

Article

Collocation Methods for High-Order Well-Balanced Methods for Systems of Balance Laws

Irene Gómez-Bueno ^{1,*} , Manuel Jesús Castro Díaz ¹ , Carlos Parés ¹  and Giovanni Russo ² 

¹ Departamento de Análisis, Estadística e I.O. y Matemática Aplicada, University of Málaga, Avda. Cervantes, 2, 29071 Málaga, Spain; mjcastro@uma.es (M.J.C.D.); pares@uma.es (C.P.)

² Dipartimento di Matematica ed Informatica, University of Catania, Viale Andrea Doria, 6, 95125 Catania, Italy; russo@dmi.unict.it

* Correspondence: igomezbueno@uma.es

Abstract: In some previous works, two of the authors introduced a technique to design high-order numerical methods for one-dimensional balance laws that preserve all their stationary solutions. The basis of these methods is a well-balanced reconstruction operator. Moreover, they introduced a procedure to modify any standard reconstruction operator, like MUSCL, ENO, CWENO, etc., in order to be well-balanced. This strategy involves a non-linear problem at every cell at every time step that consists in finding the stationary solution whose average is the given cell value. In a recent paper, a fully well-balanced method is presented where the non-linear problems to be solved in the reconstruction procedure are interpreted as control problems. The goal of this paper is to introduce a new technique to solve these local non-linear problems based on the application of the collocation RK methods. Special care is put to analyze the effects of computing the averages and the source terms using quadrature formulas. A general technique which allows us to deal with resonant problems is also introduced. To check the efficiency of the methods and their well-balance property, they have been applied to a number of tests, ranging from easy academic systems of balance laws consisting of Burgers equation with some non-linear source terms to the shallow water equations—without and with Manning friction—or Euler equations of gas dynamics with gravity effects.

Keywords: systems of balance laws; well-balanced methods; finite volume methods; high order methods; reconstruction operators; collocation methods; shallow water equations; Euler equations



Citation: Gómez-Bueno, I.; Castro Díaz, M.J.; Parés, C.; Russo, G. Collocation Methods for High-Order Well-Balanced Methods for Systems of Balance Laws. *Mathematics* **2021**, *9*, 1799. <https://doi.org/10.3390/math9151799>

Academic Editors: Begoña Cano and Mechthild Thalhammer

Received: 4 June 2021
Accepted: 25 July 2021
Published: 29 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Let us consider a PDE system of the form:

$$U_t(x, t) + f(U(x, t))_x = S(U(x, t))H_x(x), \quad x \in \mathbb{R}, t > 0, \quad (1)$$

where $U(x, t)$ takes values on an open convex set $\Omega \subset \mathbb{R}^N$, $f : \Omega \rightarrow \mathbb{R}^N$ is the flux function, $S : \Omega \rightarrow \mathbb{R}^N$, and H is a continuous known function from \mathbb{R} to \mathbb{R} (possibly the identity function $H(x) = x$). It is supposed that system (1) is strictly hyperbolic, that is, $D_f(U) = \frac{\partial f}{\partial U}(U)$ has N real distinct eigenvalues $r_1(U), \dots, r_N(U)$ and associated eigenvectors v_1, \dots, v_N .

Systems of the form (1) have non-trivial stationary solutions that satisfy the ODE system:

$$f(U)_x = S(U)H_x. \quad (2)$$

A numerical method is said to be well-balanced if it has discrete stationary solutions that approximate all the stationary solutions of the system or, at least, a relevant family of them. The design of numerical methods that have the previous property is of great importance in problems where small perturbations of a stationary solution are going to be

considered, like in many geophysical problems. Methods that are well-balanced have been proposed by many authors: for example, References [1–35] and their references.

The aim of this work is to introduce a general methodology to build well-balanced high-order finite-volume numerical methods for (1) of the form:

$$\frac{d\tilde{U}_i}{dt} = -\frac{1}{\Delta x} \left(F_{i+\frac{1}{2}}(t) - F_{i-\frac{1}{2}}(t) \right) + \frac{1}{\Delta x} S_i, \tag{3}$$

where:

- $I_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$ are the computational cells, whose length Δx is supposed to be constant for simplicity;
- $\tilde{U}_i(t)$ is the approximation of the average of the exact solution at the i th cell at time t , that is,

$$\tilde{U}_i(t) \cong \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} U(x, t) dx;$$

- $F_{i+\frac{1}{2}} = \mathbb{F}(U_{i+\frac{1}{2}}^{t,-}, U_{i+\frac{1}{2}}^{t,+})$, where \mathbb{F} is a consistent numerical flux, and $U_{i+\frac{1}{2}}^{t,\pm}$ are the reconstructed states at the intercells, i.e.,

$$U_{i+\frac{1}{2}}^{t,-} = P_i^t(x_{i+\frac{1}{2}}), \quad U_{i+\frac{1}{2}}^{t,+} = P_{i+1}^t(x_{i+\frac{1}{2}}),$$

where $P_i^t(x)$ is a reconstruction of order p of the solution at the given i th cell, computed from the sequence $\{\tilde{U}_i(t)\}$:

$$P_i^t(x) = P_i(x; \{\tilde{U}_j(t)\}_{j \in \mathcal{S}_i}),$$

- where \mathcal{S}_i is the set of cell indices belonging to the stencil associated to $P_i^t(x)$;
- finally,

$$S_i \approx \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} S(P_i^t(x)) H_x(x) dx. \tag{4}$$

Given a function U , the following notation will be used to represent its cell averages and the approximations to its cell averages or its point values at the intercells:

$$\bar{U}_i = \frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} U(x) dx, \quad \tilde{U}_i \approx \bar{U}_i, \quad U^{i+1/2} \approx U(x_{i+1/2}), \quad \forall i.$$

Following References [36,37], the well-balanced property of the methods can be transferred to the reconstruction operator:

Definition 1. A reconstruction operator $P_i(x)$ is said to be well-balanced for a function U^* if

$$P_i(x) = U^*(x), \quad \forall x \in [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}], \quad \forall i, \tag{5}$$

where P_i is the approximation of U^* computed from the vector of cell-averages $\{\bar{U}_i^*\}$ of the function U^* .

In Reference [37], it has been proved that the numerical method (3) is exactly well-balanced according to Definition 2 if the reconstruction operator is well-balanced for every stationary solution U^* and

$$S_i = \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} S(P_i^t(x)) H_x(x) dx. \tag{6}$$

Definition 2. The numerical method (3) is said to be exactly well-balanced if the sequence of cell-averages $\{\bar{U}_i^*\}$ (or the sequence of their approximations $\{\tilde{U}_i^*\}$ if a quadrature formula is used to compute them) of any stationary solution U^* of (1) is an equilibrium of the ODE system (3).

Furthermore, in References [36,37], the authors propose a method to build a well-balanced reconstruction operator from any standard one that will be recalled in the next section. As we shall see, the main difficulty in the application of this technique comes from the fact that, at every cell and at every time step, one has to solve a non-linear problem of the form:

Find U such that

$$\begin{cases} f(U)_x = S(U)H_x, \\ \frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} U(x) dx = \tilde{U}_i, \end{cases} \tag{7}$$

where \tilde{U}_i is an approximation of the average at the i -th cell of the sought solution of (1). Once this problem with given average has been solved at the cell, the solution of (7), that will be denoted by U_i^* , has to be extended to the whole stencil by solving two Cauchy problems of the form

$$\begin{cases} f(U)_x = S(U)H_x, \\ U(x_{i+1/2}) = U^{i+1/2}. \end{cases} \tag{8}$$

More precisely, (8) with initial condition $U(x_{i+1/2}) = U^*(x_{i+1/2})$ (resp. with final condition $U(x_{i-1/2}) = U^*(x_{i-1/2})$) has to be solved forward in space (resp. backward in space) in the cells at the right (resp. at the left) of the i th cell in its stencil.

The previous procedure has been applied to balance laws whose stationary solutions are known in implicit or explicit form, so that problems (7) and (8) can be easily solved: see, for example, References [32,38], and the references therein. However, when solving these problems by analytic procedures is not possible, they have to be solved numerically. Please note that the two main difficulties come from the condition on the average in (7) and from the fact that the ODE is not in normal form: special care has to be taken in *resonant problems* i.e., problems in which the Jacobian matrix of the flux, $D_f(U)$ becomes singular. When the Jacobian is regular, the ODE system (2) can be written in normal form

$$U_x = D_f(U)^{-1}S(U)H_x, \tag{9}$$

and the solution of the problems is easier. At a sonic point, i.e., a point in which one of the eigenvalues vanishes, the ODE system may have no solution or may have more than one.

In recent papers [39,40], a strategy based on the interpretation of (7) as a control problem has been proposed to deal with the first difficulty: the value of U_i^* at the leftmost intercell of the stencil is the control that has to be selected so that the condition on the average is satisfied. Newton’s method is used to find the control, where the gradient is computed by solving the adjoint problem. A standard ODE solver, namely RK4, is used to solve the state and the adjoint equations, and to extend the solution to the whole stencil.

The first goal of this paper is to present a general framework to design and analyze well-balanced high-order numerical methods in which problems (7) and (8) are numerically solved: the definition of the well-balanced property and a general result that allows one to check it will be stated. The second goal is then to introduce a methodology based on the resolution of (7) and (8) using RK collocation methods that leads to numerical methods that are well-balanced according to the given definition. As we shall see, the condition on the average in (7) will naturally lead to non-linear problems that will be solved by an iterative procedure. Finally, the third goal is to propose a strategy to deal with resonant problems that is general, albeit problem-dependent: this strategy will be illustrated in the particular case of the shallow water system.

The organization of the article is as follows: Section 2 is devoted to state the general framework. Definitions of well-balanced methods and reconstruction operators adapted

to the numerical solution of the local problems will be introduced and a general result will be stated showing that well-balanced reconstruction operators lead to well-balanced methods. In Section 3, we introduce a strategy based on the use of RK collocation methods to solve problems (7) and (8), and we show that it leads to numerical schemes that are well-balanced. Section 3 ends with the introduction of the technique that allows us to deal with resonant problems.

In Section 4, a number of numerical tests are presented to analyze the performance of the methods and to check their accuracy and well-balancedness. Some numerical tests for scalar and systems of balance laws are considered: Burgers’ equation with a non-linear source term, the shallow water system—both with and without Manning friction—or the Euler equations of gas dynamics including gravity effects are considered. In addition, we also present some numerical results to check that such numerical methods introduced in this work behave correctly in the presence of critical states. Moreover, we show that the numerical methods are able to preserve subsonic and supersonic moving stationary solutions for the compressible Euler equations with gravitational force. Finally, some conclusions are drawn in Section 5, and further developments are also discussed.

2. Well-Balanced Methods

The following strategy to obtain a reconstruction operator

$$P_i(x) = P_i(x; \{\bar{U}_j\}_{j \in \mathcal{S}_i})$$

that is well-balanced according to Definition 1, from a standard one

$$Q_i(x) = Q_i(x; \{\bar{U}_j\}_{j \in \mathcal{S}_i}),$$

was introduced in Reference [36] (see also Algorithm 1):

Algorithm 1 Exactly well-balanced methods. Reconstruction procedure using exact integration.

Given a family of cell values $\{\bar{U}_i\}$, at every cell I_i :

1. Find, if possible, a stationary solution $U_i^*(x)$ defined in the stencil of cell I_i , $\cup_{j \in \mathcal{S}_i} I_j$, such that:

$$\frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} U_i^*(x) dx = \bar{U}_i. \tag{10}$$

Otherwise, take $U_i^* \equiv 0$.

2. Compute the *fluctuations* $\{V_j\}_{j \in \mathcal{S}_i}$ given by

$$V_j = \bar{U}_j - \frac{1}{\Delta x} \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} U_i^*(x) dx, \quad j \in \mathcal{S}_i,$$

and compute the reconstruction operator:

$$Q_i(x) = Q_i(x; \{V_j\}_{j \in \mathcal{S}_i}).$$

3. Finally, define

$$P_i(x) = U_i^*(x) + Q_i(x). \tag{11}$$

According to Reference [36], the reconstruction operator P_i is well-balanced if Q_i is exact for the null function, conservative if Q_i is conservative

$$\frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} P_i(x) dx = \bar{U}_i, \quad \forall i,$$

and P_i is of the same order of accuracy p of Q_i provided that the stationary solutions are smooth.

Observe that, if it is impossible to find a stationary solution defined in the stencil that satisfies (10), then the standard reconstruction is used. Please note that this choice does not spoil the well-balanced character of the numerical method: in this case, the cell values in the stencil cannot be the averages of a stationary solution (otherwise, there would be at least one solution U_i^*); thus, there is no local equilibrium to preserve. On the other hand, if there is more than one stationary solution defined on the stencil that satisfies (10), a criterion is needed to select one of them. This criterion depends on the particular problem and it will be discussed later.

Notice that the use of quadrature formulas for the computation of the integral (3) appearing in the right-hand side of (4) may destroy the well-balanced property of the method. In Reference [37], the authors proposed to write the source term as follows:

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} S(P_i^t(x))H_x(x) dx = f(U_i^{t,*}(x_{i+\frac{1}{2}})) - f(U_i^{t,*}(x_{i-\frac{1}{2}})) + \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} ((S(P_i^t(x)) - S(U_i^{t,*}(x)))H_x(x) dx, \tag{12}$$

where $U_i^{t,*}$ is the stationary solution found in (10). Finally, the integral of the source term is approximated as follows:

$$S_i = f(U_i^{t,*}(x_{i+\frac{1}{2}})) - f(U_i^{t,*}(x_{i-\frac{1}{2}})) + \Delta x \sum_{m=1}^M b_m (S(P_i^t(x_i^m)) - S(U_i^{t,*}(x_i^m)))H_x(x_i^m), \tag{13}$$

where $x_i^m, b_m, m = 1, \dots, M$ are, respectively, the nodes and the weights of the quadrature formula chosen in the cell I_i , whose order of accuracy s is larger or equal than p . It can be checked that, if the reconstruction operator is well-balanced for U^* , then the numerical method (3)–(13) is exactly well-balanced for U^* .

If the quadrature formula is used, as well, to compute cell averages, i.e.,

$$\bar{U}_i^* \approx \tilde{U}_i^* = \sum_{m=1}^M b_m U^*(x_i^m), \tag{14}$$

the reconstruction algorithm (Algorithm 2) has to be modified as follows (see Reference [37]):

Algorithm 2 Exactly well-balanced methods. Reconstruction procedure using numerical integration.

Given a family of cell values $\{\tilde{U}_i\}$, at every cell I_i :

1. Find, if possible, a stationary solution $U_i^*(x)$ defined in the stencil of cell $I_i, (\cup_{j \in \mathcal{S}_i} I_j)$ such that

$$\sum_{m=1}^M b_m U_i^*(x_i^m) = \tilde{U}_i. \tag{15}$$

Otherwise, take $U_i^* \equiv 0$.

2. Compute the *fluctuations* $\{V_j\}_{j \in \mathcal{S}_i}$ given by

$$V_j = \tilde{U}_j - \sum_{m=1}^M b_m U_i^*(x_j^m), \quad j \in \mathcal{S}_i,$$

and compute the reconstruction operator:

$$Q_i(x) = Q_i(x; \{V_j\}_{j \in \mathcal{S}_i}).$$

3. Define

$$P_i(x) = U_i^*(x) + Q_i(x).$$

Again, it can be proved that this reconstruction operator is well-balanced and leads to numerical methods that are well-balanced according to Definition 2.

The main difficulty to implement the well-balanced reconstruction procedure given by Algorithm 2 comes from the first step in which the values at the quadrature points and the intercells of the stationary solution that satisfies (15) have to be computed in the stencil. When it is not possible to solve these problems by analytical procedure, a numerical one is required. In this case, the generic local discrete problem to be solved is the following:

Problem 1 (Local problem (LP)). *Given an index i and a state $\tilde{W} \in \Omega$, find approximations*

$$U_{i,j}^{*,m}, m = 1, \dots, M, j \in \mathcal{S}_i; \quad U_i^{*,i\pm 1/2},$$

of the values

$$U_i^*(x_i^m), m = 1, \dots, M, j \in \mathcal{S}_i; \quad U_i^*(x_{i\pm 1/2}),$$

where U_i^* is the stationary solution that satisfies

$$\sum_{m=1}^M b_m U_i^*(x_i^m) = \tilde{W}. \tag{16}$$

Although the approximations to be found depend in general on Δx and \tilde{W} , this dependency will not be explicitly written to avoid an excess of notation.

Once a numerical solver has been selected for the local discrete problems, the approximations of the stationary solution satisfying (16) at the quadrature points $U_{i,j}^{*,m}$ and at the intercells $U_i^{*,i\pm 1/2}$ are used to define the reconstruction operator as follows (Algorithm 3):

Algorithm 3 Well-balanced methods. Reconstruction procedure.

