be written under the usual form of a conservative numerical scheme. First, the numerical flux is defined by

$$G_{i+1/2} = G(W_i^n, W_{i+1}^n) = \frac{1}{2}\big(F(W_i^n) + F(W_{i+1}^n)\big) - \frac{1}{2}\big|\mathcal{A}_{i+1/2}\big|(W_{i+1}^n - W_i^n),$$

where

$$\big|\mathcal{A}_{i+1/2}\big| = \mathcal{A}_{i+1/2}^+ - \mathcal{A}_{i+1/2}^-.$$

Then, the following equalities can be easily verified:

$$(2.8) \qquad F(W_{i+1}^n) - G_{i+1/2} = \mathcal{A}_{i+1/2}^+(W_{i+1}^n - W_i^n),$$

$$(2.9) \qquad G_{i+1/2} - F(W_i^n) = \mathcal{A}_{i+1/2}^-(W_{i+1}^n - W_i^n).$$

Finally using these equalities in (2.7) we obtain:

$$W_{i+1}^n = W_i^n + \frac{\Delta t}{\Delta x}\big(G_{i-1/2} - G_{i+1/2}\big).$$

The best choice of the family of paths Ψ appearing in the definition of Roe linearization is the family Φ selected for the definition of weak solutions. In effect, Roe methods based on the family of paths Φ can correctly solve discontinuities in the following sense: let us suppose that W_i^n and W_{i+1}^n can be linked by an entropic discontinuity propagating at speed ξ; then, from (2.5) and (2.2) we deduce that

$$\mathcal{A}_{i+1/2}\big(W_{i+1}^n - W_i^n\big) = \xi\big(W_{i+1}^n - W_i^n\big),$$

i.e., ξ is an eigenvalue of the intermediate matrix and $W_{i+1}^n - W_i^n$ is an associated eigenvector. As a consequence, the solution of the linear Riemann problem corresponding to the intercell $x_{i+1/2}$,

$$\begin{cases} U_t + \mathcal{A}_{i+1/2} U_x = 0, \\ U(x, t^n) = \begin{cases} W_i^n & \text{if } x < x_{i+1/2}, \\ W_{i+1}^n & \text{if } x > x_{i+1/2}, \end{cases} \end{cases}$$

coincides with the solution of the Riemann problem

$$\begin{cases} U_t + \mathcal{A}(U) U_x = 0, \\ U(x, t^n) = \begin{cases} W_i^n & \text{if } x < x_{i+1/2}, \\ W_{i+1}^n & \text{if } x > x_{i+1/2}. \end{cases} \end{cases}$$

Both solutions consist of a discontinuity linking the states and propagating at velocity ξ.

Nevertheless, the construction of these schemes with $\Psi = \Phi$ can be difficult or very costly in practice. In this case, a simpler family of paths Ψ has to be chosen as the family of segments

$$(2.10) \qquad \Psi(s; W_L, W_R) = W_L + s(W_R - W_L), \quad s \in [0, 1].$$

In [28] it was remarked that, in this case, the convergence of the numerical scheme can fail when the weak solution to approach involves discontinuities connecting states W^- and W^+ such that the paths of the families Φ and Ψ linking them are different.

As in the case of a system of conservation laws, the scheme (2.7) has to be used with a CFL condition:

$$\max\left\{|\lambda_l^{i+1/2}|, \ 1 \le l \le N, \ i \in \mathbb{Z}\right\} \frac{\Delta t}{\Delta x} \le \gamma,$$

with $0 < \gamma \le 1$. An entropy fix technique, as the Harten-Hyman one ([24], [25]), also has to be included.

3. High order schemes based on reconstruction of states

In the case of systems of conservation laws

$$(3.1) \qquad W_t + F(W)_x = 0,$$

there are several methods to obtain higher order schemes based on the use of a reconstruction operator. In particular, methods based on the reconstruction of states are built using the following procedure: given a first order conservative scheme with numerical flux function $G(U, V)$, a reconstruction operator of order p is considered, that is, an operator that associates to a given sequence $\{W_i\}$ two new sequences, $\{W_{i+1/2}^-\}$ and $\{W_{i+1/2}^+\}$, in such a way that, whenever

$$W_i = \frac{1}{\Delta x} \int_{I_i} W(x) \, dx$$

for some smooth function W, we have that

$$W_{i+1/2}^\pm = W(x_{i+1/2}) + O(\Delta x^p).$$

Once the first order method and the reconstruction operator have been chosen, the method of lines can be used to develop high order methods for (3.1): the idea is to discretize only in space, leaving the problem continuous in time. This procedure leads to a system of ordinary differential equations which is solved using a standard numerical method. In particular, we assume here that the first order scheme is a Roe method.

Let $\overline{W}_j(t)$ denote the cell average of a regular solution W of (3.1) over the cell I_i at time t:

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

The following equation can be easily obtained for the cell averages:

$$\overline{W}_i'(t) = \frac{1}{\Delta x} \big(F(W(x_{i-1/2}, t)) - F(W(x_{i+1/2}, t)) \big).$$

This system is now approached by

$$(3.2) \qquad W_i'(t) = \frac{1}{\Delta x}\big(\widetilde{G}_{i-1/2} - \widetilde{G}_{i+1/2}\big),$$

with

$$\widetilde{G}_{i+1/2} = G(W^-_{i+1/2}(t), W^+_{i+1/2}(t)),$$

where $W_i(t)$ is the approximation to $\overline{W}_i(t)$ provided by the scheme, and $W^\pm_{i+1/2}(t)$ is the reconstruction associated to the sequence $\{W_j(t)\}$.

Let us now generalize this semi-discrete method for a nonconservative system (2.1). Observe that, in Section 2, the key point to generalize both the Rankine-Hugoniot condition (2.3) and the Roe property (2.6) to system (2.1) was to replace a difference of fluxes by an integral along a path. Let us apply the same technique here. First of all, as the first order scheme is a Roe method, using the equalities (2.8) and (2.9) (replacing W_i^n and W_{i+1}^n by $W^-_{i+1/2}(t)$ and $W^+_{i+1/2}(t)$, respectively) we can rewrite (3.2) as follows:

$$(3.3)$$
$$W_i'(t) = -\frac{1}{\Delta x}\big(\mathcal{A}^+_{i-1/2}(W^+_{i-1/2}(t) - W^-_{i-1/2}(t)) $$
$$+ \mathcal{A}^-_{i+1/2}(W^+_{i+1/2}(t) - W^-_{i+1/2}(t)) - F(W^+_{i-1/2}(t)) + F(W^-_{i+1/2}(t))),$$

where $\mathcal{A}_{i+1/2}$ is the intermediate matrix corresponding to the states $W^-_{i+1/2}(t)$ and $W^+_{i+1/2}(t)$.

Let us now introduce, at every cell I_i, *any* regular function P_i^t such that

$$(3.4) \qquad \lim_{x \to x^+_{i-1/2}} P_i^t(x) = W^+_{i-1/2}(t), \qquad \lim_{x \to x^-_{i+1/2}} P_i^t(x) = W^-_{i+1/2}(t).$$

Then, (3.3) can now be written under the form

$$(3.5)$$
$$W_i'(t) = -\frac{1}{\Delta x}\bigg(\mathcal{A}^+_{i-1/2}(W^+_{i-1/2}(t) - W^-_{i-1/2}(t)) $$
$$+ \mathcal{A}^-_{i+1/2}(W^+_{i+1/2}(t) - W^-_{i+1/2}(t)) + \int_{x_{i-1/2}}^{x_{i+1/2}} \frac{d}{dx} F(P_i^t(x))\, dx\bigg).$$

Note now that (3.5) can be easily generalized to obtain a numerical scheme for solving (2.1):

$$(3.6)$$
$$W_i'(t) = -\frac{1}{\Delta x}\bigg(\mathcal{A}^+_{i-1/2}(W^+_{i-1/2}(t) - W^-_{i-1/2}(t)) $$
$$+ \mathcal{A}^-_{i+1/2}(W^+_{i+1/2}(t) - W^-_{i+1/2}(t)) + \int_{x_{i-1/2}}^{x_{i+1/2}} \mathcal{A}(P_i^t(x))\frac{d}{dx} P_i^t(x)\, dx\bigg),$$

where the intermediate matrices are defined by means of a Roe linearization based on a family of paths Ψ and P_i^t is again a regular function satisfying (3.4).

Remark 3.1. It is important to note that for conservative problems, the numerical scheme (3.6) is equivalent to the conservative numerical scheme (3.3) if, and only if, the integral term is computed exactly. However, the formulation (3.6) is useless when working with conservative problems, as we would get involved with a more complex expression of the numerical scheme. The numerical scheme (3.6) is useful only for problems with nonconservative products, as it allows us to deduce numerical schemes for particular problems, using numerical quadratures if necessary.

There is an important difference between the conservative and the nonconservative case: in the conservative case the numerical scheme is independent of the functions P_i^t chosen at the cells, but this is not the case for nonconservative problems. As a consequence, while the numerical scheme (3.2) has the same order of the reconstruction operator, in the case of the scheme (3.6) it seems clear that, in order to have a high order scheme, together with a high order reconstruction operator, the functions P_i^t and their derivatives have to be high order approximations of $W(\cdot, t)$ and its partial derivative $W(\cdot, t)_x$.