Given a sequence of cell values $\{\tilde{U}_i\}$, at every cell I_i :

1. Apply, if possible, the local solver at the i -th cell with $\tilde{W} = \tilde{U}_i$ to obtain

$$U_{i,j}^{*,m}, m = 1, \dots, M, j \in \mathcal{S}_i; \quad U_i^{*,i\pm 1/2}.$$

Otherwise, take

$$U_{i,j}^{*,m} = 0, m = 1, \dots, M, j \in \mathcal{S}_i; \quad U_i^{*,i\pm 1/2} = 0.$$

2. Compute the fluctuations $\{V_j\}_{j \in \mathcal{S}_i}$ given by

$$V_j = \tilde{U}_i - \sum_{m=1}^M b_m U_{i,j}^{*,m}, \quad j \in \mathcal{S}_i,$$

and compute:

$$Q_i(x) = Q_i(x; \{V_j\}_{j \in \mathcal{S}_i}).$$

3. Define

$$P_i^m = U_{i,i}^{*,m} + Q_i(x_i^m), \quad m = 1, \dots, M, \tag{17}$$

$$U_{i-1/2}^+ = U_i^{*,i-1/2} + Q_i(x_{i-1/2}), \tag{18}$$

$$U_{i+1/2}^- = U_i^{*,i+1/2} + Q_i(x_{i+1/2}). \tag{19}$$

Then, the numerical method is defined by (3) with

$$S_i = f(U_i^{*,i+1/2}) - f(U_i^{*,i-1/2}) + \Delta x \sum_{m=1}^M b_m (S(P_i^m) - S(U_{i,i}^{*,m})) H_x(x_i^m), \quad (20)$$

where the dependency on t has not been explicitly written to avoid an excess of indices.

The question is now what is the well-balanced property satisfied by the previous method. The following definition is considered:

Definition 3. The numerical method (3) is said to be well-balanced with order $q \geq p$ if for every stationary solution U^* of (1), and, for every Δx , there exists a discrete stationary solution $\{\tilde{U}_{\Delta x,i}^*\}$, i.e., an equilibrium of (3), such that

$$\bar{U}_i^* = \tilde{U}_{\Delta x,i}^* + O(\Delta x^q), \quad \forall i. \quad (21)$$

In order to illustrate the definition, let us consider some examples. We consider first the equation:

$$u_t + u_x = u, \quad (22)$$

whose stationary solutions are

$$u(x) = Ce^x, \quad C \in \mathbb{R}.$$

The cell-averages will be computed by the mid-point value so that

$$\frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} u_0(x) dx \approx u_0(x_i).$$

We consider the upwind numerical flux and different discretizations of the source term:

- Method 1:

$$\frac{du_i}{dt} = \frac{1}{\Delta x} (e^{\Delta x} u_{i-1} - u_i). \quad (23)$$

This is the first-order well-balanced numerical method based on the well-balanced piecewise constant reconstruction operator described in Algorithm 2. The discrete stationary solutions satisfy

$$u_i = e^{\Delta x} u_{i-1}, \quad i = 0, 1, \dots$$

It is clear then that the set of discrete stationary solutions is given by the point values of the continuous ones:

$$\{Ce^{x_i}\}, \quad C \in \mathbb{R},$$

and the numerical method is the exactly well-balanced according to Definition (2).

- Method 2:

$$\frac{du_i}{dt} = \frac{1}{\Delta x} (u_i - u_{i-1}) + \frac{1}{2} (u_{i-1} + u_i). \quad (24)$$

This is the first-order well-balanced numerical method based on the RK2 collocation method that will be introduced in Section 3. The discrete stationary solutions satisfy

$$u_i = u_{i-1} + \frac{\Delta x}{2} (u_{i-1} + u_i),$$

which leads to

$$u_i = \frac{1 + \Delta x/2}{1 - \Delta x/2} u_{i-1}, \quad i = 1, 2, \dots$$

The set of discrete stationary solutions is given by:

$$\left\{ C \left(\frac{1 + \Delta x/2}{1 - \Delta x/2} \right)^i \right\}, \quad C \in \mathbb{R}.$$

One has for every C:

$$u(x_i) = Ce^{i\Delta x} = C \left(\frac{1 + \Delta x/2}{1 - \Delta x/2} \right)^i + O(\Delta x^2), \quad \forall i,$$

so that the numerical method is well-balanced with order 2.

- Method 3:

$$\frac{du_i}{dt} = \frac{1}{\Delta x} (u_{i-1} - u_i) + u_{i-1}. \tag{25}$$

This is the numerical method that corresponds to the upwind treatment of the source term (see Reference [41]). The discrete stationary solutions satisfy

$$u_i = u_{i-1} + \Delta x u_{i-1},$$

which leads to

$$u_i = (1 + \Delta x) u_{i-1}, \quad i = 1, 2, \dots$$

The set of discrete stationary solutions is given by:

$$\left\{ C(1 + \Delta x)^i \right\}, \quad C \in \mathbb{R}.$$

One has for every C:

$$u(x_i) = Ce^{i\Delta x} = C(1 + \Delta x)^i + O(\Delta x), \quad \forall i,$$

so that the numerical method is well-balanced with order 1.

- Consider now the equation

$$u_t + \left(\frac{u^2}{2} \right)_x = u,$$

whose stationary solutions are

$$u = x + \text{constant},$$

and $u \equiv 0$. Let us consider the numerical method:

$$\frac{du_i}{dt} = \frac{1}{\Delta x} (F_{i-1/2} - F_{i+1/2}) + \bar{u}_{i-1/2}, \tag{26}$$

where

$$\bar{u}_{i+1/2} = \frac{u_i + u_{i+1}}{2}, \quad F_{i+1/2} = \begin{cases} u_i^2/2 & \text{if } \bar{u}_{i+1/2} > 0; \\ 0 & \text{if } \bar{u}_{i+1/2} = 0; \\ u_{i+1}^2/2 & \text{if } \bar{u}_{i+1/2} < 0. \end{cases}$$

The discrete stationary solutions satisfy

$$F_{i-1/2} - F_{i+1/2} + \Delta x \bar{u}_{i-1/2} = 0.$$

Let us see that there are no discrete stationary solutions that change their sign from negative to positive (as it happens with all the stationary solutions)

$$u^*(x) = C + x, \quad C < 0$$

if the interval contains point $x_0 = -C$.

- Let us suppose that there exists a discrete stationary solution such that $\bar{u}_{i-1/2} < 0$, $\bar{u}_{i+1/2} > 0$. Then, at the i -th cell, we would have:

$$0 = \frac{u_i^2}{2} - \frac{u_i^2}{2} + \Delta x \bar{u}_{i-1/2} \implies \bar{u}_{i-1/2} = 0,$$

which is a contradiction.

- Let us suppose now that $\bar{u}_{i-3/2} < 0$, $\bar{u}_{i-1/2} = 0$, $\bar{u}_{i+1/2} > 0$. Then, at the i -th cell, we have

$$0 = -\frac{u_i^2}{2} + \Delta x \bar{u}_{i-1/2} = -\frac{u_i^2}{2} \implies u_i = 0.$$

Since $\bar{u}_{i-1/2} = 0$, we have that $u_{i-1} = 0$, as well. Therefore, at the $(i - 1)$ -th cell, we have

$$0 = \frac{u_{i-1}^2}{2} + \Delta x \bar{u}_{i-3/2} = \Delta x \bar{u}_{i-3/2} \implies \bar{u}_{i-3/2} = 0,$$

which is again a contradiction.

Therefore, it is not possible to find discrete stationary solutions that change from negative to positive. The numerical method is not well-balanced in spite of the fact that it is consistent and stable.

Remark 1. Some authors impose $q > p$ in the definition of well-balanced, i.e., stationary solutions have to be approximated with enhanced accuracy. Please note that, although any standard numerical method of order p will provide approximations of any smooth stationary solution with order of accuracy p , these approximations are not necessarily discrete stationary solutions. This is the case of the fourth example above: although the numerical method will provide first order numerical approximations of the stationary solutions of the form $u^*(x) = x + C$ with $C < 0$, these approximations cannot be discrete stationary solutions.

Remark 2. From the practical point of view, it is important to know how to compute discrete stationary solutions in order to 'prepare well' the initial conditions. Think, for instance, of a numerical experiment for Equation (22) with initial condition

$$u_0(x) = u^*(x) + \delta(x),$$

where $u^*(x)$ is a stationary solution, say $u^*(x) = e^x$, and $\delta(x)$ a small perturbation with compact support. If the initial condition is computed by

$$u_i^0 = e^{x_i} + \delta(x_i), \quad i = 1, 2, \dots,$$

and Method 1 is used, the stationary solution will remain unperturbed in regions where the wave generated by the initial perturbation is not arrived. Nevertheless, if Method 2 is used, the numerical solution will move everywhere to fit a discrete stationary solution of the problem. If the amplitude of the initial perturbation is of order $O(\Delta x)$, the wave can be lost among the numerical errors at the beginning of the experiment. If, instead, the initial condition is computed using the discrete stationary solution that approximates u^* ,

$$u_i^0 = \left(\frac{1 + \Delta x/2}{1 - \Delta x/2} \right)^i + \delta(x_i), \quad i = 1, 2, \dots$$

the numerical solution will remain again unperturbed until the arrival of the wave.

According to Definition 3, in order to be well-balanced, a numerical method has to have discrete stationary solutions. Let us discuss what are the equilibria of (3)–(20) when

the well-balanced reconstruction operator given by Algorithm 3 is used. To do this, let us first introduce the following discrete version of Definition 1:

Definition 4. The reconstruction operator defined by Algorithm 3 is said to be well-balanced for a sequence of cell values $\{\tilde{U}_i\}$ if, for every i ,

$$\sum_{m=1}^M b_m U_{i,j}^{*,m} = \tilde{U}_j, \quad j \in \mathcal{S}_i, \tag{27}$$

$$U_{i+1/2}^- = U_i^{*,i+1/2} = U_{i+1}^{*,i+1/2} = U_{i+1/2}^+, \tag{28}$$

where $U_{i,j}^{*,m}, U_i^{*,i\pm 1/2}$ is the output given by the solver of local problems for $\tilde{W} = \tilde{U}_i$ at the i th cell.

In other words, the reconstruction procedure defined by Algorithm 3 is well-balanced for a sequence $\{\tilde{U}_i\}$ if (a) the cell values are recovered when the quadrature formula is applied to the approximations at the points given by the local solver and (b) the approximations at the intercells do not depend on the stencil.

Like in the continuous case, there is a direct relation between discrete stationary solutions and well-balanced reconstruction operators:

Theorem 1. If the reconstruction operator described in Algorithm 3 is well-balanced for a sequence of cell values $\{\tilde{U}_i^*\}$, then the sequence is a discrete stationary solution, i.e., an equilibrium of the ODE system given by (3)–(20).

Proof. Let us show that the sequence $\{\tilde{U}_i^*\}$ is an equilibrium of (3)–(20). Due to the well-balancedness of the reconstruction operator, at the i th cell, we have:

$$V_j = \tilde{U}_j^* - \sum_{j=1}^M b_m U_{i,j}^{*,m} = 0, \quad j \in \mathcal{S}_i,$$

so that $Q_i(x) = 0$. Therefore,

$$\begin{aligned} P_i^m &= U_{i,i}^{*,m}, \quad m = 1, \dots, M, \\ U_{i-1/2}^+ &= U_i^{*,i-1/2}, \\ U_{i+1/2}^- &= U_i^{*,i+1/2}. \end{aligned}$$

As a consequence, (20) reduces to

$$S_i = f(U_i^{*,i+1/2}) - f(U_i^{*,i-1/2}).$$

On the other hand:

$$\begin{aligned} F_{i+\frac{1}{2}} - F_{i-\frac{1}{2}} &= \mathbb{F}(U_{i+\frac{1}{2}}^-, U_{i+\frac{1}{2}}^+) - \mathbb{F}(U_{i-\frac{1}{2}}^-, U_{i-\frac{1}{2}}^+) \\ &= \mathbb{F}(U_i^{*,i+1/2}, U_{i+1}^{*,i+1/2}) - \mathbb{F}(U_{i-1}^{*,i-1/2}, U_i^{*,i-1/2}) \\ &= f(U_i^{*,i+1/2}) - f(U_i^{*,i-1/2}). \end{aligned}$$

In the previous expression, the consistency of the numerical flux and (28) have been used. Therefore, the right-hand side of (3) vanishes. \square

The previous result gives the condition under which a sequence of cell values defines a discrete stationary solution. In order to prove that the method is well-balanced, one has to prove that, given any stationary solution U^* and Δx , it is possible to construct a discrete one that approximates U^* with order of accuracy greater or equal than p . If we observe the

three first examples above, we see that the discrete stationary solutions satisfy a consistent and well-defined ODE solver of the equation satisfied by the continuous ones

$$u_x = u,$$

the exact solver for Method 1, the trapezoidal scheme for Method 2, and forward Euler scheme for Method 3. This is not the case for the fourth example. The following result states that, if there exists an ODE solver of (2) with order of accuracy greater or equal than p that is consistent in some sense with the local problems solver, then the numerical method is well-balanced:

Theorem 2. *Let us suppose that there exists an ODE solver of (2) with order of accuracy $q \geq p$ that, given any stationary solution U^* and any value of Δx , provides approximations*

$$U_{\Delta x,i}^{*,m}, m = 1, \dots, M; U_{\Delta x}^{*,i+1/2}, \forall i,$$

of the values of U^* at the quadrature points and the intercells

$$U^*(x_i^m), m = 1, \dots, M; U^*(x_{i+1/2}), \forall i,$$

in such a way that, at every cell I_i , the solution of the local discrete problem (16) with \tilde{W} equal to

$$\tilde{U}_{\Delta x,i}^* := \sum_{m=1}^M b_m U_{\Delta x,i}^{*,m} \tag{29}$$

is given by

$$U_{i,j}^{*,m} = U_{\Delta x,j}^{*,m}, m = 1, \dots, M, j \in S_i; \tag{30}$$

$$U_i^{*,i\pm 1/2} = U_{\Delta x}^{*,i\pm 1/2}. \tag{31}$$

Then, the numerical method (3)–(20) is well-balanced.

Proof. The sequence $\{\tilde{U}_{\Delta x,i}^*\}$ given by (29) approaches the cell values of U^* with order at least p since both the orders of the ODE solver and the quadrature formula are greater or equal than p . On the other hand, it is obviously well-balanced for the reconstruction operator since

$$\sum_{m=1}^M b_m U_{i,j}^{*,m} = \sum_{m=1}^M b_m U_{\Delta x,j}^{*,m} = \tilde{U}_{\Delta x,j}^*$$

and, furthermore

$$U_i^{*,i+1/2} = U_{i+1}^{*,i+1/2} = U_{\Delta x}^{*,i+1/2}$$

so that (27) and (28) are satisfied. As a consequence of Theorem 1, $\{\tilde{U}_{\Delta x,i}^*\}$ is a discrete stationary solution, which completes the proof. \square

In a recent paper [39], the ODE solver chosen for (2) was RK4. The local problems were interpreted as control problems in which both the state and the adjoint local problems were solved using RK4, as well, which gave the consistency of both solvers and, thus, the well-balanced property of the method.

3. Collocation Methods

In this paper, we propose to build well-balanced numerical methods for (1) on the basis of collocation RK methods. Remember that these methods admit a double interpretation: on the one hand, they are standard RK methods based on a Butcher tableau

$$\begin{array}{c|ccc}
 c_1 & a_{1,1} & \dots & a_{1,s} \\
 c_2 & a_{2,1} & \dots & a_{2,s} \\
 \vdots & \vdots & \ddots & \vdots \\
 c_s & a_{s,1} & \dots & a_{s,s} \\
 \hline
 & b_1 & \dots & b_s.
 \end{array}$$

Given an index i_0 , when these methods are applied to a Cauchy problem

$$\begin{cases} U_x = G(x, U), \\ U(x_{i_0-1/2}) = U^{i_0-1/2}, \end{cases} \tag{32}$$

in a uniform mesh of nodes $x_{i+1/2} = x_{i-1/2} + \Delta x, i = i_0, i_0 + 1, \dots$, the numerical solutions are updated as follows:

$$U^{i+1/2} = U^{i-1/2} + \Delta x \Phi_{\Delta x}(U^{i-1/2}), \quad i = i_0, i_0 + 1, \dots \tag{33}$$

where

$$\Phi_{\Delta x}(U^{i-1/2}) = \sum_{j=1}^s b_j K_i^j.$$

Here, K_i^1, \dots, K_i^s solve the non-linear system

$$K_i^j = G\left(x_i^j, U^{i-1/2} + \Delta x \sum_{l=1}^s a_{j,l} K_i^l\right), \quad j = 1, \dots, s, \tag{34}$$

where

$$x_i^j = x_{i-1/2} + c_j \Delta x, \quad j = 1, \dots, s. \tag{35}$$

On the other hand, they can be interpreted as follows:

$$U^{i+1/2} = P_i(x_{i+1/2}),$$

where P_i is the only polynomial of degree s that satisfies:

$$\begin{cases} P_i(x_{i-1/2}) = U^{i-1/2}, \\ P_i'(x_i^j) = G(x_i^j, P_i(x_i^j)), \quad j = 1, \dots, s. \end{cases} \tag{36}$$

We will consider here Gauss-Legendre methods, in which x_i^1, \dots, x_i^s and b_1, \dots, b_s are, respectively, the quadrature points and the weights of the Gauss quadrature formula in the interval $[x_{i-1/2}, x_{i+1/2}]$. This quadrature formula will be used to compute the averages at the cells. The order of accuracy of these methods is $2s$.