In practice, the definition of the reconstruction operator gives the natural choice of the function P_i^t, as the usual procedure to define a reconstruction operator is the following: given a sequence $\{W_i\}$ of values at the cells, first an approximation function is constructed at every cell I_i, based on the values of W_i at some of the neighbor cells (the *stencil*):

$$P_i(x; W_{i-l}, \ldots, W_{i+r}),$$

for some natural numbers l, r. These approximations functions are calculated by means of an interpolation or approximation procedure. Once these functions have been constructed, $W_{i+1/2}^-$ (resp. $W_{i+1/2}^+$) is calculated by taking the limit of P_i (resp. P_{i+1}) to the left (resp. to the right) of $x_{i+1/2}$. If the reconstruction operator is built following this procedure (as we will assume in the sequel), the natural choice of P_i^t is

$$P_i^t(x) = P_i(x; W_{i-l}(t), \ldots, W_{i+r}(t)).$$

Let us now investigate the order of the numerical scheme (3.6). Note first that, for regular solutions W of (2.1), the cell averages at the cells $\{\overline{W}_j(t)\}$ satisfy

$$(3.7) \qquad \overline{W}_i'(t) = -\frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} \mathcal{A}(W(x,t)) W_x(x,t)\, dx.$$

Thus, (3.6) is expected to be a good approximation of (3.7). This fact is stated in the following result:

Theorem 3.2. *Let us assume that $\mathcal{A}$ is of class $\mathcal{C}^2$ with bounded derivatives and $\mathcal{A}_\Psi$ is bounded. Let us also suppose that the p-order reconstruction operator is such that, given a sequence defined by*

$$W_i = \frac{1}{\Delta x} \int_{I_i} W(x)\, dx$$

for some smooth function W, we have that

$$P_i(x; W_{i-l}, \ldots, W_{i+r}) = W(x) + O(\Delta x^q), \quad \forall\, x \in I_i,$$
$$\frac{d}{dx} P_i(x) = W'(x) + O(\Delta x^r), \quad \forall\, x \in I_i.$$

Then (3.6) *is an approximation of order at least* $\gamma = \min(p, q + 1, r + 1)$ *to the system* (3.7) *in the following sense:*

$$
\begin{aligned}
(3.8) \quad & \mathcal{A}^+_{i-1/2}(W^+_{i-1/2}(t) - W^-_{i-1/2}(t)) \\
& + \mathcal{A}^-_{i+1/2}(W^+_{i+1/2}(t) - W^-_{i+1/2}(t)) \\
& + \int_{x_{i-1/2}}^{x_{i+1/2}} \mathcal{A}(P^t_i(x)) \frac{d}{dx} P^t_i(x)\, dx \\
& = \int_{x_{i-1/2}}^{x_{i+1/2}} \mathcal{A}(W(x,t)) W_x(x,t)\, dx + O(\Delta x^\gamma),
\end{aligned}
$$

for every smooth enough solution W, $W^\pm_{i+1/2}(t)$ *being the associated reconstructions and* P^t_i *the approximation functions corresponding to the sequence*

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

Proof. On the one hand, as the reconstruction operator is of order p, we have

$$
\mathcal{A}^+_{i-1/2}\big(W^+_{i-1/2}(t) - W^-_{i-1/2}(t)\big) + \mathcal{A}^-_{i+1/2}\big(W^+_{i+1/2}(t) - W^-_{i+1/2}(t)\big) = O(\Delta x^p).
$$

On the other hand,

$$
\begin{aligned}
& \int_{x_{i-1/2}}^{x_{i+1/2}} \mathcal{A}(P^t_i(x)) \frac{d}{dx} P^t_i(x)\, dx - \int_{x_{i-1/2}}^{x_{i+1/2}} \mathcal{A}(W(x,t)) W_x(x,t)\, dx \\
& = \int_{x_{i-1/2}}^{x_{i+1/2}} \big(\mathcal{A}(P^t_i(x)) - \mathcal{A}(W(x,t))\big) \frac{d}{dx} P^t_i(x)\, dx \\
& + \int_{x_{i-1/2}}^{x_{i+1/2}} \mathcal{A}(W(x,t)) \big(W_x(x,t) - \frac{d}{dx} P^t_i(x)\big)\, dx \\
& = O(\Delta x^{r+1}) + O(\Delta x^{q+1}).
\end{aligned}
$$

The equality (3.8) is easily deduced from the above equalities. $\qquad \square$

Remark 3.3. For the usual reconstruction operators one has $r \le q \le p$, and thus the order of (3.6) is $r + 1$ for nonconservative systems and p for conservation laws. Therefore a loss of accuracy can be observed when a technique of reconstruction giving order p for systems of conservation laws is applied to a nonconservative problem.

4. Well-balanced property

In this paragraph we investigate the well-balanced properties of schemes of the form (3.6). Well-balancing is related with the numerical approximation of equilibria, i.e., steady state solutions. System (2.1) can only have nontrivial steady state solutions if it has linearly degenerate fields: if $W(x)$ is a regular steady state solution it satisfies

$$
\mathcal{A}(W(x)) \cdot W'(x) = 0, \quad \forall x \in \mathbb{R},
$$

and then 0 is an eigenvalue of $\mathcal{A}(W(x))$ for all x and $W'(x)$ is an eigenvector. Therefore, the solution can be interpreted as a parametrization of an integral curve of a linearly degenerate characteristic field whose corresponding eigenvalue takes the value 0 through the curve. In order to define the concept of well-balancing, let us introduce the set Γ of all the integral curves γ of a linearly degenerate

field of $\mathcal{A}(W)$ such that the corresponding eigenvalue vanishes on Γ. According to [28], given a curve $\gamma \in \Gamma$, a numerical scheme is said to be *exactly well-balanced* (respectively *well-balanced with order k*) for γ if it solves exactly (respectively up to order $O(\Delta x^k)$) regular stationary solutions W satisfying $W(x) \in \gamma$ for every x. The numerical scheme is said to be exactly well-balanced or well-balanced with order k if these properties are satisfied for any curve γ of Γ (see [28] for details).

In the cited article, it has been shown that a Roe scheme (2.7) based on a family of paths Ψ is exactly well balanced for a curve $\gamma \in \Gamma$ if, given two states W_L and W_R in γ, the path $\Psi(s; W_L, W_R)$ is a parametrization of the arc of γ linking the states. In particular, if the family of paths Ψ coincides with the one used in the definition of weak solutions Φ, the numerical scheme is exactly well balanced. On the other hand, the numerical scheme is well balanced with order p if $\Psi(s; W_L, W_R)$ approximates with order p a regular parametrization of the arc of γ linking the states. In particular, a Roe scheme based on the family of segments (2.10) is well balanced with order 2. Moreover, it is exactly well balanced for curves of Γ that are straight lines.

The definition of a well-balanced scheme introduced in [28] can be easily extended for semi-discrete methods.

Definition 4.1. Let us consider a semi-discrete method for solving (2.1):

$$(4.1) \qquad \begin{cases} W_i'(t) = \frac{1}{\Delta x}\mathcal{H}(\mathbf{W}(t); i), \\ \mathbf{W}(0) = \mathbf{W}_0, \end{cases}$$

where $\mathbf{W}(t) = \{W_i(t)\}$ represent the vector of approximations to the cell averages of the exact solution, and $\mathbf{W}_0 = \{W_i^0\}$ the vector of initial data. Let γ be a curve of Γ. The numerical method (4.1) is said to be *exactly well balanced* for γ if, given a regular stationary solution W such that

$$W(x) \in \gamma, \quad \forall x \in \mathbb{R},$$

the vector $\mathbf{W} = \{W(x_i)\}$ is a critical point for the system of differential equations in (4.1), i.e.,

$$\mathcal{H}(\mathbf{W}; i) = 0, \quad \forall i.$$

Also, it is said to be *well balanced with order p* if

$$\mathcal{H}(\mathbf{W}; i) = O(\Delta x^p), \quad \forall i.$$

Finally, the semi-discrete method (4.1) is said to be *exactly well balanced* or *well balanced with order p* if these properties are satisfied for every curve γ of the set Γ.

For the particular case of the numerical schemes based on reconstruction of states (3.6) we have

$$\mathcal{H}(\mathbf{W}; i) = \mathcal{A}_{i-1/2}^+(W_{i-1/2}^+ - W_{i-1/2}^-) + \mathcal{A}_{i+1/2}^-(W_{i+1/2}^+ - W_{i+1/2}^-)$$
$$+ \int_{x_{i-1/2}}^{x_{i+1/2}} \mathcal{A}(P_i(x))\frac{d}{dx}P_i(x)\,dx,$$

where $W_{i+1/2}^{\pm}$ are the reconstructions associated to the sequence $\mathbf{W}$ and P_i the corresponding approximation functions. Hereafter, we give two results concerning the well-balanced property of this scheme, but first we introduce a new definition.

Definition 4.2. A reconstruction operator based on smooth approximation functions is said to be *exactly well balanced* for a curve $\gamma \in \Gamma$ if, given a sequence $\{W_i\}$ in γ, the approximation functions satisfy

$$(4.2) \qquad P_i(x) \in \gamma, \quad \forall\, x \in [x_{i-1/2}, x_{i+1/2}],$$

for every i.

Theorem 4.3. *Let γ belong to Γ. Let us suppose that both the generalized Roe method and the reconstruction operator chosen are exactly well balanced for γ. Then the numerical scheme* (3.6) *is also exactly well balanced for γ.*

Proof. Let W be a regular stationary solution satisfying

$$W(x) \in \gamma, \quad \forall\, x,$$

and $\mathbf{W} = \{W(x_i)\}$. From (4.2) and the exactly well-balanced character of the generalized Roe method, we obtain

$$\mathcal{A}^{\pm}_{i+1/2}\big(W^{+}_{i+1/2} - W^{-}_{i+1/2}\big) = 0.$$

On the other hand, using (4.2), P_i can be understood as a parametrization of an arc of γ, which is an integral curve of a linearly degenerate field whose corresponding eigenvalue is zero. Therefore,

$$\int_{x_{i-1/2}}^{x_{i+1/2}} \mathcal{A}(P_i(x)) \frac{d}{dx} P_i(x)\, dx = 0.$$

The proof is easily deduced from the two equalities above. $\qquad\square$

Theorem 4.4. *Under the hypotheses of Theorem* 3.2*, the scheme* (3.6) *is well balanced with order at least $\gamma = \min(p, q+1, r+1)$.*

Proof. The proof is similar to that of Theorem 3.2 $\qquad\square$

Remark 4.5. Note that well-balanced properties for the Roe scheme or the reconstruction operator are not required in this latter result.

5. Applications

We consider in this section systems of the form

$$(5.1) \qquad \frac{\partial W}{\partial t} + \frac{\partial F}{\partial x}(W, \sigma) = \mathcal{B}(W, \sigma) \cdot \frac{\partial W}{\partial x} + \widetilde{S}(W, \sigma) \frac{d\sigma}{dx},$$

where

$$W(x, t) = \begin{bmatrix} w_1(x, t) \\ w_2(x, t) \\ \vdots \\ w_N(x, t) \end{bmatrix} \in \mathbb{R}^N.$$

Here, $\sigma(x)$ is a known function from $\mathbb{R}$ to $\mathbb{R}$, F is a regular function from $\Omega \times \mathbb{R}$ to $\mathbb{R}^N$, Ω is an open convex subset of $\mathbb{R}^N$, $\mathcal{B}$ is a regular matrix-valued function from $\Omega \times \mathbb{R}$ to $\mathcal{M}_N(\mathbb{R})$, and $\widetilde{S}$ is a function from Ω to $\mathbb{R}^N$. We can assume without loss of generality that $\widetilde{S}$ has the form

$$\widetilde{S}(W, \sigma) = S(W, \sigma) + \frac{\partial F}{\partial \sigma}(W, \sigma),$$

for some regular function S.

We denote by $\mathcal{J}(W, \sigma)$ the Jacobian matrix of $F(\cdot, \sigma)$:

$$\mathcal{J}(W, \sigma) = \frac{\partial F}{\partial W}(W, \sigma).$$

System (5.1) includes as particular cases systems of conservation laws ($\mathcal{B} = 0$, $S = 0$) whose flux function may depend on x via the function σ, systems of conservation laws with source term or balance laws ($\mathcal{B} = 0$), or coupled systems of conservation laws as defined in [5]. In this latter case, $\mathcal{J}$ is block-diagonal and the blocks of $\mathcal{B}$ corresponding to the nonzero diagonal blocks of $\mathcal{J}$ are zero.