Because of this double interpretation, it can be easily shown that Gauss methods are symmetric or reversible in the following sense (see Reference [42]):

$$\Phi_{\Delta x} \circ \Phi_{-\Delta x} = Id, \quad \text{or equivalently} \quad \Phi_{\Delta x} = \Phi_{-\Delta x}^{-1}. \tag{37}$$

3.1. Cauchy Problem Solver

Let us describe how the collocation RK methods will be adapted to solve problems of the form (8). The following algorithm will be used to approximate the solution of a Cauchy problem with initial condition

$$U^*(x_{i_0-1/2}) = U^{i_0-1/2}$$

forward and backward in space (Algorithm 4):

Algorithm 4 Numerical solver for the Cauchy problem (8) using collocation RK methods.

- For $i = i_0, i_0 + 1, \dots$
 - Compute $K_i^m, m = 1, \dots, s$ by solving the non-linear system:

$$D_f \left(U^{i-1/2} + \Delta x \sum_{k=1}^s a_{m,k} K_i^k \right) K_i^m = S \left(U^{i-1/2} + \Delta x \sum_{k=1}^s a_{m,k} K_i^k \right) H_x(x_i^m), \quad m = 1, \dots, s. \tag{38}$$

- Compute:

$$U^{i+1/2} = U^{i-1/2} + \Delta x \sum_{m=1}^s b_m K_i^m.$$

- Compute:

$$U_i^m = U^{i-1/2} + \Delta x \sum_{k=1}^s a_{m,k} K_i^k, \quad m = 1, \dots, s.$$

- For $i = i_0, i_0 - 1, i_0 - 2, \dots$
 - Compute K_{i-1}^m by solving the non-linear system:

$$D_f \left(U^{i-1/2} - \Delta x \sum_{k=1}^s a_{m,k} K_{i-1}^k \right) K_{i-1}^m = S \left(U^{i-1/2} - \Delta x \sum_{k=1}^s a_{m,k} K_{i-1}^k \right) H_x(x_{i-1}^m), \quad m = 1, \dots, s. \tag{39}$$

- Compute:

$$U^{i-3/2} = U^{i-1/2} - \Delta x \sum_{m=1}^s b_m K_{i-1}^m.$$

- Compute:

$$U_{i-1}^m = U^{i-1/2} - \Delta x \sum_{k=1}^s a_{m,k} K_{i-1}^k, \quad m = 1, \dots, s.$$

With the notation of Theorem 2, this ODE solver provides the following approximations of a stationary solution U^* at the quadrature points and intercells:

$$U_{\Delta x}^{*,i+1/2} = U^{i+1/2}, \quad U_{\Delta x,i}^{*,m} = U_i^m, \quad m = 1, \dots, s, \quad \forall i,$$

and the discrete stationary solutions will be then given by

$$\tilde{U}_{\Delta x,i}^* = \sum_{m=1}^s b_m U_{\Delta x,i}^{*,m}.$$

The non-linear systems (38) and (39) are solved by using a fixed-point algorithm. If the Jacobian of the flux function is never singular, this algorithm is equivalent to apply the collocation RK method to the ODE system in normal form (9). When the Jacobian may become singular, special care has to be taken when solving the linear systems involved by the fixed-point algorithm: this difficulty will be discussed in Section 3.6.

3.2. Local Problem Solver

Let us suppose that the stencils are

$$\mathcal{S}_i = \{i - l, \dots, i + r\}.$$

Remember that local problems (LP) consist in, given an index i and a state $\tilde{W} \in \Omega$, finding approximations

$$U_{i,j}^{*,m}, \quad m = 1, \dots, M, \quad j = i - l, \dots, i + r; \quad U_i^{*,i\pm 1/2},$$

of the values

$$U_i^*(x_j^m), m = 1, \dots, M, j \in i - l, \dots, i + r; \quad U_i^*(x_{i \pm 1/2}),$$

where U_i^* is the stationary solution that satisfies

$$\sum_{m=1}^M b_m U_i^*(x_i^m) = \tilde{W}. \tag{40}$$

To solve the problem, we look first for an approximation of U_i^* in the cell I_i using the collocation method, i.e., we look for a vector valued polynomial P_i of degree s such that:

$$\begin{cases} \sum_{m=1}^s b_m P_i(x_i^m) = \tilde{W}, \\ D_f(P_i(x_i^m))P_i'(x_i^m) = S(P_i(x_i^m))H_x(x_i^m), \quad m = 1, \dots, s. \end{cases} \tag{41}$$

Observe that this system has $(s + 1)$ vector unknowns (the coefficients of the polynomials) and $(s + 1)$ vector equations.

A fixed-point algorithm will be used to solve this system. Once it has been solved, the approximations of U_i^* at the intercells are computed by evaluating the polynomial P_i :

$$U^{i-1/2} = P_i(x_{i-1/2}), \quad U^{i+1/2} = P_i(x_{i+1/2}).$$

From these two values, the Cauchy solver described in the previous section will be used to compute the approximations of the stationary solution in the other cells of the stencil.

Using the standard RK notation, the local problem solver is as follows (Algorithm 5):

Algorithm 5 Numerical solver for the local problems (LP) using collocation RK methods.

- Compute $U^{i-1/2}, K_i^m, m = 1, \dots, s$ by solving the system:

$$\begin{cases} D_f\left(U^{i-1/2} + \Delta x \sum_{k=1}^s a_{m,k} K_i^k\right) K_i^m = S\left(U^{i-1/2} + \Delta x \sum_{k=1}^s a_{m,k} K_i^k\right) H_x(x_i^m), \quad m = 1, \dots, s, \\ \sum_{m=1}^s b_m \left(U^{i-1/2} + \Delta x \sum_{k=1}^s a_{m,k} K_i^k\right) = \tilde{W}. \end{cases} \tag{42}$$

- Compute:

$$\begin{aligned} U^{i+1/2} &= U^{i-1/2} + \Delta x \sum_{m=1}^s b_m K_i^m; \\ U_i^m &= U^{i-1/2} + \Delta x \sum_{k=1}^s a_{m,k} K_i^k, \quad m = 1, \dots, s. \end{aligned}$$

- Apply Algorithm 4 forward in space from $U^{i+1/2}$ to obtain for $j = i + 1, \dots, i + r$

$$U^{j+1/2}; \quad U_j^m, \quad m = 1, \dots, s.$$

- Apply Algorithm 4 backward in space from $U^{i-1/2}$ to obtain for $j = i - l, \dots, i - 1$

$$U^{j-1/2}; \quad U_j^m, \quad m = 1, \dots, s.$$

The output of the solver is then:

$$U_{i,j}^{*,m} = U_j^m, \quad m = 1, \dots, s, \quad j = i - l, \dots, i + r; \quad U_i^{*,i-1/2} = U^{i-1/2}, \quad U_i^{*,i+1/2} = U^{i+1/2}.$$

A fixed-point algorithm is used again to solve the non-linear system (42). A sensible initial guess is given by:

$$\begin{aligned} U^{i-1/2,0} &= \tilde{W}, \\ K_i^{m,0} &= K, \quad m = 1, \dots, s, \end{aligned}$$

where K is the solution of the linear system:

$$D_f(\tilde{W})K = S(\tilde{W})H_x(x_i),$$

where x_i represents the mid-point of the cell I_i . Let us mention again that special care has to be taken if the Jacobian of the flux function may become singular: this difficulty will be discussed in Section 3.6.

3.3. Well-Balanced Property

To prove that the high-order numerical methods are well-balanced, we have to check that the ODE solver and the local problem solver described in the two last subsections are consistent in the sense of the statement of Theorem 2.

Let

$$U_{\Delta x,i}^{*,m} \quad m = 1, \dots, s; \quad U_{\Delta x}^{*,i+1/2}$$

be the approximations of a stationary solution U^* at the quadrature points and intercells

$$U^*(x_i^m), \quad m = 1, \dots, s; \quad U^*(x_{i+1/2}),$$

provided by the ODE solver described in Section 3.1 in the uniform mesh of nodes $x_{i+1/2} = x_{i-1/2} + \Delta x, i = i_0, i_0 + 1, \dots$. Let P_i^* be the only polynomial of degree s such that:

$$\begin{aligned} P_i^*(x_{i-1/2}) &= U_{\Delta x}^{*,i-1/2}, \\ D_f(P_i^*(x_i^m))\partial_x P_i^*(x_i^m) &= S(P_i^*(x_i^m))H_x(x_i^m), \quad i = 1, \dots, s. \end{aligned}$$

By the double interpretation of the method, we have:

$$\begin{aligned} P_i^*(x_i^m) &= U_{\Delta x,i}^{*,m} \quad m = 1, \dots, s; \\ P_i^*(x_{i+1/2}) &= U_{\Delta x}^{*,i+1/2}. \end{aligned}$$

To solve now the local problem at the i -th cell with

$$\tilde{W} = \tilde{U}_{\Delta x,i}^* = \sum_{m=1}^s b_m U_{\Delta x,i}^{*,m}$$

we have to look for a polynomial P_i satisfying (41). Obviously, P_i^* solves this non-linear system; thus, $P_i = P_i^*$, which implies, in particular,

$$U_{i,i}^{*,m} = U_{\Delta x,i}^{*,m} \quad m = 1, \dots, s; \tag{43}$$

$$U_i^{*,i\pm 1/2} = U_{\Delta x}^{*,i\pm 1/2}. \tag{44}$$

The equalities

$$U_{i,j}^{*,m} = U_{\Delta x,j}^{*,m} \quad m = 1, \dots, s, \quad j = i + 1, \dots, i + r \tag{45}$$

are then trivial, since these values are computed by applying Algorithm 4 forward in space from the same value $U_{\Delta x, i \pm 1/2}^*$, and the equalities

$$U_{i,j}^{*,m} = U_{\Delta x, j}^{*,m} \quad m = 1, \dots, s, \quad j = i - l, \dots, i - 1, \tag{46}$$

follow from the symmetry of the collocation method. Therefore, the hypothesis of Theorem 2 is satisfied, and the numerical method is then well-balanced *provided that every stationary solution U^* can be approximated by the RK collocation method.*

3.4. First and Second Order Methods

For second order methods, the MUSCL reconstruction that uses 3-cell centered stencils ($l = r = 1$) and the second-order 1-stage collocation RK method whose Butcher tableau is

$$\begin{array}{c|c} 1/2 & 1/2 \\ \hline 0 & 1 \end{array} \tag{47}$$

will be used. The solution of local problems is particularly simple, since (42) reduces to find $U^{i-1/2}, K$ such that

$$\begin{cases} D_f \left(U^{i-1/2} + \frac{\Delta x}{2} K \right) = S \left(U^{i-1/2} + \frac{\Delta x}{2} K \right) H_x(x_i^1), \\ U^{i-1/2} + \frac{\Delta x}{2} K = \tilde{W}, \end{cases}$$

where

$$x_i^1 = x_{i-1/2} + \frac{\Delta x}{2} = x_i.$$

Therefore, it is enough to solve

$$D_f(\tilde{W})K = S(\tilde{W})H_x(x_i)$$

and then define

$$U^{i-1/2} = \tilde{W} - \frac{\Delta x}{2} K, \quad U^{i+1/2} = \tilde{W} + \frac{\Delta x}{2} K, \quad U_0^1 = \tilde{W}.$$

Then, the standard method is used forward and backward in space to compute U_{-1}^1 and U_1^1 .

Finally, for first order numerical method, the trivial pointwise constant reconstruction is considered so that the stencils for the reconstructions consist only of one cell ($l = r = 0$). The second order collocation method (47) is also used.

Remark 3. Notice that, since the 1-stage Gauss-Legendre method is second order, the schemes considered here are:

- First order numerical methods which are well-balanced with order two.
- Second order numerical methods which are well-balanced with order two.

3.5. Third-Order Methods

The third-order CWENO reconstruction operator will be used here to design third-order methods. This operator uses 3-cell centered stencils, so that $l = r = 1$. The fourth-order 2-stage Gauss-Legendre method corresponding to the choices,

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} - \frac{\sqrt{3}}{6} \\ \frac{1}{4} + \frac{\sqrt{3}}{6} & \frac{1}{4} \end{bmatrix}, \tag{48}$$

$$c_1 = \frac{1}{2} - \frac{\sqrt{3}}{6}, \quad c_2 = \frac{1}{2} + \frac{\sqrt{3}}{6}, \quad b_1 = b_2 = \frac{1}{2}, \tag{49}$$

will be used to solve both the local and global problems.

Remark 4. Notice that, since the 2-stage Gauss-Legendre method is fourth order, we consider third order numerical methods which are well-balanced with order four.

3.6. Resonant Problems

When a fixed-point algorithm is used to solve (38), (39), or (42), one or more linear systems of the form

$$D_f(W)K = S(W)H_x(\bar{x}) \tag{50}$$

have to be solved at every step, where W is a known state, and \bar{x} a quadrature point. If W is a sonic state, i.e., if one of the eigenvalues of $D_f(W)$ vanishes, the problem is said to be resonant and the system may not have solution or it may have infinitely many ones:

1. If $S(W)H_x(x)$ does not belong to the image of the linear application defined by the matrix $D_f(W)$, the system has no solution.
2. Otherwise, the system has infinitely many solutions

$$K^* + \alpha R, \quad \alpha \in \mathbb{R},$$

where K^* is a particular solution, and R is an eigenvector associated to the null eigenvalue.

Since, when applying RK collocation methods, we are looking for solutions of the differential Equation (2), a solution K of (50) is said to be *admissible* if there exists a C^1 stationary solution U^* such that

$$U^*(\bar{x}) = W, \quad U_x^*(\bar{x}) = K.$$

In case 1, there are no admissible solutions. If this situation arises when solving a local problem, it will be assumed that the problem has no solution and the standard reconstruction will be used in the corresponding cell.

In case 2, if K is an admissible solution and W is an isolated sonic point of the corresponding smooth stationary solution U^* , the following equality is satisfied:

$$U_x^*(x) = D_f(U^*)^{-1}S(U^*)H_x(x), \quad \forall x \in (\bar{x} - \varepsilon, \bar{x} + \varepsilon) - \{\bar{x}\},$$

for some $\varepsilon > 0$. Therefore, the limit

$$\lim_{x \rightarrow \bar{x}} D_f(U^*)^{-1}S(U^*)H_x(x) \tag{51}$$

has to exist. If the available information about the PDE system and its stationary solution allows one to compute this limit by any analytic procedure, then the possible values of (51) constitute the set of admissible solutions, and the solution of (50) will then be selected in this set. This technique will be illustrated in the particular case of the shallow water equations in Section 4.3.

Remark 5. Observe that, if W is sonic and $S(W)$ does not belong to the image of $D_f(W)$, case 2 can only happen if $H_x(\bar{x}) = 0$. In this case, if there exists a C^1 stationary solution U^* such that $U^*(\bar{x}) = W$, then $U^*(\bar{x})$ would be an admissible solution of (50) which implies that necessarily $H_x(x^*) = 0$, i.e., a smooth stationary solution of (1) can only reach a sonic state W such that $S(W)$ does not belong to the image of $D_f(W)$ at a critical point of H . In this case, (51) will imply a 0/0 indeterminate form that can be solved by L'Hôpital rule. This is the case for many systems, including the shallow water equations, for which a smooth stationary solution can only reach a sonic (critical) state in the minimum points of the depth function H .

4. Numerical Experiments

Let us describe first the choices that have been made to implement the well-balanced scheme previously described:

- The second-order 1-stage collocation RK method is chosen for the first and second order well-balanced numerical schemes and the 2-stage Gauss-Legendre collocation method for third order ones.
- Concerning the quadrature rules, the midpoint rule for first and second order schemes are used, and the 2-point Gauss quadrature rule for third order schemes.

In all the numerical tests, Rusanov numerical flux is chosen and MUSCL (see Reference [43]) and CWENO reconstructions operators (see References [44,45]) are used, respectively, for the second and third order schemes. Finally, TVD Runge-Kutta schemes of first, second, and third [46] are used to integrate the ODE system (3)–(20).

In those problems where the initial condition is a steady state U^* , the approximation of its cell averages are computed using a quadrature formula if the exact solution is known, or by approximating the solution of the Cauchy problem (2) with initial condition

$$U(a) = U^*(a),$$

using the RK collocation method: see Remark 2.

Note that the procedure that we have been describing in this work is quite general and works for a wide class of f , S , and H .

In this section, we will denote the different numerical methods that we compare as follows:

- NWBM i , $i = 1, 2, 3$: a standard non-well balanced numerical scheme of order i .
- WBM i , $i = 1, 2, 3$: well-balanced scheme of order i with reconstruction operator given by Algorithm 2.
- CWBM i , $i = 1, 2, 3$: well-balanced scheme of order i with reconstruction operator Algorithm 3 in which the Cauchy problems are solved following Reference [39].
- CLWBM i , $i = 1, 2, 3$: well-balanced scheme of order i with reconstruction operator defined by Algorithm 3 in which Cauchy and local problems are integrated by using the Gauss-Legendre collocation method described in the previous section.

The numerical tests have been run on a laptop with Intel(R) Xeon(R) CPU E3-1220 v3 @ 3.10 GHz with 8 Mb cache using one single core.

4.1. Problem 1: Burgers Equation with a Non-linear Source Term I

We consider the Burgers equation with a non-linear source term:

$$\begin{cases} u_t + \left(\frac{u^2}{2}\right)_x = u^2, & x \in \mathbb{R}, t > 0, \\ u(x, 0) = u_0(x). \end{cases} \quad (52)$$

This problem corresponds to the choice

$$U = u, \quad f(U) = \frac{u^2}{2}, \quad S(U) = u^2, \quad H(x) = x.$$

The stationary solutions satisfy the ODE

$$u' = u, \quad (53)$$

whose general solution is

$$u(x) = Ce^x, \quad C \in \mathbb{R}.$$

As the expression of the steady states is known, the stationary solution u_i^* satisfying (15) could be easily computed for any quadrature formula

$$u_i^*(x) = \frac{u_i}{\sum_{m=0}^M \alpha_m^i e^{x_m^i}} e^x. \tag{54}$$

Although in this test WBMi can be easily computed, we also consider CWBMi and CLWBMi to compare the efficiency of the methods.

Test 1.1

Let us consider the space interval $[-1, 1]$ and the time interval $[0, 5]$. $u_0 = e^x$ is set as initial condition and the CFL parameter is set to 0.9. The boundary condition is imposed at the left boundary and open boundary conditions at the right boundary.

The convergence criterion chosen in Algorithm 5 is the following: we stop it if the maximum absolute value between the approximations of the values of the stationary solution at the extremes of the cell and the quadrature points of the cells belonging to the stencil for two consecutive iterations is less than a threshold.

In what follows, the tolerance considered to stop Algorithm 5 is $\varepsilon = 10^{-15}$.

Note that, at the discrete level, the cell-averages of $u_0(x) = e^x$ have to be computed. We consider 4 different ways of computing them:

- (a) Exactly computed:

$$u_i^0 = \frac{e^{\Delta x} - 1}{\Delta x} e^{x_{i-1/2}}, \quad \forall i.$$

- (b) Using a quadrature formula:

$$u_i^0 = \sum_{m=0}^M \alpha_m^i e^{x_m^i}, \quad \forall i.$$

- (c) The cell averages are approximated by

$$u_i^0 = \sum_{m=0}^M \alpha_m^i u_{h,m}^{*,i}, \quad \forall i, \tag{55}$$

where $u_{h,m}^{*,i}$ are the approximation the Cauchy problem

$$\begin{cases} u' = u, \\ u(-1) = e^{-1}, \end{cases} \tag{56}$$

computed with the standard RK4 method on the mesh selected to solve the control problems described in Reference [39].

- (d) The cell averages are approximated by (55), where now $u_{h,m}^{*,i}$ are the approximation of the exponential at the quadrature points obtained by solving (56) with the RK collocation method.

The numerical methods are integrated until the final time $t = 5$, which is more than twice the travel time of a small perturbation of the steady state along the domain. L^1 errors between the initial and final cell-averages will be computed. According to Theorem 1, errors should be of the order of the machine precision (MP) if option (d) is considered to compute the initial condition for CLWBMi, $i = 1, 2, 3$. Figure 1 (right) shows the errors corresponding to CLWBMi, $i = 1, 2, 3$ for a 200-cell mesh. As expected, they are of the order of MP.

Notice that, in the other cases, if (a), (b), or (c) are used for CLWBMi, $i = 1, 2, 3$, the errors decrease with Δx (see Table 1). Thus, if option (b) is chosen and a third order reconstruction operator is used, then the convergence rate is expected to be 4 as errors are

generated by the use of the 2-stage collocation RK method for solving the local problems. If option (c) is used, the same convergence rate is expected, and the errors are due to the use of RK4 for computing the initial condition and the use of the 2-stage collocation RK method. Finally, in the case of option (a), the errors are due to the use of both numerical integration and the 2-stage collocation method in the solution of local problems: again, the errors are expected to be of order 4.

Similar results are obtained for first and second-order methods: the order of convergence is two if option (a), (b), and (c) are used since the mid-point method is second-order accurate, and we consider the 1-stage collocation RK method.

Table 1. Test 1.1. L^1 errors at $t = 5$ and convergence rates for CLWBM3 when option (a), (b), (c), and (d) are chosen to compute the initial cell averages.

Cells	Option (a)		Option (b)		Option (c)		Option (d)	
	Error	Order	Error	Order	Error	Order	Error	Order
5	1.96×10^{-4}	-	1.82×10^{-4}	-	1.46×10^{-4}	-	4.88×10^{-16}	-
10	1.21×10^{-5}	4.015	1.12×10^{-5}	4.016	8.81×10^{-6}	4.054	7.44×10^{-16}	-
20	7.56×10^{-7}	4.003	7.01×10^{-7}	4.003	5.42×10^{-7}	4.023	3.00×10^{-16}	-
40	4.72×10^{-8}	4.001	4.38×10^{-8}	4.001	3.36×10^{-8}	4.011	7.94×10^{-15}	-
80	2.95×10^{-9}	4.000	2.74×10^{-9}	4.000	2.09×10^{-9}	4.006	9.00×10^{-15}	-
160	1.84×10^{-10}	4.000	1.71×10^{-10}	4.000	1.30×10^{-10}	4.003	1.23×10^{-14}	-

Now, we compare the results obtained with NWBM $_i$, WBM $_i$, CWBM $_i$, CLWBM $_i$, $i = 1, 2, 3$. The cell-averages are computed using option (b) for NWBM $_i$ and WBM $_i$, option (c) for CWBM $_i$, and option (d) for CLWBM $_i$. L^1 errors with respect to the initial condition are computed.

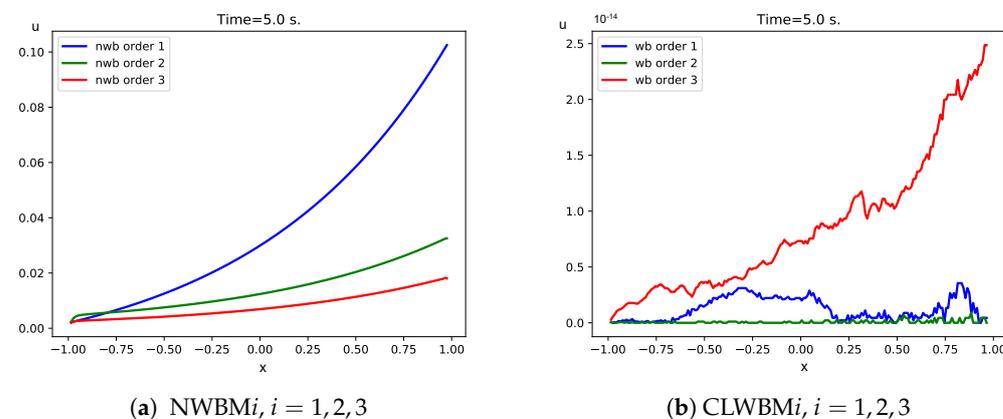


Figure 1. Test 1.1. Differences between the stationary and the numerical solutions at time $t = 5$ s for a 200–cell mesh.

Figure 1 shows the differences between the stationary and the numerical solutions obtained with NWBM $_i$ and CLWBM $_i$, $i = 1, 2, 3$. Similar results are obtained for WBM $_i$ and CWBM $_i$, $i = 1, 2, 3$. Tables 2 and 3 show the errors corresponding to NWBM $_i$, WBM $_i$, CWBM $_i$, CLWBM $_i$, $i = 1, 2, 3$.

Table 2. Test 1.1. Differences in L^1 -norm with respect to the stationary solution and convergence rates for NWBM $_i$, $i = 1, 2, 3$.

Cells	NWBM1: Error	Order	NWBM2: Error	Order	NWBM3: Error	Order
100	7.53×10^{-2}	-	2.44×10^{-3}	-	7.66×10^{-6}	-
200	3.78×10^{-2}	0.995	8.09×10^{-4}	1.591	9.62×10^{-7}	2.993
400	1.89×10^{-2}	1.002	2.16×10^{-4}	1.905	1.21×10^{-7}	2.995
800	9.43×10^{-3}	1.000	5.54×10^{-5}	1.963	1.51×10^{-8}	2.998

Table 3. Test 1.1. Differences in L^1 -norm with respect to the stationary solution for WBMi, CWBMi, CLWBMi, $i = 1, 2, 3$.

Cells	Order 1: Error			Order 2: Error			Order 3: Error		
	WBM	CWBM	CLWBM	WBM	CWBM	CLWBM	WBM	CWBM	CLWBM
100	4.21×10^{-15}	3.55×10^{-15}	2.50×10^{-15}	8.87×10^{-16}	3.63×10^{-16}	2.03×10^{-16}	3.20×10^{-16}	1.43×10^{-14}	8.17×10^{-15}
200	2.90×10^{-15}	5.54×10^{-13}	2.51×10^{-15}	4.42×10^{-16}	1.23×10^{-15}	1.66×10^{-16}	2.54×10^{-16}	2.43×10^{-14}	1.76×10^{-14}
400	1.84×10^{-14}	2.05×10^{-14}	1.12×10^{-15}	1.82×10^{-15}	3.64×10^{-16}	2.89×10^{-16}	7.40×10^{-14}	4.47×10^{-14}	4.45×10^{-14}
800	4.45×10^{-16}	2.67×10^{-15}	2.77×10^{-15}	1.83×10^{-16}	2.03×10^{-16}	2.05×10^{-16}	2.61×10^{-15}	9.48×10^{-14}	7.88×10^{-14}

Notice that the errors for NWBM decrease with Δx at the expected rate. WBM preserve the exact solution up to MP, as well as CWBM and CLWBM, provided that the tolerance used in the Newton’s method for the former and Algorithm 5 for the latter is small enough. The computational costs are shown in Table 4. Nonetheless, since the goal of carrying out these experiments is to prove that the introduced well-balanced technique works, the implementation is not optimized. It can be seen that the well-balanced reconstruction given by Algorithm 2 increases the computational cost by a factor ranging from 1.5 to 7.5 (see Table 4). On the other hand, the use of Algorithm 3 increases the computational cost by a factor ranging from 1 to 1.8 if collocation RK method are used for solving the local problems and a factor ranging from 1.1 to 3.8 if control techniques with $N_p = 1$ (see Reference [39]) are used. Therefore, the CLWBM schemes introduced in this work are more efficient than the CWBM schemes presented in Reference [39].

Table 4. Test 1.1. Computational times (milliseconds).

Cells	Order (i)	NWBMi	WBMi	CWBMi	CLWBMi
100	1	20	30	70	30
	2	30	60	140	110
	3	40	190	240	230
200	1	20	60	230	100
	2	40	190	330	280
	3	110	480	530	520
400	1	50	180	520	220
	2	100	530	1150	720
	3	350	1680	1980	1870
800	1	140	570	2020	870
	2	270	2040	3580	2820
	3	1080	5540	6600	5960

4.2. Problem 2: Burgers Equation with a Non-linear Source Term II

We now consider Burgers equation with another non-linear source term:

$$\begin{cases} u_t + \left(\frac{u^2}{2}\right)_x = \sin(u), & x \in \mathbb{R}, t > 0, \\ u(x, 0) = u_0(x). \end{cases} \tag{57}$$

The stationary solutions satisfy the ODE

$$u' = \frac{\sin(u)}{u}. \tag{58}$$

Note that no simple expression of the stationary solutions is known so (10) has to be computed numerically.

4.2.1. Test 2.1

We consider again the same computational domain, and the system is integrated up to $T = 5$. The CFL parameter is imposed to be 0.9. As initial condition, we consider the stationary solution that corresponds to the solution of the Cauchy problem (58) with

$$u(-1) = 2,$$

that is approximated by the Gauss-Legendre collocation method. The left boundary condition is set to $u(-1, t) = 2$, and open boundary conditions are considered upstream.

Figure 2 shows the differences between the stationary and the numerical solutions computed with NWB M_i , $i = 1, 2, 3$ (left) and CLW BM_i , $i = 1, 2, 3$ (right), (the figures corresponding to CW BM_i , $i = 1, 2, 3$ are similar).

The errors corresponding to NW BM_i , CW BM_i , and CLW BM_i , $i = 1, 2, 3$, respectively, are shown in Tables 5 and 6. Computational costs are shown in Table 7. Note that the collocation strategy increases the computational cost from 7.3 to 14, whereas the method proposed in Reference [39] increases the computational cost by a factor ranging from 27 to 40. After the study of these two experiments, one can conclude that the results obtained with the collocation strategy introduced in this paper are better than the ones got by CW BM_i , $i = 1, 2, 3$ as described in Reference [39], considering that the errors are similar for both procedures, but the first strategy is much more efficient. Consequently, we will only consider the collocation strategy in the following experiments.

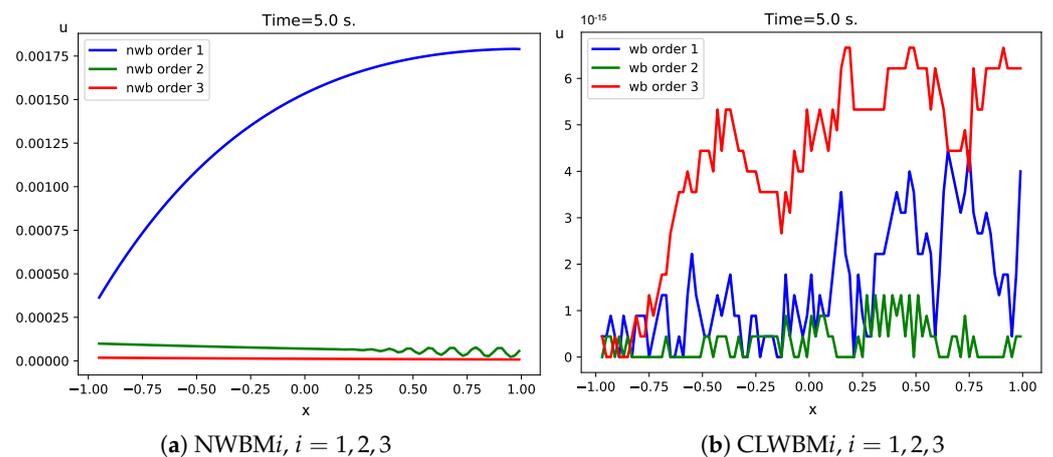


Figure 2. Test 2.1. Differences between the stationary and the numerical solutions at $t = 5$ s using a mesh with 100 cells.

Table 5. Test 2.1. Differences in L^1 -norm with respect to the stationary solution and convergence rates for NW BM_i , $i = 1, 2, 3$.

Cells	Error ($i = 1$)	Order	Error ($i = 2$)	Order	Error ($i = 3$)	Order
100	2.72×10^{-3}	-	1.43×10^{-4}	-	2.53×10^{-5}	-
200	1.34×10^{-3}	1.021	2.43×10^{-6}	5.879	1.74×10^{-8}	10.503
400	6.58×10^{-4}	1.026	8.19×10^{-7}	1.569	1.14×10^{-10}	7.250
800	3.24×10^{-4}	1.022	2.34×10^{-7}	1.806	1.41×10^{-11}	3.016

Table 6. Test 2.1. Differences in L^1 -norm with respect to the stationary solution for CW BM_i , CLW BM_i , $i = 1, 2, 3$.

Cells	Order 1: Error		Order 2: Error		Order 3: Error	
	CWBM	CLWBM	CWBM	CLWBM	CWBM	CLWBM
100	9.71×10^{-14}	3.00×10^{-15}	1.76×10^{-13}	6.39×10^{-16}	1.99×10^{-13}	8.50×10^{-15}
200	7.56×10^{-15}	5.37×10^{-15}	3.46×10^{-15}	5.15×10^{-16}	2.97×10^{-14}	2.51×10^{-14}
400	4.00×10^{-15}	5.68×10^{-15}	7.53×10^{-16}	5.73×10^{-16}	3.31×10^{-14}	4.85×10^{-14}
800	5.97×10^{-15}	4.63×10^{-15}	8.54×10^{-16}	5.31×10^{-16}	6.63×10^{-14}	9.61×10^{-14}

Table 7. Test 2.1. Computational cost (milliseconds). $t = 5$ s.

Cells	Order(i)	NWBM i	CWBM i	CLWBM i
100	1	10	340	130
	2	20	690	280
	3	40	1390	490
200	1	30	1280	230
	2	60	2350	440
	3	180	5190	1230

4.2.2. Test 2.2

Now, we consider the evolution of a small perturbation of the previous stationary solution. Thus, the initial solution is set to

$$u_0(x) = u^*(x) + 0.3e^{-200(x+0.5)^2},$$

where $u^*(x)$ is the previous stationary solution.