Following the idea developed in [17], [18] for conservation laws with source terms, if we add to (5.1) the trivial equation

$$\frac{\partial \sigma}{\partial t} = 0,$$

the problem can be written under the form (2.1):

(5.2) $$\widetilde{W}_t + \widetilde{\mathcal{A}}(\widetilde{W}) \cdot \widetilde{W}_x = 0,$$

where $\widetilde{W}$ is the augmented vector

$$\widetilde{W} = \left[\begin{array}{c} W \\ \sigma \end{array} \right],$$

and the block structure of the $(N + 1) \times (N + 1)$ matrix $\widetilde{\mathcal{A}}(\widetilde{W})$ is given by

$$\widetilde{\mathcal{A}}(\widetilde{W}) = \left[\begin{array}{c|c} \mathcal{A}(\widetilde{W}) & -\widetilde{S}(\widetilde{W}, \sigma) \\ 0 & 0 \end{array} \right].$$

Here $\mathcal{A}(\widetilde{W})$ represents the $N \times N$ matrix

$$\mathcal{A}(\widetilde{W}) = \mathcal{J}(W, \sigma) - \mathcal{B}(W, \sigma).$$

We assume that the matrix $\mathcal{A}(\widetilde{W})$ has N real distinct eigenvalues

$$\lambda_1(\widetilde{W}) < \cdots < \lambda_N(\widetilde{W})$$

and associated eigenvectors $R_j(\widetilde{W})$, $j = 1, \ldots, N$. If these eigenvalues do not vanish, (5.2) is a strictly hyperbolic system: $\widetilde{\mathcal{A}}(\widetilde{W})$ has $N + 1$ distinct real eigenvalues

$$\lambda_1(\widetilde{W}), \ldots, \lambda_N(\widetilde{W}), 0,$$

with associated eigenvectors

$$\widetilde{R}_1(\widetilde{W}), \ldots, \widetilde{R}_{N+1}(\widetilde{W}),$$

given by

$$\widetilde{R}_i(\widetilde{W}) = \left[\begin{array}{c} R_i(\widetilde{W}) \\ 0 \end{array} \right], \ i = 1, \ldots, N, \quad \widetilde{R}_{N+1}(\widetilde{W}) = \left[\begin{array}{c} \mathcal{A}(\widetilde{W})^{-1} \cdot S(\widetilde{W}) \\ 1 \end{array} \right].$$

Clearly, the $(N+1)$-th field is linearly degenerate and, for the sake of simplicity, we will suppose that it is the only one. The integral curves of the linearly degenerate field are given by those of the ODE system

$$\frac{d\widetilde{W}}{ds} = \widetilde{R}_{N+1}(\widetilde{W}).$$

Remark 5.1. Note that, in this case, the set Γ defined in the previous section is simply the set of all the integral curves of the linearly degenerate field, as the corresponding eigenvalues always take the value 0. Let us illustrate in this case the relation between these integral curves and the stationary solutions. Let γ be an integral curve of the linearly degenerate field and let us suppose that it can be described implicitly by a system of N equations:

$$(5.3) \qquad g_j(w_1, \ldots, w_N, \sigma) = 0, \quad 1 \le j \le N.$$

As σ is a known function, for every x, (5.3) is a system of N equations with N unknowns $w_1, \ldots, w_N$. The stationary solutions associated to the curve γ are obtained by searching solutions $\{w_1(x), \ldots, w_N(x)\}$ of system (5.3) which depend smoothly on x.

For the definition of weak solutions of system (5.2) and the choice of the family of paths, we refer the interested reader to [28] and the references therein. Let us only mention that the complete definition of the path linking two states is not easy, as it requires the explicit knowledge of the solution of the corresponding Riemann problem. Therefore, the construction of Roe schemes based on the family of paths used in the definition of weak solutions is, in general, a difficult task.

Thus we consider the general case in which the family of paths $\widetilde{\Psi}$ used for the construction of Roe matrices is different to that used in the definition of weak solutions. In particular, in the applications the family of segments (2.10) has been considered.

The following notation will be used:

$$\widetilde{\Psi}(s; \widetilde{W}_i^n, \widetilde{W}_{i+1}^n) = \begin{bmatrix} \Psi(s; \widetilde{W}_i^n, \widetilde{W}_{i+1}^n) \\ \Psi_{N+1}(s; \widetilde{W}_i^n, \widetilde{W}_{i+1}^n) \end{bmatrix} = \begin{bmatrix} \Psi_1(s; \widetilde{W}_i^n, \widetilde{W}_{i+1}^n) \\ \vdots \\ \Psi_N(s; \widetilde{W}_i^n, \widetilde{W}_{i+1}^n) \\ \Psi_{N+1}(s; \widetilde{W}_i^n, \widetilde{W}_{i+1}^n) \end{bmatrix}.$$

Let us suppose that, for any fixed value of σ, Roe matrices can be calculated for the system of conservation laws corresponding to $\mathcal{B} = 0$ and $S = 0$, i.e., we assume that, given W_i^n, W_{i+1}^n and σ, we can calculate a matrix $\mathcal{J}_{i+1/2}^\sigma$ such that

$$\mathcal{J}_{i+1/2}^\sigma \cdot (W_{i+1}^n - W_i^n) = F(W_{i+1}^n, \sigma) - F(W_i^n, \sigma).$$

Let us also suppose that it is possible to calculate a value $\sigma_{i+1/2}$ of σ, a $N \times N$ matrix $\mathcal{B}_{i+1/2}$, and a vector $S_{i+1/2}$, such that the following identities hold:

$$(5.4) \quad F(W_{i+1}^n, \sigma_{i+1}) - F(W_{i+1}^n, \sigma_{i+1/2}) + F(W_i^n, \sigma_{i+1/2}) - F(W_i^n, \sigma_i)$$

$$= \int_0^1 \frac{\partial F}{\partial \sigma}\big(\widetilde{\Psi}(s; \widetilde{W}_i^n, \widetilde{W}_{i+1}^n)\big) \cdot \frac{\partial \Psi_{N+1}}{\partial s}(s; \widetilde{W}_i^n, \widetilde{W}_{i+1}^n)\, ds,$$

$$\mathcal{B}_{i+1/2} \cdot (W_{i+1}^n - W_i^n) = \int_0^1 \mathcal{B}\big(\widetilde{\Psi}(s; \widetilde{W}_i^n, \widetilde{W}_{i+1}^n)\big) \cdot \frac{\partial \Psi}{\partial s}(s; \widetilde{W}_i^n, \widetilde{W}_{i+1}^n)\, ds,$$

$$S_{i+1/2}(\sigma_{i+1} - \sigma_i) = \int_0^1 S\big(\widetilde{\Psi}(s; \widetilde{W}_i^n, \widetilde{W}_{i+1}^n)\big) \cdot \frac{\partial \Psi_{N+1}}{\partial s}(s; \widetilde{W}_i^n, \widetilde{W}_{i+1}^n)\, ds.$$

Then, it can be easily shown (see [28]) that the matrix

$$\widetilde{\mathcal{A}}_{i+1/2} = \left[\begin{array}{c|c} \mathcal{A}_{i+1/2} & -S_{i+1/2} \\ \hline 0 & 0 \end{array} \right],$$

where

$$\mathcal{A}_{i+1/2} = \mathcal{J}_{i+1/2}^{\sigma_{i+1/2}} - \mathcal{B}_{i+1/2},$$

is a Roe matrix provided that it has $N+1$ distinct real eigenvalues.

Once the Roe matrices have been calculated, the reconstructions are added to go to higher order. We will use the following notation:

$$\widetilde{P}_i^t = \begin{bmatrix} P_i^t \\ p_{i,N+1}^t \end{bmatrix} = \begin{bmatrix} p_{i,1}^t \\ \vdots \\ p_{i,N}^t \\ p_{i,N+1}^t \end{bmatrix}.$$

Some straightforward calculations allow us to rewrite the scheme (3.6) under a form closer to that of WENO-Roe methods for conservation laws:

(5.5)

$$\begin{aligned}
W_i' = & \frac{\Delta t}{\Delta x}\left(\widetilde{G}_{i-1/2} - \widetilde{G}_{i+1/2}\right) \\
& + \frac{\Delta t}{2\Delta x}\left(\mathcal{B}_{i-1/2} \cdot (W_{i-1/2}^+ - W_{i-1/2}^-) + \mathcal{B}_{i+1/2} \cdot (W_{i+1/2}^+ - W_{i+1/2}^-)\right) \\
& + \frac{\Delta t}{\Delta x}\left(\mathcal{P}_{i-1/2}^+ S_{i-1/2}(\sigma_{i-1/2}^+ - \sigma_{i-1/2}^-) + \mathcal{P}_{i+1/2}^- S_{i+1/2}(\sigma_{i+1/2}^+ - \sigma_{i+1/2}^-)\right) \\
& + \frac{\Delta t}{2\Delta x}\left(V_{i-1/2} + V_{i+1/2}\right) \\
& + \frac{\Delta t}{\Delta x}\left(\mathcal{I}_{B,i} + \mathcal{I}_{S,i}\right),
\end{aligned}$$

where

$$\begin{aligned}
\widetilde{G}_{i+1/2} = & \frac{1}{2}\left(F(W_{i+1/2}^-, \sigma_{i+1/2}^-) + F(W_{i+1/2}^+, \sigma_{i+1/2}^+)\right) \\
& - \frac{1}{2}\left|\mathcal{A}_{i+1/2}\right| \cdot (W_{i+1/2}^+ - W_{i+1/2}^-)
\end{aligned}$$

and

$$\mathcal{P}_{i+1/2}^\pm = \frac{1}{2}\left(\mathcal{I} \pm \left|\mathcal{A}_{i+1/2}\right| \mathcal{A}_{i+1/2}^{-1}\right).$$

These latter matrices can be also be written under the form

$$\mathcal{P}_{i+1/2}^\pm = \frac{1}{2}\mathcal{K}_{i+1/2}\left(\mathcal{I} \pm \mathrm{sgn}(\mathcal{L})_{i+1/2}\right)\mathcal{K}_{i+1/2}^{-1},$$

where $\mathcal{K}_{i+1/2}$ is the $N \times N$ matrix whose columns are the eigenvectors $R_{i+1/2,1}, \ldots, R_{i+1/2,N}$ and $\mathrm{sgn}(\mathcal{L})_{i+1/2}$ is the diagonal matrix whose coefficients are the signs of the eigenvalues $\lambda_{i+1/2,1}, \ldots, \lambda_{i+1/2,N}$. Besides,

$$\begin{aligned}
V_{i+1/2} = & F(W_{i+1/2}^+, \sigma_{i+1/2}^+) - F(W_{i+1/2}^+, \sigma_{i+1/2}) \\
& + F(W_{i+1/2}^-, \sigma_{i+1/2}) - F(W_{i+1/2}^-, \sigma_{i+1/2}^-),
\end{aligned}$$

or, equivalently (see (5.4)),

$$V_{i+1/2} = \int_0^1 \frac{\partial F}{\partial \sigma}\left(\widetilde{\Psi}(s; \widetilde{W}_{i+1/2}^-, \widetilde{W}_{i+1/2}^+)\right) \cdot \frac{\partial \Psi_{N+1}}{\partial s}(s; \widetilde{W}_{i+1/2}^-, \widetilde{W}_{i+1/2}^+) \, ds.$$

Finally,

$$\mathcal{I}_{B,i} = \int_{x_{i-1/2}}^{x_{i+1/2}} \mathcal{B}(\widetilde{P}_i^t(x)) \frac{d}{dx} P_i^t(x)\, dx,$$

$$\mathcal{I}_{S,i} = \int_{x_{i-1/2}}^{x_{i+1/2}} S(\widetilde{P}_i^t(x)) \frac{d}{dx} p_{i,N+1}^t(x)\, dx.$$

Remark 5.2. In this context, the meaning of the well-balanced property of the reconstruction operator can be understood as follows: let us suppose, as in Remark 5.1, that an integral curve γ of the linearly degenerate field can be described by a system of equations (5.3). Let us suppose that $\widetilde{W}(x) = (w_1(x), \ldots, w_N(x), \sigma(x))$ is a stationary solution such that $\widetilde{W}(x) \in \gamma$ for all x, i.e.,

$$g_j(w_1(x), \ldots, w_N(x), \sigma(x)) = 0, \quad j = 1, \ldots N, \quad \forall\, x.$$

If we now apply a well-balanced reconstruction operator to the sequence $\{\widetilde{W}(x_i)\}$, then the approximation functions $\widetilde{P}_i$ have to satisfy

$$g_j(p_{i,1}(x), \ldots, p_{i,N}(x), p_{i,N+1}(x)) = 0, \quad j = 1, \ldots N, \quad \forall\, x \in I_i.$$

6. WENO-Roe methods

In this section we consider numerical schemes of the form (3.6), in which the approximation functions used in the reconstruction operator are built by means of a WENO interpolation procedure using stencils with r points; we denote this method simply as r-WENO, and the resulting scheme as r-WENO-Roe. For the details about WENO interpolation, see [23], [27], [33], [34]. The reconstructions proposed in the r-WENO method are as follows:

$$W_{i+1/2}^- = \sum_{k=0}^{r-1} \omega_k^- q_k(x_{i+1/2}), \quad W_{i-1/2}^+ = \sum_{k=0}^{r-1} \omega_k^+ q_k(x_{i-1/2}),$$

where each q_k is the derivative of an interpolation polynomial that uses the values of the sequence W_i^n at the stencil

$$S_k^r = \{x_{i-k}, \ldots, x_{i-k+r-1}\}.$$

The weights $\omega_k^{\pm}$ satisfy

$$w_k^{\pm} \geq 0, \quad \sum_{k=0}^{r-1} w_k^{\pm} = 1.$$

These weights are calculated so that, on the one hand, the reconstruction operator is of order $2r - 1$ and, on the other hand, the weight ω_k is near to zero when the data on the stencil S_k^r indicate the presence of a discontinuity.

In order to construct the approximation function at the cells, let us first define

$$P_{i+1/2}^-(x) = \sum_{k=0}^{r-1} \omega_k^- q_k(x), \quad P_{i-1/2}^+(x) = \sum_{k=0}^{r-1} \omega_k^+ q_k(x)$$

(see Figure 1).

We have to define a function P_i at the cell I_i satisfying

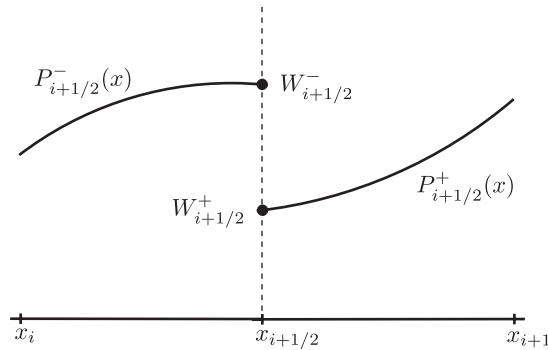$$\lim_{x \to x_{i-1/2}^+} P_i(x) = W_{i-1/2}^+, \quad \lim_{x \to x_{i+1/2}^-} P_i(x) = W_{i+1/2}^-.$$

FIGURE 1. Approximation functions $P^{\pm}_{i+1/2}(x)$.

A first possibility is given by

(6.1)
$$P_i(x) = \begin{cases} P^+_{i-1/2}(x) & \text{if } x \in [x_{i-1/2}, x_i), \\ P^-_{i+1/2}(x) & \text{if } x \in (x_i, x_{i+1/2}]. \end{cases}$$

This first definition does not fit into the framework defined in Section 3, as P_i is, in general, discontinuous:

$$W^-_i = P^+_{i-1/2}(x_i) \neq P^-_{i+1/2}(x_i) = W^+_i.$$

Due to this fact, if a WENO-Roe scheme (3.6) is used to design a high order numerical method for a problem of the form (5.1), when the numerical scheme is written under the form (5.5), an extra term has to be added at the right-hand side:

$$F(W^+_i) - F(W^-_i).$$

Nevertheless, this difference of fluxes is of order r, and it can be neglected.

A second definition avoiding this discontinuity is the following:

(6.2)
$$P_i(x) = \frac{1}{\Delta x}\big((x_{i+1/2} - x)P^+_{i-1/2}(x) + (x - x_{i-1/2})P^-_{i+1/2}(x)\big).$$

Due to the definition of the reconstruction operator, the functions P_i given by (6.1) or (6.2) provide only approximations of order r at the interior points of the cells, while their derivatives give approximations of order $r-1$. Therefore, applying Theorem 3.2, the method (3.6) has only order r, while it has order $2r-1$ when it is applied to systems of conservation laws.

Remark 6.1. If, instead of a WENO method the r-ENO reconstruction operator is chosen, the expected order of the numerical scheme is r, since in this case the approximation functions coincide with interpolation polynomials constructed on the basis of stencils with r points. Nevertheless, as commented in [34], the use of WENO approximations has several advantadges: it gives smoother operators, it is less sensible to round-off errors, and it avoids the use of conditionals in its practical implementation, being optimal for the vectorization of the algorithms.

It is however possible, performing some slight modifications on the WENO interpolation procedure, to obtain a method of order $2r-1$. The idea is as follows:

instead of choosing the usual WENO reconstructions we consider the functions

$$\widetilde{P}^-_{i+1/2}(x) = \sum_{k=0}^{r-1} \widetilde{\omega}^-_k(x) q_k(x), \quad \widetilde{P}^+_{i-1/2}(x) = \sum_{k=0}^{r-1} \widetilde{\omega}^+_k(x) q_k(x),$$

where the weights now depend on x and are calculated following the usual procedure in WENO reconstruction, so that the order of approximation is $2r - 1$ in the cell. Unfortunately, the derivatives of these approximation functions are not easy to obtain. Instead, we substitute these derivatives by new WENO approximation functions

$$\widetilde{Q}^-_{i+1/2}(x) = \sum_{k=0}^{r-1} \widetilde{\gamma}^-_k(x) q'_k(x), \quad \widetilde{Q}^+_{i-1/2}(x) = \sum_{k=0}^{r-1} \widetilde{\gamma}^+_k(x) q'_k(x),$$

where, again, the weights are calculated, for every x, following the usual procedure in WENO reconstruction. Therefore, we again obtain order $2r - 2$ in the cell.

Once these functions have been defined, we introduce the new approximation functions at the cells given either by

$$\widetilde{P}_i(x) = \begin{cases} \widetilde{P}^+_{i-1/2}(x) & \text{if } x \in [x_{i-1/2}, x_i), \\ \widetilde{P}^-_{i+1/2}(x) & \text{if } x \in (x_i, x_{i+1/2}], \end{cases}$$

$$\widetilde{Q}_i(x) = \begin{cases} \widetilde{Q}^+_{i-1/2}(x) & \text{if } x \in [x_{i-1/2}, x_i), \\ \widetilde{Q}^-_{i+1/2}(x) & \text{if } x \in (x_i, x_{i+1/2}], \end{cases}$$

or

$$\widetilde{P}_i(x) = \frac{1}{\Delta x}\big((x_{i+1/2} - x)\widetilde{P}^+_{i-1/2}(x) + (x - x_{i-1/2})\widetilde{P}^-_{i+1/2}(x)\big),$$

$$\widetilde{Q}_i(x) = \frac{1}{\Delta x}\big((x_{i+1/2} - x)\widetilde{Q}^+_{i-1/2}(x) + (x - x_{i-1/2})\widetilde{Q}^-_{i+1/2}(x)\big),$$

depending on the chosen approach.

Once these functions have been defined, the integral appearing in (3.6) is replaced by

$$\int_{x_{i-1/2}}^{x_{i+1/2}} \mathcal{A}(\widetilde{P}_i(x))\widetilde{Q}_i(x)\,dx.$$

In practice, this integral is approached by means of a Gaussian quadrature of order at least $2r-1$. As a consequence, the weights $\omega^\pm(x)$ and $\gamma^\pm(x)$ have to be calculated only at the quadrature points.

Following the same steps as in the case of the r-WENO-Roe method, it can be easily shown that the resulting scheme (that will be denoted as *modified r-WENO-Roe*) is well balanced with order $2r - 1$.

The computational cost of this modified numerical scheme is higher than those corresponding to standard WENO reconstructions, as two set of weights have to be calculated at every quadrature point. Moreover, the positivity of the weights is only ensured at the intercells, due to the choice of stencils. Therefore, in some cases negative weights may appear at interior quadrature points giving rise to oscillations and instabilities. For handling these negative weights, if necessary, the splitting technique of Shi, Hu and Shu ([32]) can be applied. However, in some cases (see, e.g., Section 7.7) this technique does not completely remove the oscillations, and the scheme eventually crashes. The causes of this problem are currently under investigation.

We finish this section with a remark about time-stepping. As is usual in WENO interpolation based schemes, in order to obtain a full high resolution scheme it is necessary to use a high order method to advance in time. In the schemes considered here we have taken optimal high order TVD Runge-Kutta schemes ([19], [33]).