Figure 3 shows the difference between the stationary and the numerical solutions computed with NWBM i , $i = 1, 2, 3$ and CLWBM i , $i = 1, 2, 3$ at times $t = 0.3, 5$ (the results are similar for CWBM i , $i = 1, 2, 3$). In black, we show a reference solution computed on a fine mesh with 12,800 cells using a first order well-balanced method.

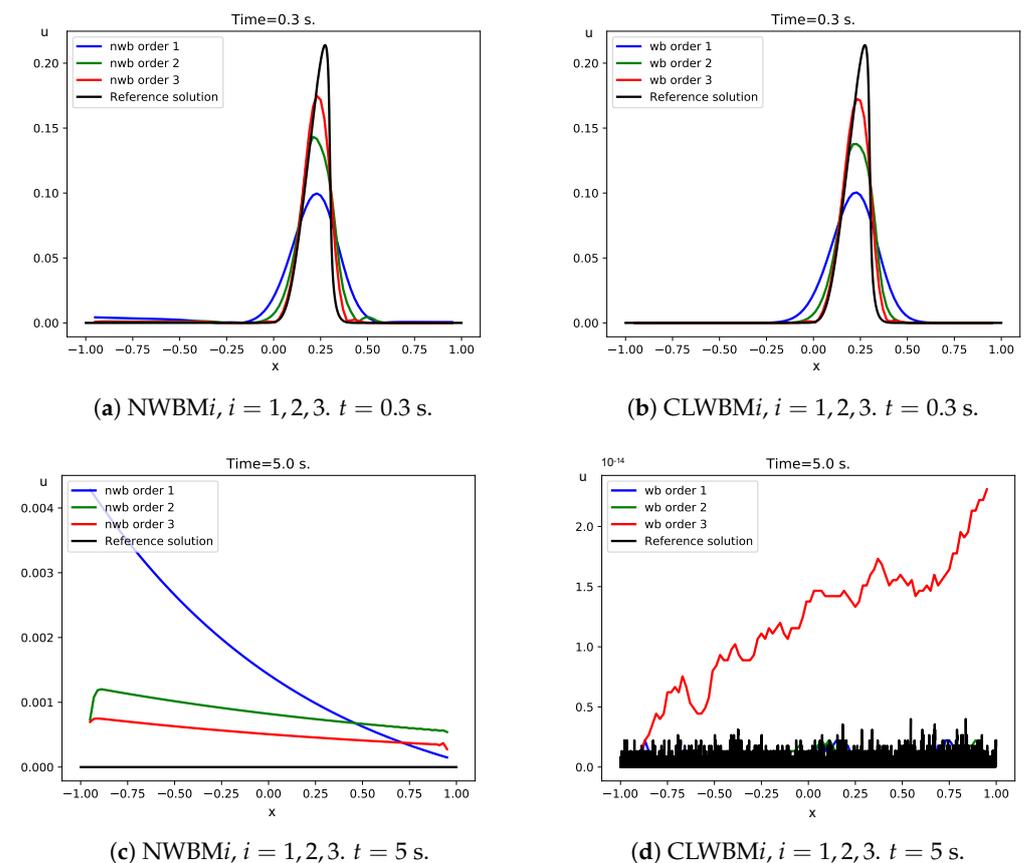


Figure 3. Test 2.2. Reference and numerical solutions: differences with the stationary solution at times $t = 0.3, 5$ s for a 100-cell mesh.

Observe that, as expected, the well-balanced methods are able to recover the stationary solution after the perturbation left the domain, while non-well balanced methods introduce an error that remains in the domain. This could be also observed in Table 8, where the errors with respect to the stationary solution at time $t = 5$ s are shown for the 100-cell mesh.

Table 8. Test 2.2. Differences in L^1 -norm with respect to the stationary solution for NWB*Mi*, CWB*Mi*, and CLW*BMi* ($i = 1, 2, 3$) at time $t = 5$ s.

Method	Error ($i = 1$)	Error ($i = 2$)	Error ($i = 3$)
NWB <i>Mi</i>	2.43×10^{-3}	1.71×10^{-3}	1.05×10^{-3}
CWB <i>Mi</i>	1.09×10^{-13}	4.47×10^{-15}	3.34×10^{-14}
CLW <i>BMi</i>	2.52×10^{-15}	1.19×10^{-15}	1.24×10^{-14}

4.3. Problem 3: Shallow Water Equations

Let us consider now the 1D shallow water system which corresponds to the choices $N = 2$, with

$$U = \begin{pmatrix} h \\ q \end{pmatrix}, \quad f(U) = \begin{pmatrix} q \\ \frac{q^2}{h} + \frac{g}{2}h^2 \end{pmatrix}, \quad S(U) = \begin{pmatrix} 0 \\ gh \end{pmatrix},$$

where x corresponds to to the axis of the channel, and t is the time. The unknown $h(x, t)$ and $q(x, t)$ are the thickness and discharge, respectively; the function $H(x)$ is the depth function, and g is the gravity; $u = q/h$ is the depth-averaged velocity, and $c = \sqrt{gh}$.

The eigenvalues of the system are $\lambda^\pm = u \pm c$, and the Froude number defined by

$$Fr(U) = \frac{|u|}{c}, \tag{59}$$

characterizes the flow regime: subcritical ($Fr < 1$), critical ($Fr = 1$), or supercritical ($Fr > 1$).

The stationary solutions U^* satisfy the ODE system :

$$\begin{cases} q_x = 0, \\ \left(\frac{q^2}{h} + \frac{1}{2}gh^2\right)_x = ghH_x, \end{cases} \tag{60}$$

which can be written as follows:

$$\begin{cases} (-u^2 + gh)h_x = ghH_x, \\ q_x = 0. \end{cases} \tag{61}$$

The stationary solutions are given in implicit form by:

$$q = C_1, \quad \frac{q^2}{2h^2} + gh - gH = C_2, \quad C_1, C_2 \in \mathbb{R}. \tag{62}$$

In Reference [38], a family of high-order well-balanced numerical methods was presented, based in the Algorithm 2, in which local problems were solved on the basis of this implicit form. Here, Algorithm 3 with collocation RK methods to solve the local problems is used.

The numerical treatment of resonant situations discussed in Section 3.6 will be illustrated here in the particular case of the shallow water system. It can be easily checked that system (50), which, in our case, becomes

$$\begin{bmatrix} 0 & 1 \\ -u^2 + gh & 2u \end{bmatrix} \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} = \begin{bmatrix} 0 \\ ghH_x(x) \end{bmatrix}, \tag{63}$$

when U^* is a critical state, i.e., when $u^2 = gh$, reduces to:

$$\begin{bmatrix} 0 & 1 \\ 0 & 2u \end{bmatrix} \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} = \begin{bmatrix} 0 \\ ghH_x(x) \end{bmatrix}. \tag{64}$$

Therefore, the system has solutions only if $H_x(x) = 0$: in this case, the solutions are

$$K = \alpha[1, 0]^T, \quad \alpha \in \mathbb{R}.$$

In fact, it can be shown (see Reference [38], for instance) that a smooth stationary solution can only reach a critical state at a minimum point x^c of the depth function H . Differentiating the second relation of (62) with respect to x , and using the relation $H_x(x_c) = 0$, one obtains, at $x = x_c$,

$$\frac{qq_x}{h^2} + h_x \left(g - \frac{q^2}{h^3} \right) = 0,$$

which, assuming $h_x(x_c) \neq 0$, implies

$$h^c(q) = \frac{q^{2/3}}{g^{1/3}}. \tag{65}$$

With this information in mind, let us compute the limit

$$\lim_{x \rightarrow x^c} \frac{ghH_x}{-u^2 + gh'}$$

to determine the value of h_x at x^c . This is a 0/0 indeterminate limit, and L'Hôpital's rule can be applied. To do this, first, the limit is rewritten as follows:

$$\lim_{x \rightarrow x^c} \frac{gH_x}{g - \frac{q^2}{h^3}}. \tag{66}$$

Some easy computation leads to

$$h_x(x^c) = \frac{h^c(q)^4}{3h_x(x^c)q^2} gH_{xx}(x^c); \tag{67}$$

therefore:

$$h_x(x^c) = \pm \frac{h^c(q)^2}{\sqrt{3}q} g^{1/2} \sqrt{H_{xx}(x^c)} = \pm \sqrt{\frac{h^c(q)H_{xx}(x^c)}{3}} = \pm \sqrt{\frac{q^{2/3}H_{xx}(x^c)}{3g^{1/3}}}. \tag{68}$$

The above expression shows that, if $q \neq 0$ and $H_{xx}(x_c) \neq 0$, it is $h_x(x_c) \neq 0$, thus justifying the assumption used before. Then, the chosen solution of (50) will be

$$K = \left[\pm \sqrt{\frac{q^{2/3}H_{xx}(x^c)}{3g^{1/3}}}, 0 \right]^T.$$

More precisely, when applying Algorithm 5 or Algorithm 4, a threshold $\epsilon > 0$ is selected to detect critical states. If a system (50) has to be solved in which

$$|Fr(U^*) - 1| < \epsilon,$$

at a point \bar{x} , then

- If \bar{x} is not close to a minimum point of H , it is assumed that is not possible to find a smooth stationary solution that solves the problem and the algorithm is stopped.
- Otherwise, the selected solution of (50) is

$$- \quad K = \left[\sqrt{\frac{q^{2/3}H_{xx}(\bar{x})}{3g^{1/3}}}, 0 \right]^T, \text{ if } h \text{ is increasing close to } \bar{x},$$

$$- K = \left[-\sqrt{\frac{q^{2/3}H_{xx}(\bar{x})}{3g^{1/3}}}, 0 \right]^T, \text{ if } h \text{ is decreasing close to } \bar{x},$$

and the algorithm goes on.

4.3.1. Test 3.1: Transcritical Smooth Stationary Solution

The objective with this test is to check the correct behavior of the numerical methods of this work in presence of critical states. Following Reference [38], we consider here a smooth transcritical stationary solution for the 1D shallow water system on the interval [0, 3]. The integration time is $T = 1$, and the bathymetry is given by

$$H(x) = \begin{cases} -0.25(1 + \cos(5\pi(x + 0.5))) & \text{if } 1.3 \leq x \leq 1.7, \\ 0 & \text{otherwise.} \end{cases} \tag{69}$$

The initial condition corresponds to the solution of the Cauchy problem (61) with initial condition $h(0) = 1.67750727$ and $q(0) = 2.5$. It can be checked that it has a critical state at $x^c = 1.5$. The shallow water system is integrated up to time $T = 1$, setting q downstream, and the water height upstream. The difference between the stationary and the numerical solutions computed with NWBM*i*, $i = 1, 2, 3$ and CLWBM*i*, $i = 1, 2, 3$ are shown in Figure 4. L^1 Errors are shown in Tables 9 and 10. Notice that Reference [38] obtained similar results. The computational costs are shown in Table 11.

Table 9. Test 3.1. Differences in L^1 -norm with respect to the stationary solution and convergence rates for NWBM*i*, $i = 1, 2, 3$.

Cells	Error ($i = 1$) h	Order	Error ($i = 2$) h	Order	Error ($i = 3$) h	Order
100	4.99×10^{-2}	-	7.63×10^{-3}	-	5.99×10^{-3}	-
200	1.31×10^{-2}	1.923	1.27×10^{-3}	2.583	8.86×10^{-4}	2.757
400	3.87×10^{-3}	1.766	1.84×10^{-4}	2.790	6.20×10^{-5}	3.838
800	1.56×10^{-3}	1.314	5.31×10^{-5}	1.794	6.88×10^{-6}	3.172
Cells	Error ($i = 1$) q	Order	Error ($i = 2$) q	Order	Error ($i = 3$) q	Order
100	1.28×10^{-1}	-	1.89×10^{-2}	-	1.72×10^{-2}	-
200	2.81×10^{-2}	2.188	3.18×10^{-3}	2.575	2.35×10^{-3}	2.866
400	9.29×10^{-3}	1.161	4.81×10^{-4}	2.724	2.09×10^{-4}	3.489
800	4.09×10^{-3}	1.168	1.37×10^{-4}	1.812	2.29×10^{-5}	3.190

Table 10. Test 3.1. Differences in L^1 -norm with respect to the stationary solution for CLWBM*i*, $i = 1, 2, 3$.

Cells	Error ($i = 1$)		Error ($i = 2$)		Error ($i = 3$)	
	h	q	h	q	h	q
100	1.46×10^{-15}	2.13×10^{-15}	2.80×10^{-16}	1.44×10^{-15}	2.88×10^{-16}	3.63×10^{-15}
200	4.95×10^{-16}	3.00×10^{-16}	3.03×10^{-15}	1.44×10^{-14}	3.94×10^{-14}	5.53×10^{-14}
400	2.94×10^{-16}	1.74×10^{-15}	4.75×10^{-16}	1.20×10^{-15}	8.89×10^{-14}	1.45×10^{-13}
800	1.50×10^{-15}	6.92×10^{-15}	3.25×10^{-16}	1.21×10^{-15}	1.04×10^{-13}	1.55×10^{-13}

Table 11. Test 3.1. Computational cost (milliseconds). $t = 1$ s.

Cells	Order (i)	NWBM i	CLWBM i
100	1	20	30
	2	30	130
	3	70	320
200	1	30	60
	2	80	370
	3	200	1100

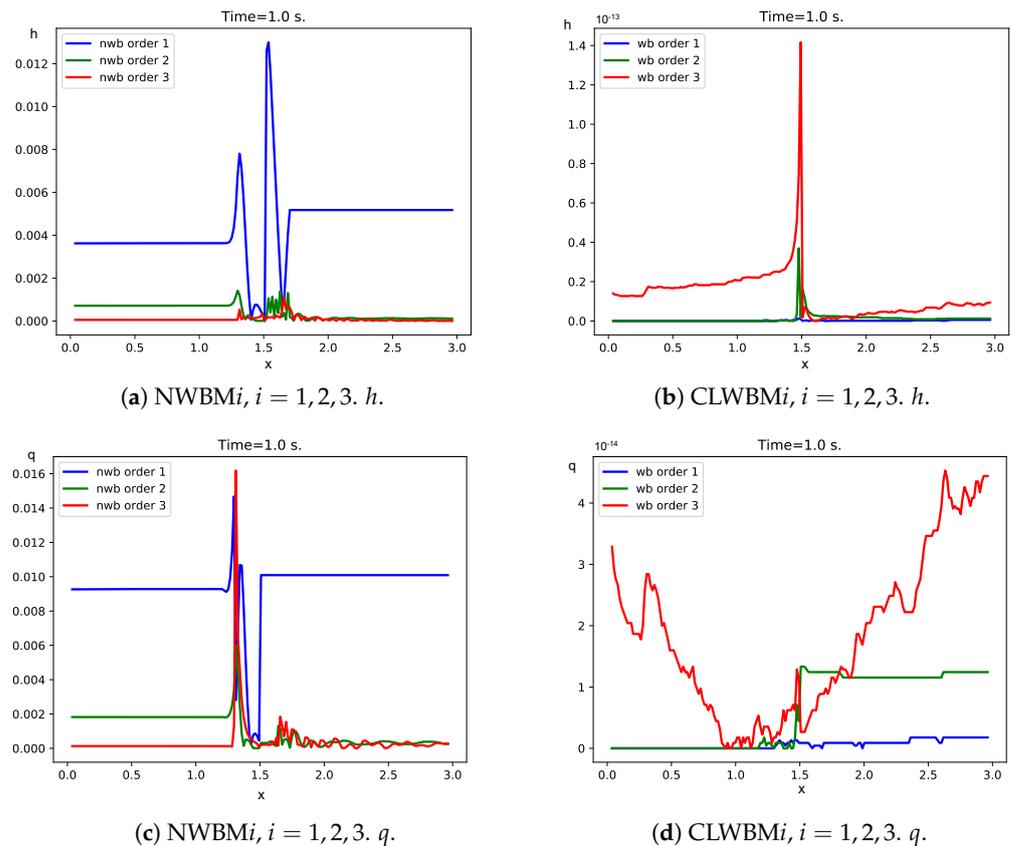


Figure 4. Test 3.1. Reference and numerical solutions: differences with the stationary solution at time $t = 1$ s for the 200-cell mesh.

4.3.2. Test 3.2: Perturbation of a Smooth Transcritical Stationary Solution

In this test, we consider the evolution of a small perturbation of the previous transcritical stationary solution. More precisely, a small perturbation of the water thickness of amplitude $\Delta h = 0.02$ is set at the interval $[1.1, 1.2]$, that is in the subcritical region. Therefore, this perturbation splits in two waves, one moving downstream and the other upstream, as shown in Figures 5 and 6 at time $t = 0.1$ s. Figures 5 and 6 show the difference between the stationary and the numerical solutions computed with NWBM i , $i = 1, 2, 3$ and CLWBM i , $i = 1, 2, 3$ at times $t = 0.1, 5$ s. A reference solution is also computed using a first order well-balanced method with fine mesh (2000 cells).

Observe that, as in the previous examples, only the well-balanced methods are able to preserve the stationary solution. The non-well-balanced method generates strong perturbations that are bigger than the initial perturbation. Differences between the stationary and the numerical solutions in L^1 -norm at time $t = 5$ s for a 200-cell mesh are shown in Table 12. Finally, let us point out that the third order well-balanced method presents a small spurious oscillation in the location of the critical point, that disappears with time. This is not the case for the first and second order well-balanced methods.

Table 12. Test 3.2. Differences in L^1 -norm for NWB M_i and CLW $B M_i$ ($i = 1, 2, 3$) with respect to the stationary solution using a 200-cell mesh at time $t = 5$ s.

Method	Error ($i = 1$)		Error ($i = 2$)		Error ($i = 3$)	
	h	q	h	q	h	q
NW $B M_i$	1.01×10^{-2}	2.13×10^{-2}	4.83×10^{-4}	2.00×10^{-3}	2.70×10^{-4}	1.30×10^{-3}
CLW $B M_i$	5.11×10^{-14}	4.32×10^{-16}	1.95×10^{-14}	6.47×10^{-15}	7.89×10^{-14}	1.38×10^{-13}

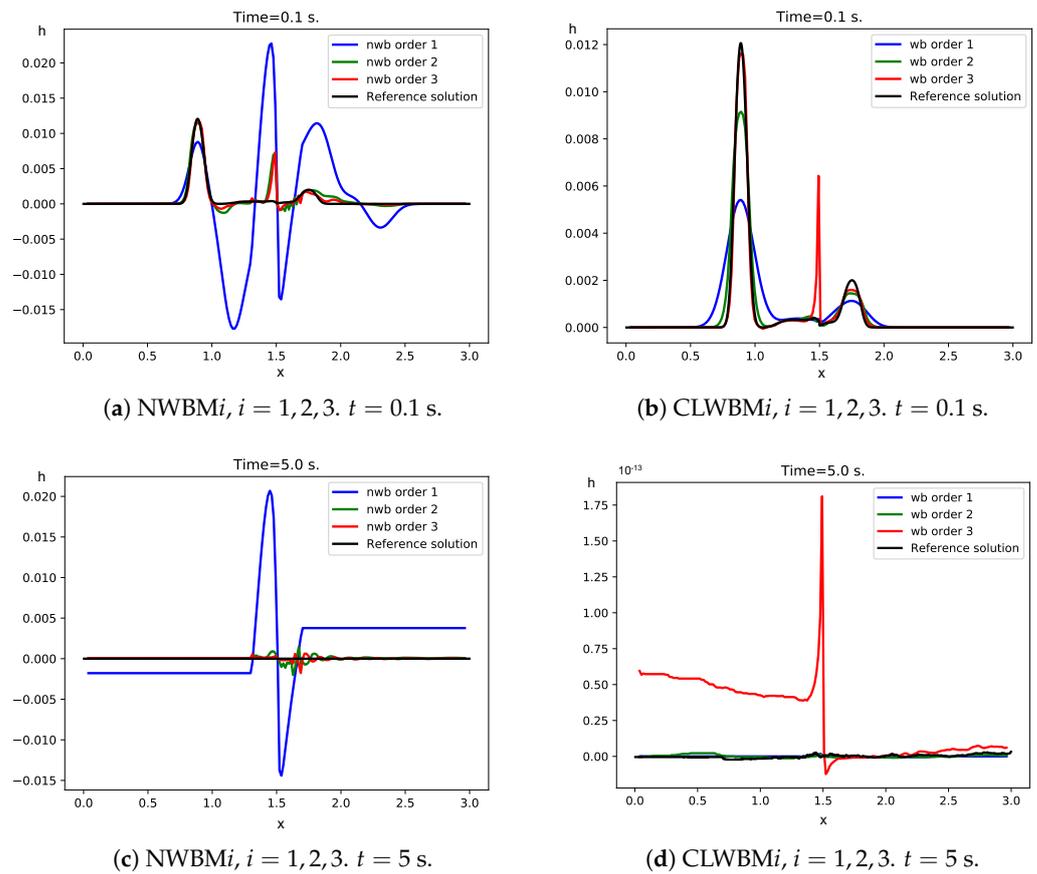


Figure 5. Test 3.2. Reference and numerical solutions: differences with the stationary solution at times $t = 0.1, 5$ s for h with a 200-cell mesh.

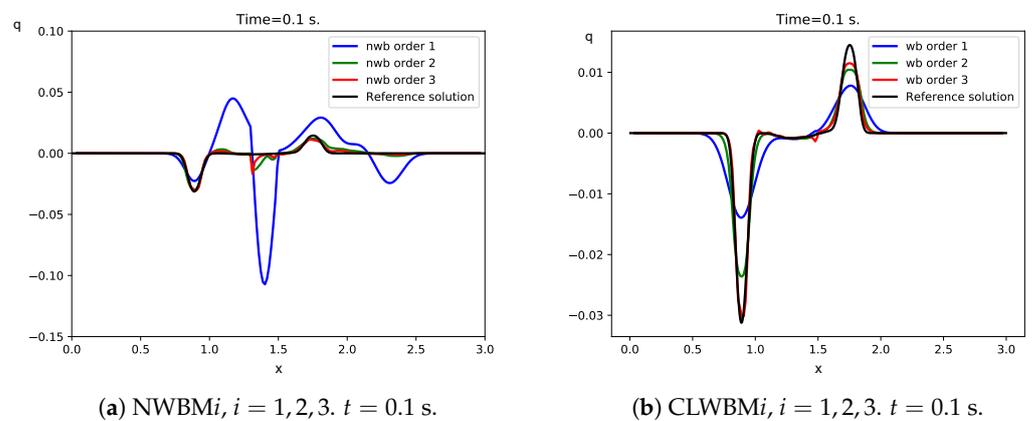


Figure 6. Cont.

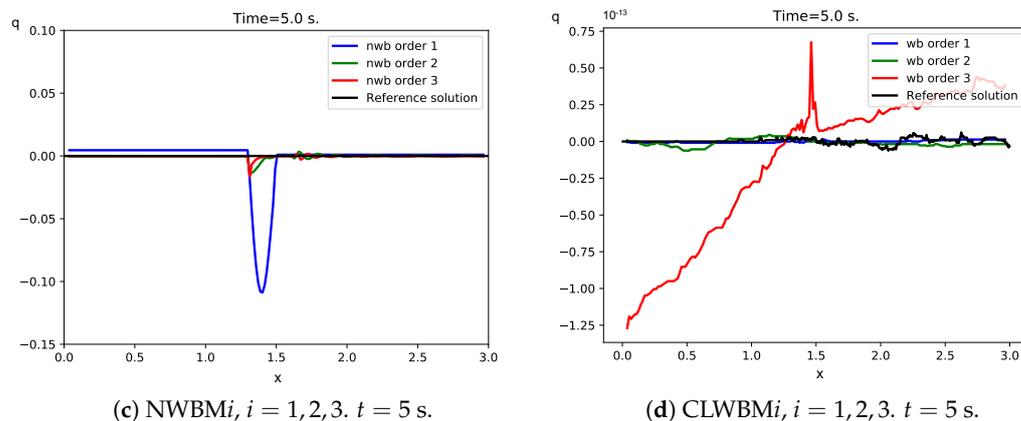


Figure 6. Test 3.2. Reference and numerical solutions: differences with the stationary solution at times $t = 0.1, 5$ s for q for a 200-cell mesh.

4.3.3. Test 3.3: Perturbation of a Smooth Subcritical Stationary Solution

Let us consider evolution of a small perturbation of a smooth subcritical stationary solution. The initial condition $U_0(x) = (h_0(x), q_0(x))^T$ is given by:

$$h_0(x) = \begin{cases} h^*(x) + 0.02, & \text{if } 0.7 \leq x \leq 1.0, \\ h^*(x), & \text{otherwise,} \end{cases}$$

$$q_0(x) = q^*(x),$$

where $U^*(x) = (h^*(x), q^*(x))^T$ is the solution of the Cauchy problem

$$\begin{cases} q_x = 0, \\ (-u^2 + gh)h_x = ghH_x, \\ h(0) = 2, q(0) = 3.5. \end{cases} \tag{70}$$

The depth function is given again by (69).

Figure 7 shows the differences between the stationary solution and the numerical solutions obtained with NWBMi, $i = 1, 2, 3$ and CLWBMi, $i = 1, 2, 3$ at times $t = 0.1, 5$ for h (the graphs are similar for q). A reference solution has been computed with a first order well-balanced scheme on a fine mesh (3200 cells).

Notice that, since the perturbation amplitude is small enough, the well-balanced methods perform much better than the non-well-balanced methods, that generates spurious waves of amplitudes bigger than the original perturbation. In Table 13, the differences with respect to the stationary solution at time $t = 5$ s are shown for the 100-cell mesh. In order to check the efficiency of the methods, Figure 8 shows the errors in L^1 -norm with respect to the reference solution versus the CPU times in milliseconds for the NWBM2 and CLWBM2 at time $t = 1$ s, when the perturbation is still in the domain. Similar results have been obtained for the first and third order methods. Notice that the CLWBM are computationally more efficient. See, for instance, that, in order to obtain an error of approximately 2×10^{-4} , the computational cost for the NWBM2 is increased by a factor of about 2.2 with respect to the CLWBM2.

Table 13. Test 3.3. Differences in L^1 -norm for NWBMi and CLWBMi ($i = 1, 2, 3$) with respect to the stationary solution for the 100-cell mesh at time $t = 5$ s.

Method	Error ($i = 1$)		Error ($i = 2$)		Error ($i = 3$)	
	h	q	h	q	h	q
NWBMi	5.08×10^{-2}	1.94×10^{-1}	9.33×10^{-3}	3.51×10^{-2}	6.02×10^{-3}	2.12×10^{-2}
CLWBMi	1.81×10^{-15}	5.54×10^{-15}	1.95×10^{-15}	4.45×10^{-15}	2.46×10^{-14}	5.20×10^{-14}

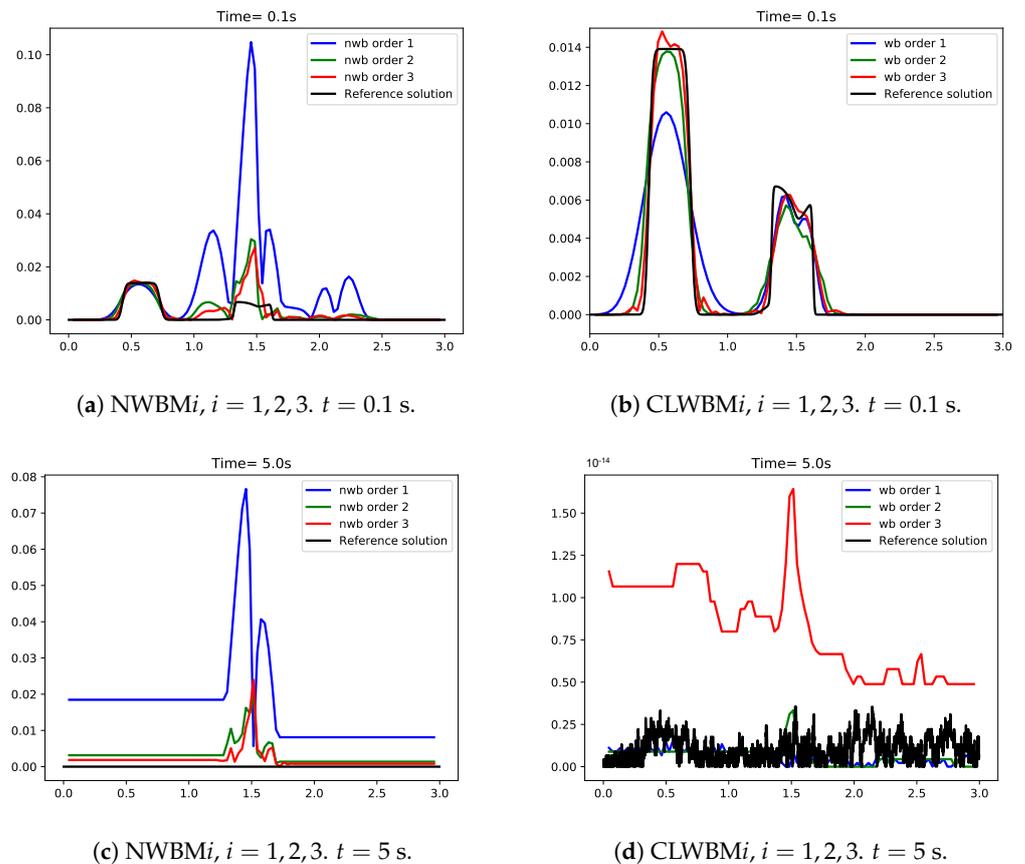


Figure 7. Test 3.3. Reference and numerical solutions: differences with the stationary solution at times $t = 0.1$ s and $t = 5$ s for h for a 100-cell mesh.

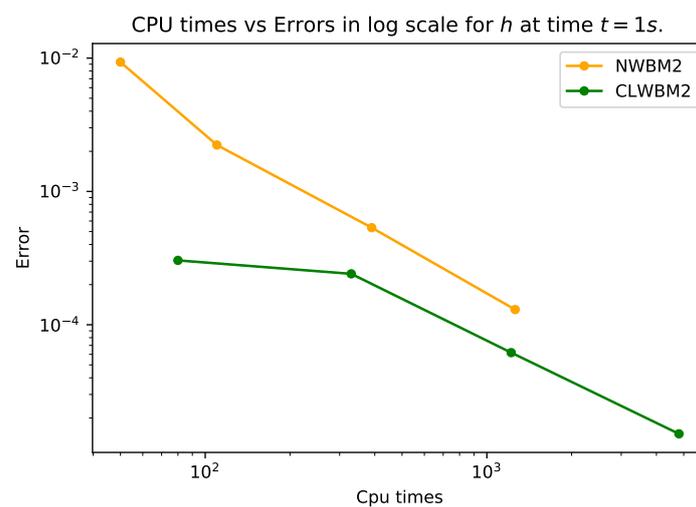


Figure 8. Test 3.3. Errors in L^1 -norm with respect to the reference solution versus CPU times in milliseconds for the NWBM2 and CLWBM2 at time $t = 1$ s.

4.4. Problem 4: Shallow Water Equations with Manning Friction

Let us consider the shallow water equations with Manning friction:

$$\begin{cases} h_t + q_x = 0, \\ q_t + \left(\frac{q^2}{h} + \frac{1}{2}gh^2 \right)_x = ghH_x - \frac{kq|q|}{h^\eta}. \end{cases} \quad (71)$$

In the previous, k is the Manning friction coefficient, and η is set to $\frac{7}{3}$. The stationary solutions satisfy the ODE system

$$\begin{cases} (-u^2 + gh)h_x = ghH_x - \frac{kq|q|}{h^\eta}, \\ q_x = 0. \end{cases} \tag{72}$$

In what follows, we consider some experiments taken from Reference [47].

4.4.1. Test 4.1

Let us consider first a moving equilibria test with constant bathymetry. Therefore, the smooth stationary solutions satisfy

$$\begin{cases} (-q^2/h^2 + gh)h_x = -\frac{kq|q|}{h^\eta}, \\ q_x = 0. \end{cases} \tag{73}$$

As in Reference [47], we consider the interval the space interval $[0, 1]$, discretized with 200 cells. Now, we consider a subcritical steady state with $q(0) = -1$ and $h(0)$ a positive root of the non-linear function $\xi(h)$, obtained by integrating (73), given by

$$\xi(h) = -\frac{q_0^2}{\eta - 1} (h^{\eta-1} - h_0^{\eta-1}) + \frac{g}{\eta + 2} (h^{\eta+2} - h_0^{\eta+2}) + kq_0|q_0|(x - x_0), \tag{74}$$

where $x_0 = -\Delta x$, $x = 0$, $q_0 = q(0)$ and $h_0 = h^c$, with h^c defined by

$$h^c = \left(\frac{q_0^2}{g} \right)^{\frac{1}{3}}.$$

In this experiment, $k = 1$, which enforces the stiff character of the friction term. As boundary conditions, q is set downstream and h upstream, and the system is integrated up to $t = 1$. Due to the stiff character of the friction term, the explicit non-well-balanced methods turn out to be unstable for small values of h and/or big values of k , unless a very small time step is used. If an implicit discretization of the stiff source term is considered, NWBMi, $i = 1, 2, 3$ methods do not explode, yet there is not convergence to any stationary solution. The stability restriction is much less severe for the well-balanced schemes. For these reasons, only the numerical solutions obtained with well-balanced schemes are shown.

Figure 9 shows the discrepancies between the stationary and the numerical solutions at $t = 1$ s with CLWBMi, $i = 1, 2, 3$; Table 14 shows the L^1 errors for a mesh with 200 cells.