6.1. Shallow-water equations with depth variations.
The equations governing the flow of a shallow-water layer of fluid through a straight channel with constant rectangular cross-section can be written as

$$
(6.3) \qquad
\begin{cases}
\dfrac{\partial h}{\partial t} + \dfrac{\partial q}{\partial x} = 0, \\[2mm]
\dfrac{\partial q}{\partial t} + \dfrac{\partial}{\partial x}\left(\dfrac{q^2}{h} + \dfrac{g}{2}h^2 \right) = gh\dfrac{dH}{dx}.
\end{cases}
$$

The variable x makes reference to the axis of the channel and t is time, $q(x,t)$ and $h(x,t)$ represent the mass-flow and the thickness, respectively, g is gravity, and $H(x)$ is the depth function measured from a fixed level of reference. The fluid is supposed to be homogeneous and inviscid.

The system (6.3) can be rewritten under the form (5.1) with $N = 2$,

$$
W = \begin{bmatrix} h \\ q \end{bmatrix}, \quad
F(W) = \begin{bmatrix} q \\ \dfrac{q^2}{h} + \dfrac{g}{2}h^2 \end{bmatrix}, \quad
S(W) = \begin{bmatrix} 0 \\ -gh \end{bmatrix},
$$

$\mathcal{B} = 0$ and $\sigma = H$. Observe that, in this case, the flux and the coefficients of the source term do not depend on σ.

We can also write system (6.3) under the nonconservative form (5.2) with

$$
\widetilde{W} = \begin{bmatrix} h \\ q \\ H \end{bmatrix}, \quad
\widetilde{\mathcal{A}}(\widetilde{W}) = \begin{bmatrix} 0 & 1 & 0 \\ -u^2 + c^2 & 2u & -c^2 \\ 0 & 0 & 0 \end{bmatrix},
$$

where $u = q/h$ is the averaged velocity and $c = \sqrt{gh}$.

If the family of segments (2.10) is chosen as the family of paths, a family of Roe matrices for system (6.3) is given by (see [28])

$$
\widetilde{\mathcal{A}}(\widetilde{W}_0, \widetilde{W}_1) = \begin{bmatrix} 1 & 0 & 0 \\ -\tilde{u}^2 + \tilde{c}^2 & 2\tilde{u} & -\tilde{c}^2 \\ 0 & 0 & 0 \end{bmatrix},
$$

where

$$
\tilde{u} = \frac{\sqrt{h_0}\,u_0 + \sqrt{h_1}\,u_1}{\sqrt{h_0} + \sqrt{h_1}}, \qquad
\tilde{c} = \sqrt{g\frac{h_0 + h_1}{2}}.
$$

For system (6.3), stationary solutions are given by

$$
(6.4) \qquad q = q_0, \quad h + \frac{q_0^2}{2gh^2} - H = C,
$$

where q_0 and C are constants. In the particular case of water at rest, we have the solutions

$$
(6.5) \qquad q = 0, \quad h - H = C.
$$

Therefore, solutions corresponding to still water define straight lines in the h-q-H space. As a consequence, Roe methods based on the family of segments are exactly well balanced for still-water solutions and well balanced with order 2 for general stationary solutions (see [2], [28]).

The reconstruction operator proposed here to get higher order schemes is based on WENO reconstruction related to the variables q, H and $\eta = h - H$ (this variable represents the water surface elevation). That is, given a sequence (q_i, h_i, H_i) we consider the new sequence (q_i, η_i, H_i) with $\eta_i = h_i - H_i$ and apply the r-WENO reconstruction operator to obtain polynomials

$$p^{\pm}_{i+1/2,q}, \quad p^{\pm}_{i+1/2,\eta}, \quad p^{\pm}_{i+1/2,H};$$

then, we define

$$p^{\pm}_{i+1/2,h} = p^{\pm}_{i+1/2,\eta} + p^{\pm}_{i+1/2,H}.$$

This reconstruction is exactly well balanced for stationary solutions corresponding to water at rest. In effect, if the sequence (q_i, h_i, H_i) lie on the curve defined by (6.5), then $q_i = 0$ and $\eta_i = C$. As a consequence,

$$p^{\pm}_{i+1/2,q} \equiv 0, \quad p^{\pm}_{i+1/2,\eta} \equiv C,$$

so we have

$$p^{\pm}_{i+1/2,q} \equiv 0, \quad p^{\pm}_{i+1/2,h} - p^{\pm}_{i+1/2,H} = C,$$

and thus the reconstruction operator is well balanced (see Remark 5.2).

Applying Theorems 4.3 and 4.4, we deduce that the corresponding WENO-Roe schemes satisfy the $\mathcal{C}$-property, i.e., they are exactly well balanced for still-water solutions, and well balanced with order r for general stationary solutions. To obtain a well-balanced numerical scheme with order $2r-1$, we have to add to the numerical scheme the modifications proposed in Section 6.

6.2. **The two-layer shallow-water system.** We now consider the equations of a one-dimensional flow of two superposed inmiscible layers of shallow-water fluids studied in [5]:

$$(6.6) \qquad \begin{cases} \dfrac{\partial h_1}{\partial t} + \dfrac{\partial q_1}{\partial x} = 0, \\ \dfrac{\partial q_1}{\partial t} + \dfrac{\partial}{\partial x}\left(\dfrac{q_1^2}{h_1} + \dfrac{g}{2}h_1^2 \right) = -gh_1 \dfrac{\partial h_2}{\partial x} + gh_1 \dfrac{dH}{dx}, \\ \dfrac{\partial h_2}{\partial t} + \dfrac{\partial q_2}{\partial x} = 0, \\ \dfrac{\partial q_2}{\partial t} + \dfrac{\partial}{\partial x}\left(\dfrac{q_2^2}{h_2} + \dfrac{g}{2}h_2^2 \right) = -\dfrac{\rho_1}{\rho_2}gh_2 \dfrac{\partial h_1}{\partial x} + gh_2 \dfrac{dH}{dx}. \end{cases}$$

In the equations, index 1 refers to the upper layer and index 2 to the lower one. We assume that the fluid occupies a straight channel with constant rectangular cross-section and constant width. The variable x refers to the axis of the channel, t is time, g is gravity and $H(x)$ is the depth function measured from a fixed level of reference. The constants ρ_1 and ρ_2 are the densities of each layer, where it is supposed that $\rho_1 < \rho_2$. Finally, $h_i(x,t)$ and $q_i(x,t)$ are, respectively, the thickness and the mass-flow of the ith layer at the section of coordinate x at time t.

Problem (6.6) can be written again under the form (5.1) with $N = 4$, $\sigma = H$,

$$W = \begin{bmatrix} h_1 \\ q_1 \\ h_2 \\ q_2 \end{bmatrix}, \quad F(W) = \begin{bmatrix} q_1 \\ \dfrac{q_1^2}{h_1} + \dfrac{g}{2}h_1^2 \\ q_2 \\ \dfrac{q_2^2}{h_2} + \dfrac{g}{2}h_2^2 \end{bmatrix}, \quad S(W) = \begin{bmatrix} 0 \\ gh_1 \\ 0 \\ gh_2 \end{bmatrix},$$

and

$$\mathcal{B}(W) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -gh_1 & 0 \\ 0 & 0 & 0 & 0 \\ -\rho gh_2 & 0 & 0 & 0 \end{bmatrix},$$

where $\rho = \rho_1/\rho_2$.

Also it can be put into the nonconservative form (5.2) with

$$\widetilde{W} = \begin{bmatrix} h_1 \\ q_1 \\ h_2 \\ q_2 \\ H \end{bmatrix}, \quad \widetilde{\mathcal{A}}(\widetilde{W}) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ -u_1^2 + c_1^2 & 2u_1 & c_1^2 & 0 & -c_1^2 \\ 0 & 0 & 0 & 1 & 0 \\ \rho c_2^2 & 0 & -u_2^2 + c_2^2 & 2u_2 & -c_2^2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

where $u_i = q_i/h_i$ and $c_i = \sqrt{gh_i}$.

The stationary solutions of system (6.6) are given by

(6.7)
$$\begin{cases} q_1 = \text{constant}, \\ \dfrac{u_1^2}{2} - \dfrac{u_2^2}{2} + g_\rho h_1 = \text{constant}, \\ q_2 = \text{constant}, \\ \dfrac{u_1^2}{2} + g(h_1 + h_2 - H) = \text{constant}, \end{cases}$$

where $g_\rho = (1 - \rho)g$ is the reduced gravity.

In [28] Roe matrices for system (6.6), based on the choice of paths as segments, were constructed. The resulting Roe scheme was proved to be well balanced with order 2 for general stationary solutions, and exactly well balanced for solutions representing water at rest or vacuum.

For the reconstruction process, a strategy similar to that in Section 6.1 is followed. Specifically, we reconstruct the elevations

$$\eta_1 = h_1 + h_2 - H, \quad \eta_2 = h_2 - H,$$

instead of the thickness h_1 and h_2. As a consequence, the resulting WENO-Roe scheme (3.6) is exactly well balanced for solutions representing water at rest or vacuum, and well balanced with order r (or $2r - 1$ if the modified WENO method is used) for general stationary solutions.

7. Numerical results

7.1. Verification of the $\mathcal{C}$-property. The objective of this section is to test the $\mathcal{C}$-property of the scheme (3.6) stated in Section 6.1. We have considered two examples, corresponding to one-layer and two-layer shallow-water systems.

We first consider the one-layer system given by (6.3). The depth function $H(x)$ is given by an exponential perturbed with a random noise (see Figure 2); as initial conditions we have taken $h(x,0) = H(x)$ and $q(x,0) = 0$, $x \in [0,1]$. The modified 3-WENO-Roe scheme has been applied with $\Delta x = 0.01$ and CFL number 0.9. For time-stepping, an optimal three-step TVD Runge-Kutta method ([19], [33]) has been considered.

As expected from Section 6.1, the numerical scheme (3.6) preserves the steady state solution exactly up to machine accuracy. This fact can be observed in Table 1.

TABLE 1. Verification of the $\mathcal{C}$-property: one layer.

Precision	L^1 error h	L^1 error q
single	7.33×10^{-7}	5.12×10^{-7}
double	1.28×10^{-15}	3.65×10^{-15}

TABLE 2. Verification of the $\mathcal{C}$-property: two layers.

Precision	L^1 error h_1	L^1 error q_1	L^1 error h_2	L^1 error q_2
single	8.21×10^{-7}	1.33×10^{-7}	6.27×10^{-7}	5.97×10^{-7}
double	1.42×10^{-15}	6.64×10^{-16}	2.47×10^{-15}	2.65×10^{-15}
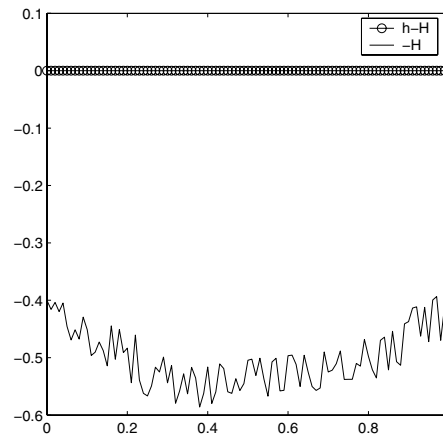
FIGURE 2. Stationary solution in test case 7.1 for the one-layer system. Elevation $\eta = h - H$ and bottom topography $-H$.