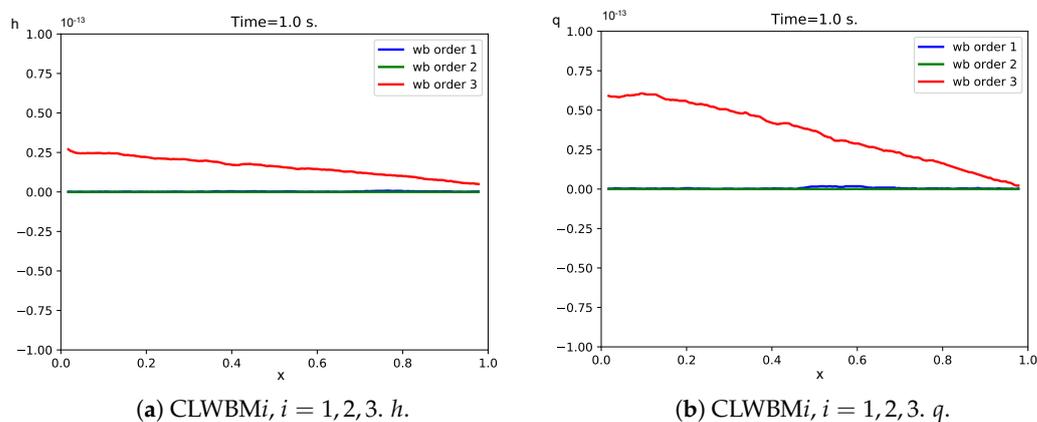


Figure 9. Test 4.1. Differences between the stationary and the numerical solutions for CLWBMi, $i = 1, 2, 3$ at time $t = 1$ s for a 200-cell mesh.

Table 14. Test 4.1. Differences in L^1 -norm with respect to the stationary solution for CLWB*Mi* ($i = 1, 2, 3$) for the 200-cell mesh at time $t = 1$ s.

Method	Error ($i = 1$)		Error ($i = 2$)		Error ($i = 3$)	
	h	q	h	q	h	q
CLWB <i>Mi</i>	2.24×10^{-16}	5.06×10^{-16}	0.00	5.56×10^{-19}	1.58×10^{-14}	3.38×10^{-14}

4.4.2. Test 4.2

Now, we consider a numerical experiment that corresponds to a small perturbation of the previous stationary solution $U^*(x) = [h^*(x), q^*(x)]^T$:

$$h_0(x) = \begin{cases} h^*(x) + 0.2, & \text{if } \frac{3}{7} \leq x \leq \frac{4}{7}, \\ h^*(x), & \text{otherwise,} \end{cases}$$

$$q_0(x) = q^*(x).$$

As in Reference [47], we consider a mesh with 100 cells, and the model is run up to $t = 9$ s.

As in the previous test case, only the well-balanced schemes will be considered. Figure 10 shows the discrepancies between the stationary and the numerical solutions for the water depth computed with CLWB*Mi*, $i = 1, 2, 3$ at the time steps $t = 0.06$ s and $t = 9$ s (the figures are similar for q). In black, we plot a reference solution computed with a 1st-order well-balanced scheme with a 1600-cell mesh.

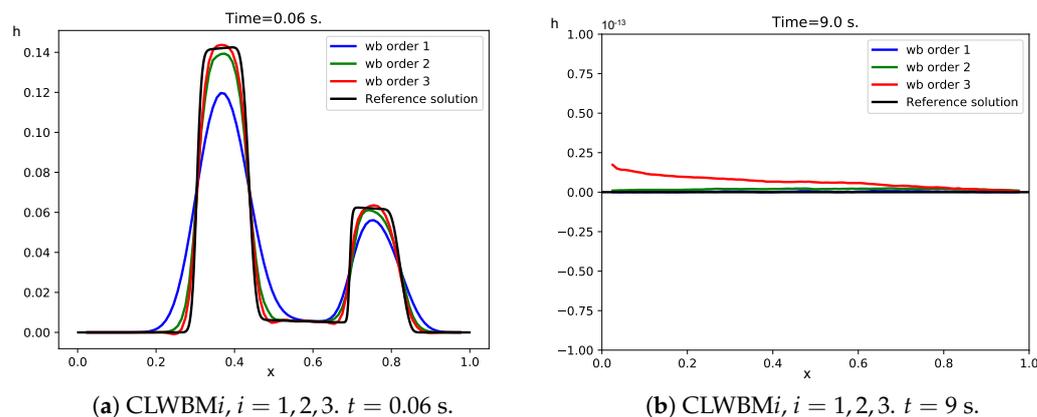


Figure 10. Test 4.2. Reference and numerical solutions: differences with the stationary solution for CLWB*Mi*, $i = 1, 2, 3$, at times $t = 0.06, 9$ s for h on a 100-cell mesh.

As expected, the stationary solution is preserved by the well-balanced schemes after the perturbation left the domain. Table 15 shows the L^1 errors with respect to the steady state at $t = 9$ s for 100-cell mesh.

Table 15. Test 4.2. Differences in L^1 -norm with respect to the stationary solution for CLWB*Mi* ($i = 1, 2, 3$) for the 100-cell mesh at time $t = 9$ s.

Method	Error ($i = 1$)		Error ($i = 2$)		Error ($i = 3$)	
	h	q	h	q	h	q
CLWB <i>Mi</i>	2.99×10^{-16}	3.97×10^{-16}	1.81×10^{-15}	2.76×10^{-15}	6.50×10^{-14}	1.77×10^{-14}

4.4.3. Test 4.3

We end this subsection with two numerical experiments whose objective is show the efficiency of the schemes to preserve steady states that involve a varying bathymetry and

friction. In this case, k takes the value of 0.01 and the space domain is again the interval $[0, 1]$. The bathymetry is given by

$$H(x) = 1 - \frac{1}{2} \frac{e^{\cos(4\pi x)} - e^{-1}}{e - e^{-1}}. \tag{75}$$

The first experiment concerns the preservation of the supercritical stationary solution corresponding to (72) with initial conditions $q(0) = 1$ and $h(0) = 0.3$.

Following Reference [47], we take 100 discretization cells, and the numerical simulation of this experiment is run until $t = 1$ s.

Figure 11 shows the numerical solutions at $t = 1$ s for NWBM i , $i = 1, 2, 3$ and CLWBM i , $i = 1, 2, 3$. Notice that NWBM1 is completely wrong: due to the numerical diffusion, the supercritical regime is lost at the right side of the domain, which causes a shock wave traveling to the left (of course, this behavior disappears when the mesh is fine enough). Therefore, in Figure 12, where the differences with the stationary solution are plotted, only the results for NWBM i , $i = 2, 3$ and for CLWBM i , $i = 1, 2, 3$ are shown. Table 16 shows the errors corresponding to the different methods for the 100-cell mesh. We also consider a reference solution computed with a well-balanced first order scheme on a 1600-cell mesh.

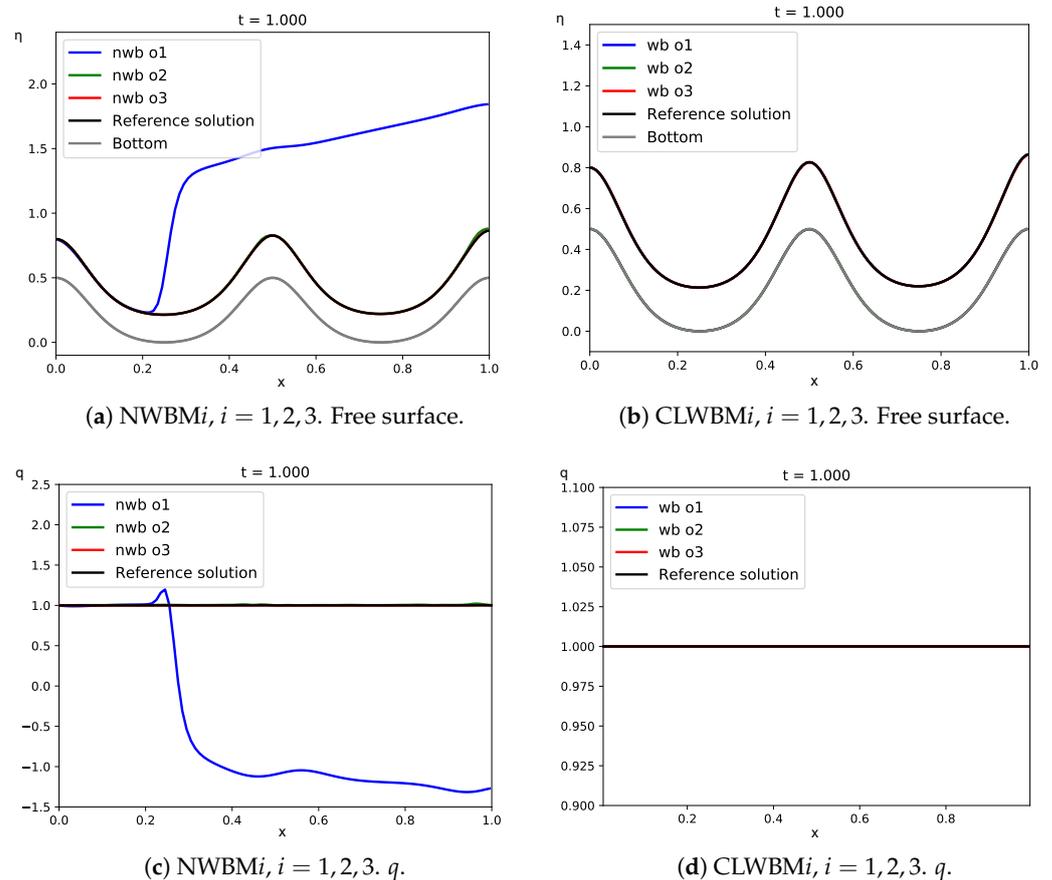


Figure 11. Test 4.3. 100-cell numerical solutions and reference solution at time $t = 1$ s.

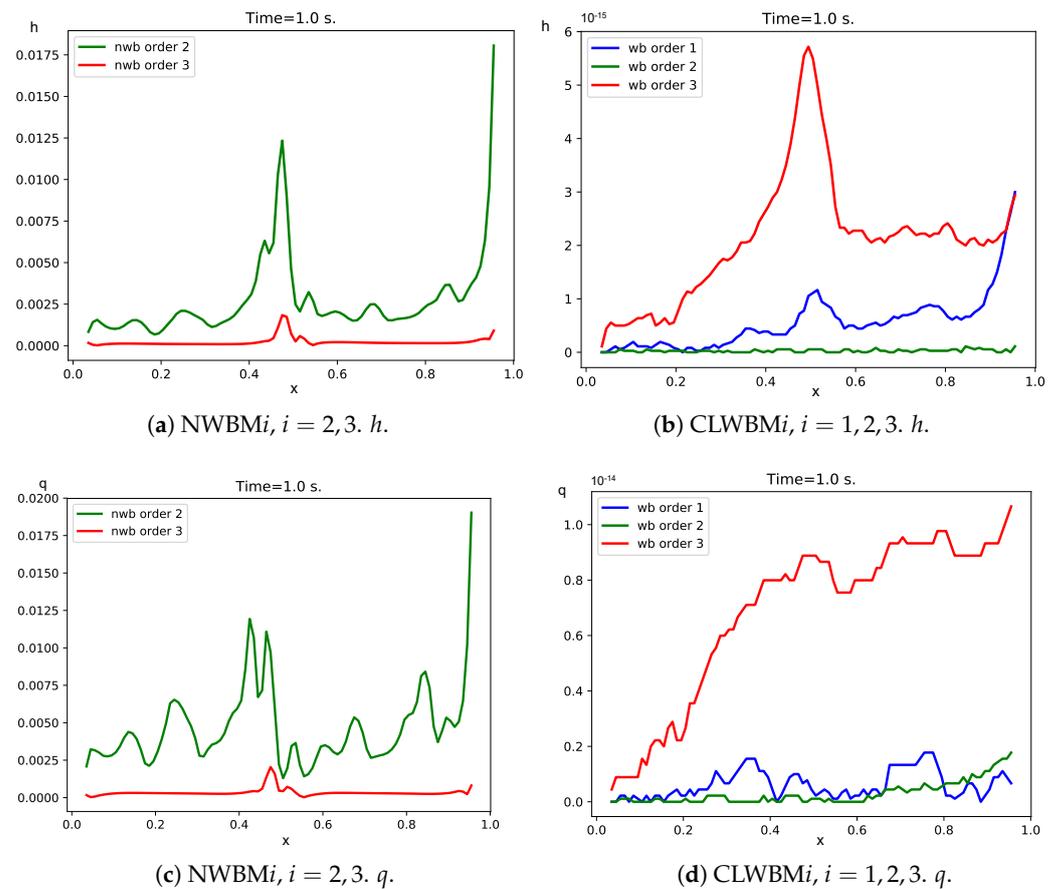


Figure 12. Test 4.3. Differences between the stationary and the numerical solutions at time $t = 1$ s with a 100-cell mesh.

Table 16. Test 4.3. Differences in L^1 -norm with respect to the stationary solution for NWBMi and CLWBMi ($i = 1, 2, 3$) for the 100-cell mesh at time $t = 1$ s.

Method	Error ($i = 1$)		Error ($i = 2$)		Error ($i = 3$)	
	h	q	h	q	h	q
NWBMi	8.28×10^{-1}	1.54	3.61×10^{-3}	4.88×10^{-3}	3.65×10^{-3}	4.30×10^{-4}
CLWBMi	7.03×10^{-16}	5.85×10^{-16}	3.22×10^{-17}	3.75×10^{-16}	2.14×10^{-15}	6.87×10^{-15}

4.4.4. Test 4.4

Following Reference [47], the very last experiment in this subsection focuses on a perturbation of the aforementioned steady state. The initial condition $U_0(x) = [h_0(x), q_0(x)]^T$ is given by:

$$h_0(x) = \begin{cases} h^*(x) + 0.05, & \text{if } x \in \left[\frac{2}{7}, \frac{3}{7}\right] \cup \left[\frac{4}{7}, \frac{5}{7}\right], \\ h^*(x), & \text{otherwise,} \end{cases}$$

$$q_0(x) = \begin{cases} q^*(x) + 0.5, & \text{if } x \in \left[\frac{2}{7}, \frac{3}{7}\right] \cup \left[\frac{4}{7}, \frac{5}{7}\right], \\ q^*(x), & \text{otherwise,} \end{cases}$$

where $U^*(x) = [h^*(x), q^*(x)]^T$ is the stationary solution considered in the previous experiment. As in Reference [47], we use a 100-cell mesh for the numerical simulation on the domain $[0, 1]$, and the computations are carried out until the final time $t = 2$ s.

Observe that there are not big differences among the solutions during for small times. Nevertheless, for the NNWBM1, the supercritical regime is lost on the right of the domain,

causing a shock traveling to the left, as in Test 4.3. Figure 13 shows then the differences between the numerical solutions and the stationary solution for h with NWB M_i , $i = 1, 2, 3$ at time $t = 0.015$ and $i = 2, 3$ at time $t = 2$, and with CLW BM_i , $i = 1, 2, 3$ at times $t = 0.015$ and $t = 2$ (the plots are similar for q). Again, in black, we plot a reference solution computed on a 1600-cell mesh with a first order well-balanced scheme.

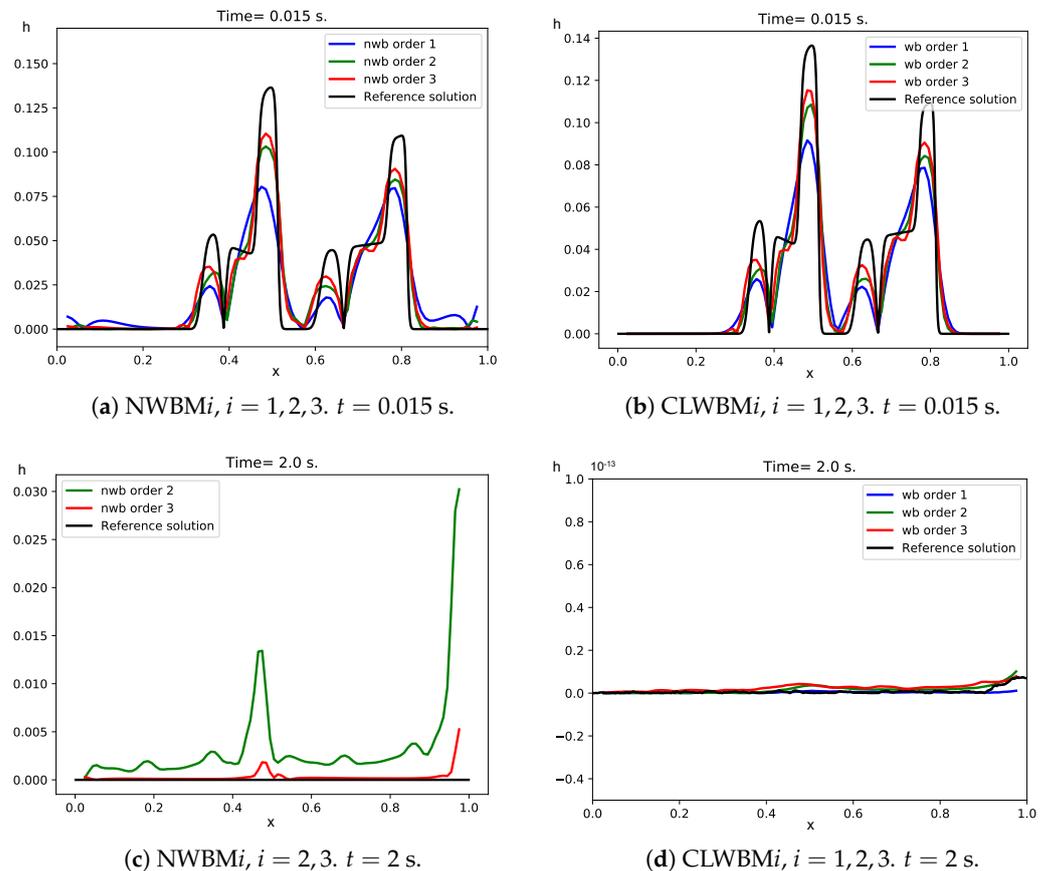


Figure 13. Test 4.4. Reference and numerical solutions: differences with the stationary solution at times $t = 0.015, 2$ s for h , for a 100-cell mesh.