Next, we consider the two-layer shallow-water system (6.6) with the same depth function as before, and initial conditions $h_1(x,0) = 0.4$, $h_2(x,0) = H(x)$, $q_1(x,0) = q_2(x,0) = 0$, $x \in [0,1]$; the ratio of densities is $\rho = 0.99805$. In Table 2 we show the results obtained with the modified 3-WENO-Roe method, with the same settings as in the previous example.

In both examples, similar results are obtained when the 3-WENO-Roe method is applied.

7.2. **Well-balancing test: Steady subcritical flow.** The purpose of this experiment is to verify numerically the well-balanced property of scheme (3.6). We consider the shallow-water system (6.3) with the depth function given by

$$H(x) = 2 - 0.2\,e^{-0.16(x-10)^2}, \quad x \in [0, 20],$$

and the initial condition corresponding to the steady subcritical flow with discharge $q(x,0) = 4.42$. The solution is represented in Figure 3, and it can be computed analytically using (6.4). This solution should be preserved.

TABLE 3. Test case 7.2 solved with the 3-WENO-Roe method.

N. cells	L^1 error h	L^1 order h	L^1 error q	L^1 order q
20	6.11×10^{-4}	–	4.48×10^{-3}	–
40	5.14×10^{-5}	3.57	2.77×10^{-4}	4.01
80	1.44×10^{-6}	5.15	2.03×10^{-5}	3.77
160	2.30×10^{-8}	5.96	2.61×10^{-6}	2.96
320	1.99×10^{-9}	3.55	3.34×10^{-7}	2.96

TABLE 4. Test case 7.2 solved with the modified 3-WENO-Roe method.

N. cells	L^1 error h	L^1 order h	L^1 error q	L^1 order q
20	2.55×10^{-3}	–	4.48×10^{-2}	–
40	1.01×10^{-4}	4.66	1.67×10^{-3}	4.75
80	1.84×10^{-6}	5.78	4.34×10^{-5}	5.26
160	3.78×10^{-8}	5.61	1.54×10^{-6}	4.81
320	7.29×10^{-10}	5.69	5.01×10^{-8}	4.95
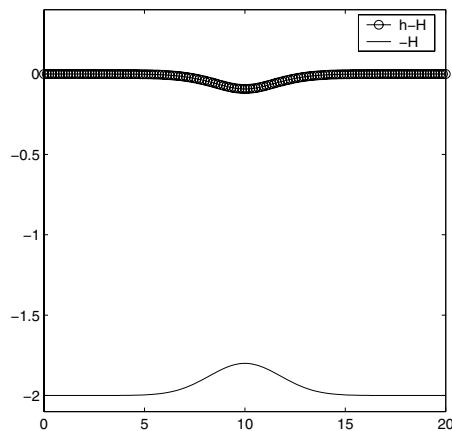
FIGURE 3. Stationary solution in test case 7.2. Elevation $\eta = h - H$ and bottom topography $-H$.

In the experiment, we have taken CFL coefficient 0.9 and WENO reconstructions with $r = 3$. The integral terms have been approximated by means of a Gaussian quadrature with three points. To advance in time, an optimal three-step TVD Runge-Kutta method has been applied.

Following Section 6.1, the 3-WENO-Roe scheme (3.6) should be well balanced with order 3, and the modified 3-WENO-Roe scheme with order 5. The numerical results obtained with each scheme can be seen in Tables 3 and 4, respectively.

In this case, the negative weights appearing in the modified 3-WENO-Roe method do not cause any instability, so no special treatment has been applied to handle them.

TABLE 5. Test case 7.3 solved with the 3-WENO-Roe method.

N. cells	L^1 error h	L^1 order h	L^1 error q	L^1 order q
50	1.66×10^0	–	9.78×10^{-1}	–
100	3.84×10^{-1}	2.11	2.33×10^{-1}	2.08
200	7.69×10^{-2}	2.32	4.83×10^{-2}	2.26
400	1.32×10^{-2}	2.54	8.71×10^{-3}	2.47
800	1.74×10^{-3}	2.92	1.19×10^{-3}	2.88
1600	2.17×10^{-4}	3.01	1.51×10^{-4}	2.97

TABLE 6. Test case 7.3 solved with the modified 3-WENO-Roe method.

N. cells	L^1 error h	L^1 order h	L^1 error q	L^1 order q
50	1.51×10^0	–	6.75×10^{-1}	–
100	1.96×10^{-1}	2.95	1.02×10^{-1}	2.72
200	2.68×10^{-2}	2.87	1.80×10^{-2}	2.51
400	2.21×10^{-3}	3.60	1.50×10^{-3}	3.58
800	9.75×10^{-5}	4.50	6.80×10^{-5}	4.47
1600	3.28×10^{-6}	4.89	2.35×10^{-6}	4.85

7.3. **Accuracy test.** In order to verify numerically that the proposed schemes in this work are indeed high order accurate, here we consider a test problem with depth function

$$H(x) = 5 - \cos(\pi x/5)$$

and initial conditions

$$h(x,0) = 5 - \cos(\pi x/5) + 0.1 \sin(\pi x/5), \qquad q(x,0) = 0,$$

in the domain $x \in [0, 20]$, with periodic boundary conditions. The solution of this problem is smooth.

As the exact solution for this problem is not known, we use as a reference solution a numerical solution computed with the modified 3-WENO-Roe scheme with 25600 cells. In Tables 5 and 6, the L^1 errors obtained at time $t = 1$ with the 3-WENO-Roe and the modified 3-WENO-Roe schemes, respectively, are shown. As can be observed, the schemes give the order of accuracy predicted by the theory.

Again, in this case, the negative weights appearing in the modified 3-WENO-Roe method do not cause any instability, so no special treatment has been applied to handle them.

7.4. **Steady flow over a bump.** The test problem analyzed in this section is the classical one of a steady flow over a bump in a rectangular channel with constant breadth [37], [38].

System (6.3) is considered in the computational domain $[0, 20]$ with depth function given by

$$H(x) = \begin{cases} 0.05(x - 10)^2 & \text{if } 8 \le x \le 12, \\ 0.2 & \text{otherwise,} \end{cases}$$

TABLE 7. Test case 7.4: comparison with respect to the exact solution.

N. cells	L^1 error h	L^1 error q
100	1.28×10^{-2}	1.43×10^{-2}
200	4.78×10^{-3}	1.44×10^{-3}
300	4.29×10^{-3}	3.57×10^{-3}
400	2.24×10^{-3}	2.74×10^{-3}
500	1.83×10^{-3}	4.84×10^{-4}

(a) Elevation $\eta = h - H$ and bottom $-H$.
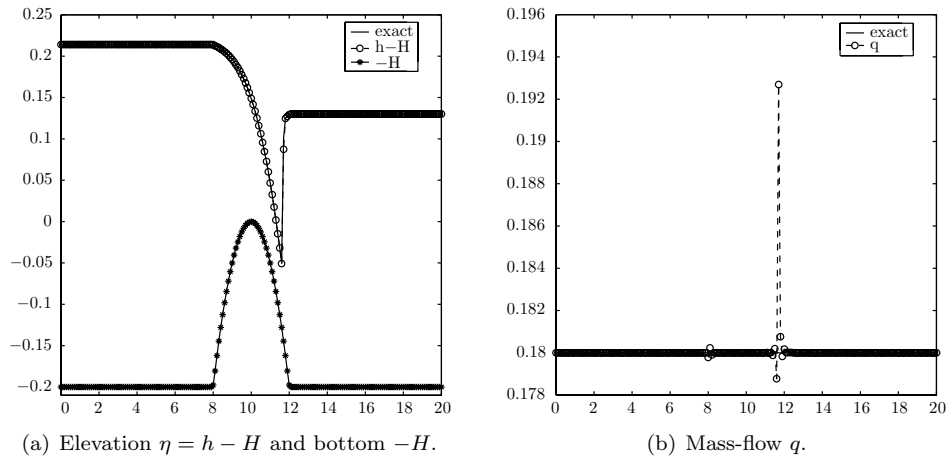
(b) Mass-flow q.

FIGURE 4. Hydraulic jump over a bump. Comparison between the solution computed with the modified 3-WENO-Roe method and the exact solution, at time $T = 200$.

and initial conditions $q(x, 0) = 0$, $h(x, 0) = 0.13 + H(x)$. With respect to the boundary conditions, a water thickness of $h = 0.33$ is imposed downstream and a mass-flow of $q = 0.18$ upstream. The solution of this problem consists of a steady transcritical flow with a smooth transition followed by a hydraulic jump.

We have considered space step $\Delta x = 0.1$ and the same settings as in Section 7.2. The results obtained with the modified 3-WENO-Roe scheme at time $T = 200$, where the steady state has been reached, are shown in Figure 4.

We have also compared the solution obtained at time $T = 200$ with the exact steady state solution of the problem (see [20]). In Table 7 the L^1 errors obtained with a different number of cells are shown.

Observe that, in this case, as the solution is not smooth, the high order convergence is not achieved. As in previous sections, no special treatment of negative weights is needed, as the modified WENO scheme has a good behavior.

7.5. **Small perturbation of steady state water.** In order to test the performances of our schemes on a rapidly varying flow over a smooth bed, we consider a problem proposed by LeVeque in [26]. Specifically, a steady state solution is perturbed by a pulse that splits into two waves propagating in opposite directions over a continuous bed. The left-going wave travels over a horizontal bottom while the right-going wave propagates over a bump.
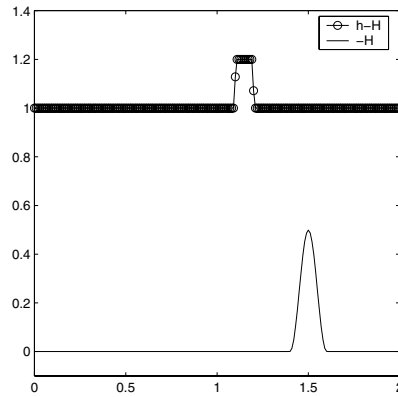
FIGURE 5. Initial condition in LeVeque's problem with $\Delta h = 0.2$.

System (6.3) is considered on the computational domain $[0, 2]$. The depth function is given by

$$H(x) = \begin{cases} -0.25(\cos(10\pi(x - 0.5)) + 1) & \text{if } 1.4 \leq x \leq 1.6, \\ 0 & \text{otherwise,} \end{cases}$$

while the initial conditions are $q(x, 0) = 0$ and

$$h(x, 0) = \begin{cases} 1 + \Delta h + H(x) & \text{if } 1.1 \leq x \leq 1.2, \\ 1 + H(x) & \text{otherwise.} \end{cases}$$

Here Δh is the height of the perturbation that takes the values $\Delta h = 0.2$ (big pulse; see Figure 5) or $\Delta h = 0.001$ (small pulse). Outflow boundary conditions have been considered.

(a) Elevation $\eta = h - H$ and bottom $-H$.
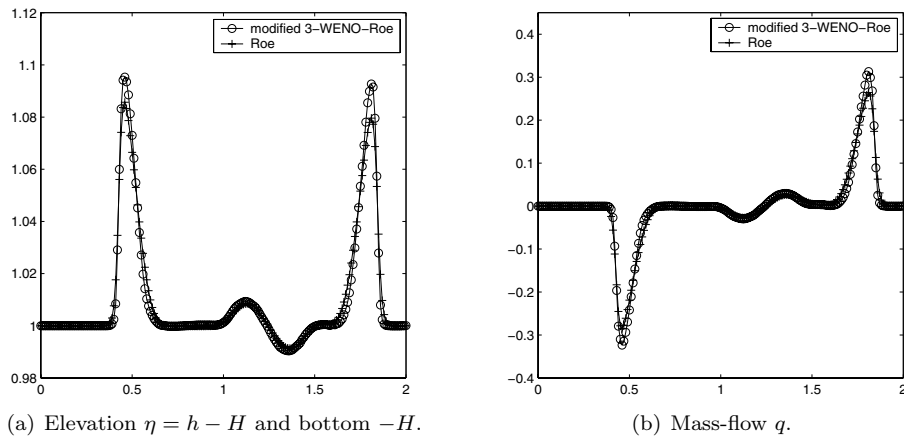

(b) Mass-flow q.

FIGURE 6. LeVeque's problem with big pulse $\Delta h = 0.2$. Comparison between the modified 3-WENO-Roe method and the Roe method at time $T = 0.2$.
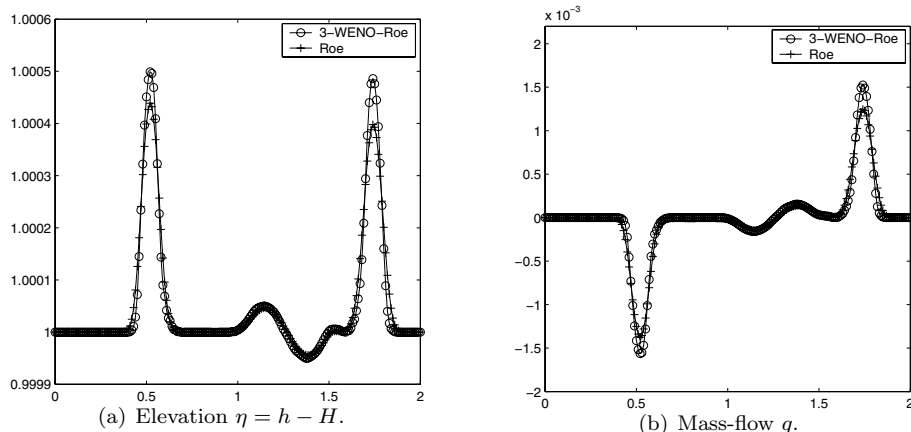
FIGURE 7. LeVeque's problem with small pulse $\Delta h = 0.01$. Comparison between the 3-WENO-Roe method and the Roe method at time $T = 0.2$.
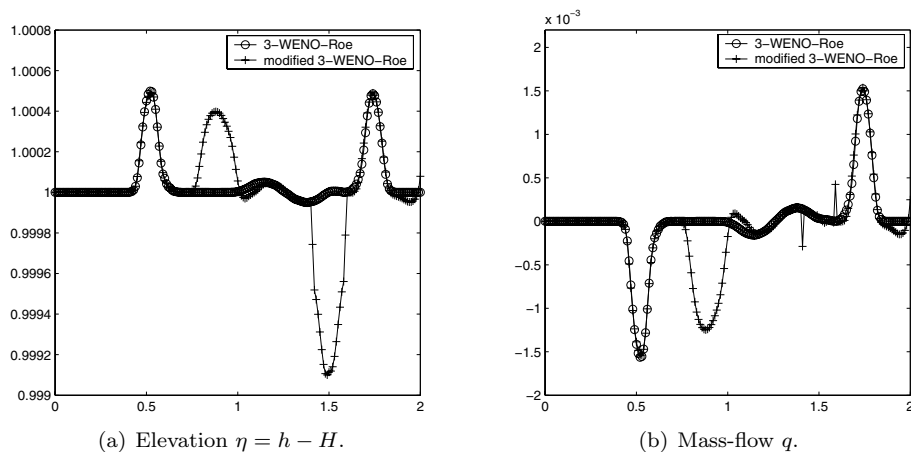

FIGURE 8. LeVeque's problem with small pulse $\Delta h = 0.01$. Comparison between the 3-WENO-Roe methods with and without modifications. Note that the modified 3-WENO-Roe method produces incorrect results

The computations have been performed with $\Delta x = 0.01$, CFL coefficient 0.9, WENO interpolation with $r = 3$ and three-step TVD Runge-Kutta time integration. In this case, the use of the modified WENO procedure gives rise to incorrect results when the scheme is applied to the problem with small pulse, even if the technique of Shi et al. ([32]) is applied; see Figure 8. Thus, for the small pulse problem we only consider the 3-WENO-Roe method.

Remark 7.1. Note that the numerical treatment of the source term is identical in both the 3-WENO-Roe and the modified 3-WENO-Roe schemes. As the first scheme works well for this problem, we think that the incorrect results produced by the modified scheme are due to the appearance of negative weights. Although

negative weights are always present in the modified 3-WENO-Roe scheme, their influence in the problem with the small pulse seems to be stronger than in the big pulse case. The cause of this fact is currently under study.

The results obtained at time $T = 0.2$ with the scheme (3.6) are shown in Figures 6 and 7, for $\Delta h = 0.2$ and $\Delta h = 0.01$, respectively. We have compared these solutions with that produced by the first order Roe method.

7.6. **Well-balancing test for the two-layer system.** In this section, we test the well-balanced property of scheme (3.6) when it is applied to the two-layer shallow-water system introduced in Section 6.2. We consider the steady state solution (6.7) with depth function

$$H(x) = 2 - 0.5e^{-x^2}, \quad x \in [-3, 3].$$

To compute the constants in (6.7), we consider the initial data

$$h_1(0, 0) = 0.5, \quad h_2(0, 0) = H(0) - 0.5, \quad q_1(0, 0) = 0.15 \quad \text{and} q_2(0, 0) = -0.15.$$

The density ratio has been taken as $\rho = 0.98$. The solution is shown in Figure 9.

As in Section 7.2, we have applied both the 3-WENO-Roe and the modified 3-WENO-Roe methods, a three-step TVD Runge-Kutta method, and a Gaussian quadrature with three points. The CFL coefficient has been taken as 0.9. The results obtained with the 3-WENO-Roe method are shown in Tables 8 and 9, while those corresponding to the modified 3-WENO-Roe method are given in Tables 10 and 11. It can be seen that the schemes are well balanced with the predicted order.

Again, no special treatment of negative weights in the WENO procedure was needed.

In this case, the exact solution is not easy to compute because a nonlinear system of algebraic equations must be solved.
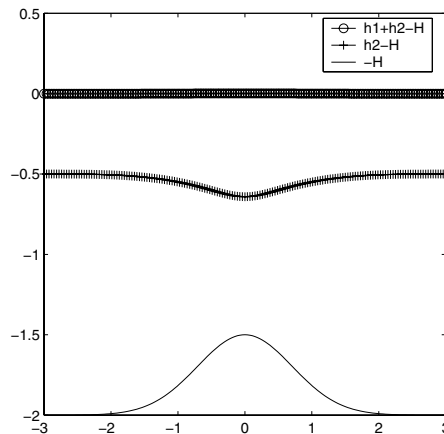

FIGURE 9. Stationary solution in test case 7.6. Elevations $\eta_1 = h_1 + h_2 - H$, $\eta_2 = h_2 - H$ and bottom topography $-H$.

TABLE 8. Test case 7.6 solved with the 3-WENO-Roe method. h_1 and h_2.

N. cells	L^1 error h_1	L^1 order h_1	L^1 error h_2	L^1 order h_2
20	1.20×10^{-3}	–	1.50×10^{-3}	–
40	1.20×10^{-4}	3.32	1.18×10^{-4}	3.66
80	1.22×10^{-5}	3.30	1.28×10^{-5}	3.21
160	1.52×10^{-6}	3.00	1.41×10^{-6}	3.17

TABLE 9. Test case 7.6 solved with the 3-WENO-Roe method. q_1 and q_2.

N. cells	L^1 error q_1	L^1 order q_1	L^1 error q_2	L^1 order q_2
20	4.65×10^{-4}	–	2.64×10^{-4}	–
40	5.12×10^{-5}	3.18	6.12×10^{-6}	5.43
80	2.62×10^{-6}	4.28	6.01×10^{-7}	3.34
160	6.13×10^{-8}	5.42	5.65×10^{-8}	3.41

TABLE 10. Test case 7.6 solved with the modified 3-WENO-Roe method. h_1 and h_2.

N. cells	L^1 error h_1	L^1 order h_1	L^1 error h_2	L^1 order h_2
20	1.25×10^{-3}	–	1.67×10^{-3}	–
30	1.86×10^{-4}	4.69	1.84×10^{-4}	5.44
40	4.73×10^{-5}	4.76	4.65×10^{-5}	4.78
60	6.74×10^{-6}	4.81	6.61×10^{-6}	4.81
80	1.66×10^{-6}	4.87	1.63×10^{-6}	4.87

TABLE 11. Test case 7.6 solved with the modified 3-WENO-Roe method. q_1 and q_2.

N. cells	L^1 error q_1	L^1 order q_1	L^1 error q_2	L^1 order q_2
20	5.10×10^{-4}	–	2.48×10^{-4}	–
30	1.51×10^{-5}	8.68	1.90×10^{-5}	6.34
40	5.10×10^{-6}	3.77	6.09×10^{-6}	3.95
60	7.90×10^{-7}	4.60	9.45×10^{-7}	4.60
80	1.85×10^{-7}	5.05	2.27×10^{-7}	4.96

7.7. **Internal dam break.** We consider the two-layer shallow-water system in Section 6.2, with constant depth function $H(x) = 2$ and initial conditions given by

$$h_1(x,0) = \begin{cases} 1.8 & \text{if } x < 0, \\ 0.2 & \text{if } x \geq 0, \end{cases}$$

$h_2(x,0) = 2 - h_1(x,0)$ and $q_1(x,0) = q_2(x,0) = 0$ (see Figure 10(a)).

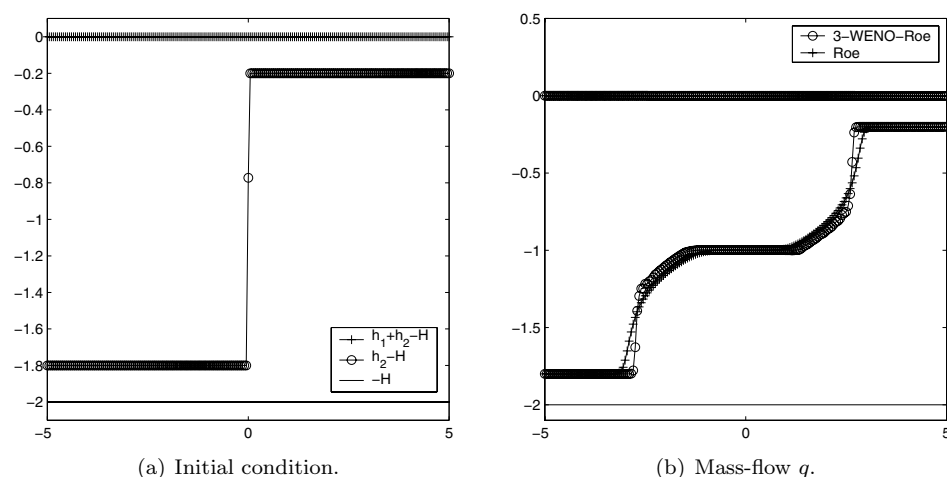(a) Initial condition.

(b) Mass-flow q.

FIGURE 10. Internal dam break problem. Elevations $\eta_1 = h_1 + h_2 - H$, $\eta_2 = h_2 - H$ and bottom topography $-H$.

In Figure 10(b), the results obtained at time $T = 25$ with the 3-WENO-Roe method and the Roe method in nonconservative form ([28]) are shown. The interface of the solution consist in three constant states, with two jumps and two rarefaction waves between them. We have considered $\Delta x = 0.05$, CFL coefficient 0.9 and density ratio $\rho = 0.99805$. As commented in Section 6, in this case the modified 3-WENO-Roe method produces oscillations and instabilities, even if the technique of [32] is applied, leading to a crash of the scheme at time $T \cong 4.5$. For this reason, we have considered only the 3-WENO-Roe method, so the accuracy reduces to third order. For time-stepping, we have again used a three-step TVD Runge-Kutta method.

REFERENCES

[1] N. Andrianov, *CONSTRUCT: a collection of MATLAB routines for constructing the exact solution to the Riemann problem for the shallow water equations*, avalaible at http://www-ian.math.unimagdeburg.de/home/andriano/CONSTRUCT.

[2] A. Bermúdez and M.E. Vázquez, *Upwind methods for hyperbolic conservation laws with source terms*. Comp. & Fluids **23** (1994), 1049–1971. MR1314237 (95i:76065)

[3] F. Bouchut, *Nonlinear Stability of Finite Volume Methods for Hyperbolic Conservation Laws and Well-Balanced Schemes for Sources*, Birkhäuser, 2004. MR2128209

[4] V. Caselles, R. Donat and G. Haro, *Flux-gradient and source term balancing for certain high resolution shock-capturing schemes*, submitted.

[5] M. J. Castro, J. Macías and C. Parés, *A Q-Scheme for a class of systems of coupled conservation laws with source term. Application to a two-layer 1-D shallow water system*. ESAIM: M2AN **35**(1) (2001), 107–127. MR1811983 (2001m:76063)

[6] T. Chacón, A. Domínguez and E.D. Fernández, *A family of stable numerical solvers for shallow water equations with source terms*. Comp. Meth. Appl. Mech. Eng. **192** (2003), 203–225. MR1951407 (2003m:76116)

[7] T. Chacón, A. Domínguez and E.D. Fernández, *An entropy correction free solver for non-homegeneous shallow water equations*. Math. Mod. Num. Anal. **37**(3) (2003), 755–772. MR2020863 (2004m:76021)

[8] T. Chacón, E.D. Fernández, M.J. Castro and C. Parés, *On well-balanced finite volume methods for non-homogeneous non-conservative hyperbolic systems*. Preprint, 2005.

[9] A. Chinnaya and A.Y. LeRoux, *A new general Riemann solver for the shallow-water equations with friction and topography*, available at `http://www.math.ntnu.no/conservation/1999/021.html`.

[10] A. Chinnaya and A.Y. LeRoux, *A well-balanced numerical scheme for the approximation of the shallow-water equations with topography: the resonance phenomenon*. To appear in Int. J. Finite Volume, 2004.

[11] G. Dal Maso, Ph. LeFloch and F. Murat, *Definition and weak stability of nonconservative products*. J. Math. Pures Appl. **74** (1995), 483–548.

[12] E.D. Fernández Nieto, *Aproximación Numérica de Leyes de Conservación Hiperbólicas No Homogéneas. Aplicación a las Ecuaciones de Aguas Someras*. Ph.D. thesis, Universidad de Sevilla, 2003.

[13] A.C. Fowler, *Mathematical Models in the Applied Sciences*, Cambridge, 1997. MR1483893 (98j:00012)

[14] P. García-Navarro and M.E. Vázquez-Cendón, *On numerical treatment of the source terms in the shallow water equations*. Comp. & Fluids **29**(8) (2000), 17–45.

[15] G. Godinaud, A.Y. LeRoux and M.N. LeRoux, *Generation of new solvers involving the source term for a class of nonhomogeneous hyperbolic systems*, available at `http://www.math.ntnu.no/conservation/2000/029.html`.

[16] E. Godlewski and P.A. Raviart, *Numerical Approximation of Hyperbolic Systems of Conservation Laws*, Springer, 1996. MR1410987 (98d:65109)

[17] L. Gosse, *A well-balanced flux-vector splitting scheme designed for hyperbolic systems of conservation laws with source terms*. Comput. Math. Appl. **39** (2000), 135–159. MR1753567 (2002a:65130)

[18] L. Gosse, *A well-balanced scheme using non-conservative products designed for hyperbolic systems of conservation laws with source terms*. Math. Models Methods Appl. Sci. **11** (2001), 339–365. MR1820677 (2002g:65097)

[19] S. Gottlieb and C.-W. Shu, *Total variation diminishing Runge-Kutta schemes*. Math. Comp. **67** (1998), 73–85. MR1443118 (98c:65122)

[20] N. Goutal and F. Maurel, *Proceedings of the Second Workshop on Dam-Break Wave Simulation*, Technical Report HE-43/97/016/A, Electrité de France, Département Laboratoire National d'Hydraulique, Groupe Hydraulique Fluviale, 1997.

[21] 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), 1–16. MR1377240 (97c:65144)

[22] J.M. Greenberg, A.Y. LeRoux, R. Baraille and A. Noussar, *Analysis and approximation of conservation laws with source terms*. SIAM J. Numer. Anal. **34** (1997), 1980–2007. MR1472206 (98k:65049)

[23] G. Jiang and C.-W. Shu, *Efficient implementation of weighted ENO schemes*. J. Comput. Phys. **126** (1996), 202–228. MR1391627 (97e:65081)

[24] A. Harten and J.M. Hyman, *Self-adjusting grid methods for one-dimensional hyperbolic conservation laws*. J. Comput. Phys. **50** (1983), 235–269. MR0707200 (85g:65111)

[25] R.J. LeVeque, *Numerical Methods for Conservation Laws*, Birkhäuser, 1990. MR1077828 (91j:65142)

[26] R.J. LeVeque, *Balancing source terms and flux gradients in high-resolution Godunov methods: the quasi-steady wave-propagation algorithm*. J. Comput. Phys. **146** (1998), 346–365. MR1650496 (99j:65182)

[27] X.D. Liu, S. Osher and T. Chan, *Weighted essentially nonoscillatory schemes*. J. Comput. Phys. **115** (1994), 200–212. MR1300340

[28] C. Parés and M. Castro, *On the well-balanced property of Roe's method for nonconservative hyperbolic systems. Applications to shallow water systems*. ESAIM: M2AN **38**(5) (2004), 821–852. MR2104431

[29] B. Perthame and C. Simeoni, *A kinetic scheme for the Saint-Venant system with a source term*. Calcolo **38**(4) (2001), 201–231. MR1890353 (2002k:76109)

[30] B. Perthame and C. Simeoni, *Convergence of the upwind interface source method for hyperbolic conservation laws*, Proc. of Hyp. 2002, Hou and Tadmor ed., Springer, 2003. MR2053160

[31] P.L. Roe, *Approximate Riemann solvers, parameter vectors and difference schemes*. J. Comput. Phys. **43** (1981), 357–371. MR0640362 (82k:65055)

[32] J. Shi, C. Hu and C.-W. Shu, *A technique of treating negative weights in WENO schemes.* J. Comput. Phys. **175** (2002), 108–127.

[33] C.-W. Shu and S. Osher, *Efficient implementation of essentially non-oscillatory shock capturing schemes.* J. Comput. Phys. **77** (1998), 439–471. MR0954915 (89g:65113)

[34] C.-W. Shu, *Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws*, ICASE Report 97-65, 1997.

[35] E.F. Toro and M.E. Vázquez-Cendón, *Model hyperbolic systems with source terms: exact and numerical solutions*, Proceedings of "Godunov methods: Theory and Applications", 2000. MR1963646 (2004b:35215)

[36] I. Toumi, *A weak formulation of Roe's approximate Riemann solver.* J. Comput. Phys. **102**(2) (1992), 360–373. MR1187694 (93h:65126)

[37] M.E. Vázquez-Cendón, *Improved treatment of source terms in upwind schemes for the shallow water equations in channels with irregular geometry.* J. Comput. Phys. **148** (1999), 497–526. MR1669644 (99i:76102)

[38] S. Vukovic and L. Sopta, *ENO and WENO schemes with the exact conservation property for one-dimensional shallow water equations.* J. Comput. Phys. **179** (2002), 593–621. MR1911377 (2003c:65079)

[39] Y. Xing and C.-W. Shu, *High order finite difference WENO schemes with the exact conservation property for the shallow water equations*, J. Comput. Phys. **208** (2005), 206–227. MR2144699

Departamento de Análisis Matemático, Facultad de Ciencias, Universidad de Málaga, 29071-Málaga, Spain

E-mail address: castro@anamat.cie.uma.es

Departamento de Análisis Matemático, Facultad de Ciencias, Universidad de Málaga, 29071-Málaga, Spain

E-mail address: gallardo@anamat.cie.uma.es

Departamento de Análisis Matemático, Facultad de Ciencias, Universidad de Málaga, 29071-Málaga, Spain

E-mail address: pares@anamat.cie.uma.es