Once more, well-balanced schemes are able to recover the stationary states after the perturbation left the domain. Table 17 shows the differences in L^1 -norm with respect to the stationary solution at time $t = 2$ s for a mesh with 100 cells.

Table 17. Test 4.4. Differences in L^1 -norm with respect to the stationary solution for NWB M_i and CLW BM_i ($i = 1, 2, 3$) at time $t = 2$ s for a 100-cell mesh.

Method	Error ($i = 1$)		Error ($i = 2$)		Error ($i = 3$)	
	h	q	h	q	h	q
NW BM_i	2.42	6.12	3.57×10^{-3}	4.87×10^{-3}	1.39×10^{-3}	4.30×10^{-4}
CLW BM_i	3.73×10^{-16}	3.60×10^{-16}	1.80×10^{-15}	1.99×10^{-15}	2.64×10^{-15}	8.93×10^{-15}

4.5. Problem 5: Compressible Euler Equations with Gravitational Force

Finally, let us consider the compressible Euler equations for gas dynamics with gravity:

$$\begin{cases} \rho_t + (\rho u)_x = 0, \\ (\rho u)_t + (\rho u^2 + p)_x = -\rho H_x, \\ E_t + (u(E + p))_x = -\rho u H_x. \end{cases} \tag{76}$$

Here, u is the velocity, $p \geq 0$ the pressure, $\rho \geq 0$ the density, $q = \rho u$ the momentum, E the total energy per unit volume, and $H(x)$ is the gravitational potential. The internal energy e is defined by $\rho e = E - \frac{1}{2}\rho u^2$, and pressure could be determined from the internal energy using the equation of state. In this work, we suppose an ideal gas; therefore,

$$p = (\gamma - 1)\rho e, \tag{77}$$

$\gamma > 1$ being the adiabatic constant. Here, it is set to $\gamma = 1.5$.

Note that (76) is a particular case of (1) with $N = 3$,

$$U = \begin{pmatrix} \rho \\ \rho u \\ E \end{pmatrix}, \quad f(U) = \begin{pmatrix} \rho u \\ \rho u^2 + p \\ u(E + p) \end{pmatrix}, \quad S(U) = \begin{pmatrix} 0 \\ -\rho \\ -\rho u \end{pmatrix}.$$

Supposing that the system is strictly hyperbolic and considering (77), the stationary solutions satisfy the ODE system:

$$\begin{cases} q_x = 0, \\ \frac{d\hat{U}}{dx} = G(x, \hat{U}), \end{cases} \tag{78}$$

where

$$\hat{U} = \begin{pmatrix} \rho \\ E \end{pmatrix}, \quad G(x, \hat{U}) = - \begin{pmatrix} \frac{\rho}{c^2 - u^2} \\ \frac{\rho}{\gamma - 1} \left(1 + \frac{3 - \gamma}{2} \frac{u^2}{c^2 - u^2} \right) \end{pmatrix} H_x,$$

where $c = \sqrt{\gamma p / \rho}$ is the sound speed, which is valid for stationary regular solution. Notice that, for regular solutions of the Euler equations, the entropy is constant along material lines. So that, for stationary solutions, $us_x = 0$, where s denotes the entropy density. Therefore, if u does not vanish, it follows that these stationary solutions are isentropic.

Test 5.1

In this test, we consider the space domain $[-1, 1]$, and the system is integrated up to time $t = 5$ s. The gravity potential is $H(x) = x$.

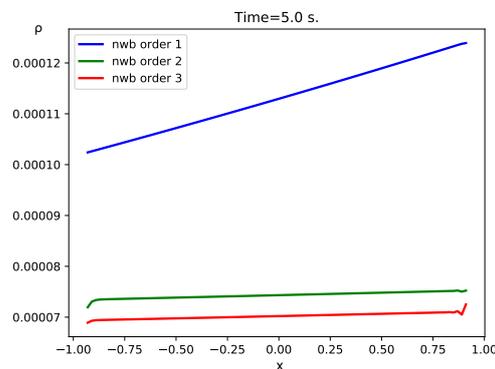
As initial condition, a supersonic stationary solution corresponding to the solution of the Cauchy problem

$$\begin{cases} q_x = 0, \\ \frac{d\hat{U}}{dx} = G(x, \hat{U}), \\ \rho(-1) = 1, q(-1) = 10, E(-1) = 52 \end{cases} \tag{79}$$

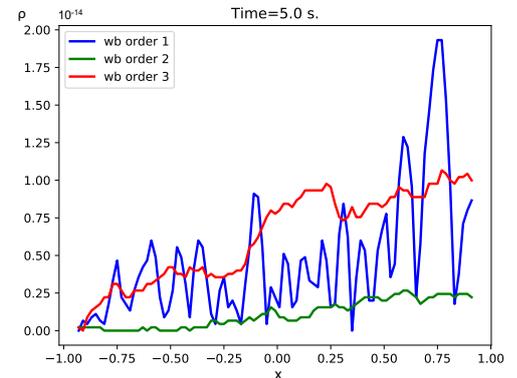
is computed with the Gauss-Legendre collocation method. Density, momentum, and energy is set downstream, and open free boundary conditions upstream. CFL parameter is set to 0.9.

Figures 14–16 show the differences between the stationary and the numerical solutions for NWB*M*_{*i*} and CLWBM*i*, $i = 1, 2, 3$ for ρ, q , and E , respectively; Table 18 shows the errors for for the 100-cell mesh with NWB*M*_{*i*} and CLWBM*i*, $i = 1, 2, 3$ at time $t = 5$ s. Table 19 shows the computational cost for NWB*M*_{*i*}, CWBM*i*, and CLWBM*i*, $i = 1, 2, 3$. Notice that CLWBM are significantly more efficient than CWBM. We have carried out similar experiments with a perturbation of this stationary solution as initial condition, and we obtained the expected results, i.e., only the well-balanced methods are able to exactly

recover the steady state solution. For the sake of conciseness, we do not present the tests here.

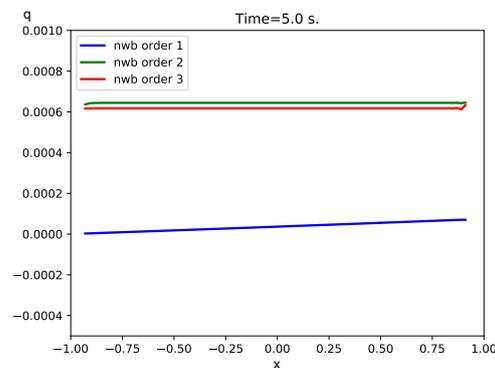


(a) NWB $M_i, i = 1, 2, 3$. Density

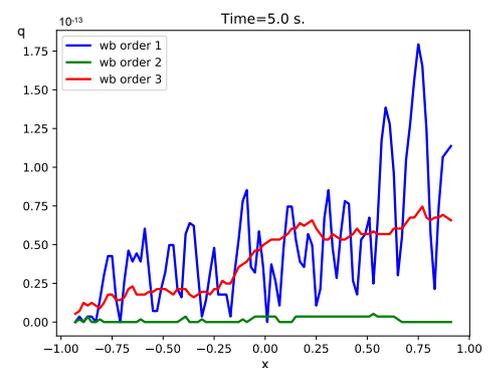


(b) CLWBM $i, i = 1, 2, 3$. Density

Figure 14. Test 5.1. Differences between the stationary and the numerical solutions at time $t = 5$ s for ρ with a mesh of 100 cells.

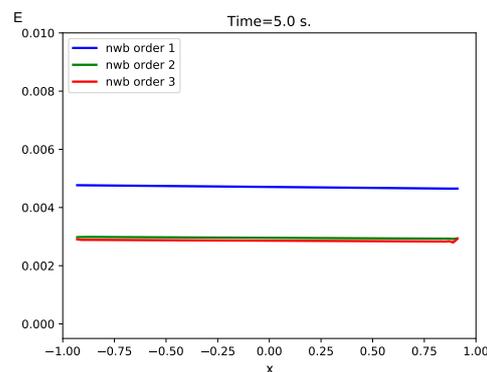


(a) NWB $M_i, i = 1, 2, 3$. Momentum

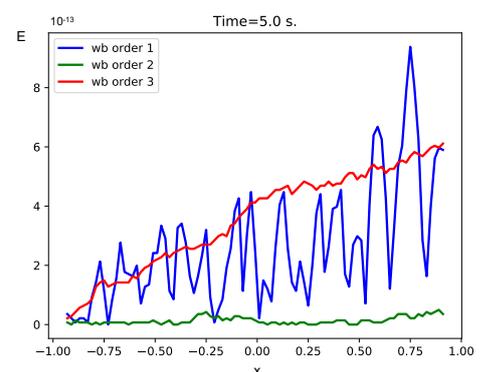


(b) CLWBM $i, i = 1, 2, 3$. Momentum

Figure 15. Test 5.1. Differences between the stationary and the numerical solutions at time $t = 5$ s for q with a mesh of 100 cells.



(a) NWB $M_i, i = 1, 2, 3$. Energy



(b) CLWBM $i, i = 1, 2, 3$. Energy

Figure 16. Test 5.1. Differences between the stationary and the numerical solutions at time $t = 5$ s for E with a mesh of 100 cells.

Table 18. Test 5.1. Differences in L^1 -norm with respect to the stationary solution for NWBM*i* and CLWBM*i* ($i = 1, 2, 3$) for the mesh with 100 cells at time $t = 5$ s.

Method	Error ($i = 1$)	Error ($i = 2$)	Error ($i = 3$)
h			
NWBM <i>i</i>	2.23×10^{-4}	9.41×10^{-5}	9.51×10^{-3}
CLWBM <i>i</i>	6.97×10^{-15}	6.58×10^{-14}	3.20×10^{-13}
q			
NWBM <i>i</i>	1.45×10^{-4}	1.28×10^{-3}	5.95×10^{-3}
CLWBM <i>i</i>	2.22×10^{-15}	2.81×10^{-15}	2.77×10^{-14}
E			
NWBM <i>i</i>	1.38×10^{-4}	1.22×10^{-3}	5.78×10^{-3}
CLWBM <i>i</i>	1.24×10^{-14}	8.13×10^{-14}	7.15×10^{-13}

Table 19. Test 5.1. Computational cost (milliseconds) for the mesh with 100 cells. $t = 5$ s.

Order (i)	NWBM <i>i</i>	CWBM <i>i</i>	CLWBM <i>i</i>
1	120	4480	220
2	250	8960	1440
3	670	19430	4750

5. Conclusions

Following Reference [36], we propose a family of well-balanced high-order numerical schemes that can be applied to general 1D balance laws. Note that the main difficulty of the method proposed in Reference [36] comes from the fact that local stationary solutions must be computed at every cell that may require the knowledge of the stationary states. When solving the ODE (2) for the stationary solution is not possible or too costly, the first step has to be computed numerically: what we propose here is to build both the stationary solutions and the local solvers for the first step on the basis of the collocation RK methods. The well-balanced property of these numerical methods is precisely stated, and we have also introduced a general technique that allows us to deal with resonant problems for any 1D systems of balance laws.

We have considered a good number of examples: Burgers equation, shallow water system with friction, and Euler equations with gravity to check the well-balanced property of the methods and to test their efficiency. In some of these test cases, the stationary solutions are known either in implicit or explicit form, while, in others, the only information comes from the ODE (2) that the stationary solutions satisfy: in the former case, we can compare the efficiency of the new implementation, while, in the latter, we can show the generality of the methods.

Although well-balanced schemes are more costly, they are more effective than standard non-well-balanced methods when they are applied to the propagation of small perturbation around an equilibria or when long-time integration is required. Furthermore, the tests show that the strategy based on the collocation RK methods introduced in this work is more efficient than the one presented in Reference [39].

The methods considered here can be applied in principle to any system of balance laws. For instance, in the context of Euler equations, in Reference [48], a first and second order numerical scheme that is well-balanced for particular 1D radial solutions for the Euler equations in cylindrical coordinates has been developed following the strategy proposed by two of the authors that was later reviewed in Reference [37]. It is possible to use the technique described here to extend the methods in Reference [48] to more general source terms and higher-order. This will be considered in a future work.

Further developments will include applications to:

- Balance laws (1) with non-regular source terms, that is systems for which H has jump discontinuities.
- Multidimensional problems.

Author Contributions: Conceptualization, M.J.C.D., I.G.-B., C.P. and G.R.; methodology, M.J.C.D., I.G.-B., C.P. and G.R.; software, I.G.-B.; validation, M.J.C.D., I.G.-B., C.P. and G.R.; formal analysis, M.J.C.D., I.G.-B., C.P. and G.R.; writing—original draft preparation, I.G.-B.; writing—review and editing, M.J.C.D., I.G.-B., C.P. and G.R.; funding acquisition, M.J.C.D., I.G.-B. and C.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research has been partially supported by the Spanish Government and FEDER through the coordinated Research project RTI2018-096064-B-C1, The Junta de Andalucía research project P18-RT-3163, the Junta de Andalucía-FEDER-University of Málaga Research project UMA18-FEDERJA-161 and the University of Málaga. G. Russo acknowledges partial support from ITN-ETN Horizon 2020 Project ModCompShock, “Modeling and Computation of Shocks and Interfaces”, Project Reference 642768, and from the Italian Ministry of University and Research (MIUR), PRIN Project 2017 (No. 2017KKJP4X entitled “Innovative numerical methods for evolutionary partial differential equations and applications”. I. Gómez-Bueno is also supported by a Grant from “El Ministerio de Ciencia, Innovación y Universidades”, Spain (FPU2019/01541).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Audusse, E.; Bouchut, F.; Bristeau, M.O.; Klein, R.; Perthame, B. A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows. *SIAM J. Sci. Comput.* **2004**, *25*, 2050–2065. [\[CrossRef\]](#)
2. Bouchut, F. *Nonlinear Stability of Finite Volume Methods for Hyperbolic Conservation Laws: And Well-Balanced Schemes for Sources*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2004.
3. Castro, M.; Chacón Rebollo, T.; Fernández-Nieto, E.; Parés, C. On Well-Balanced Finite Volume Methods for Nonconservative Nonhomogeneous Hyperbolic Systems. *SIAM J. Sci. Comput.* **2007**, *29*, 1093–1126. [\[CrossRef\]](#)
4. Chacón Rebollo, T.; Delgado, A.; Fernández-Nieto, E. A family of stable numerical solvers for the shallow water equations with source terms. *Comput. Methods Appl. Mech. Eng.* **2003**, *192*, 203–225. [\[CrossRef\]](#)
5. Chacón Rebollo, T.; Delgado, A.; Fernández-Nieto, E. Asymptotically balanced schemes for non-homogeneous hyperbolic systems—Application to the Shallow Water equations. *C. R. Math.* **2004**, *338*, 85–90. [\[CrossRef\]](#)
6. Varma, R.S.D.; Chandrashekar, P. A second-order well-balanced finite volume scheme for Euler equations with gravity. *Comput. Fluids* **2018**, *181*. [\[CrossRef\]](#)
7. Chandrashekar, P.; Zenk, M. Well-Balanced Nodal Discontinuous Galerkin Method for Euler Equations with Gravity. *J. Sci. Comput.* **2015**, *71*. [\[CrossRef\]](#)
8. Desveaux, V.; Zenk, M.; Berthon, C.; Klingenberg, C. A well-balanced scheme to capture non-explicit steady states in the Euler equations with gravity. *Int. J. Numer. Methods Fluids* **2016**, *81*, 104–127. [\[CrossRef\]](#)
9. Desveaux, V.; Zenk, M.; Berthon, C.; Klingenberg, C. Well-balanced schemes to capture non-explicit steady states. RIPA model. *Math. Comput.* **2015**, *85*, 1. [\[CrossRef\]](#)
10. Gosse, L. A well-balanced flux-vector splitting scheme designed for hyperbolic systems of conservation laws with source terms* 1. *Comput. Math. Appl.* **2000**, *39*, 135–159. [\[CrossRef\]](#)
11. Gosse, L. A Well-Balanced Scheme Using Non-Conservative Products Designed for Hyperbolic Systems of Conservation Laws With Source Terms. *Math. Model. Methods Appl. Sci.* **1999**, *11*. [\[CrossRef\]](#)
12. Gosse, L. Localization effects and measure source terms in numerical schemes for balance laws. *Math. Comput.* **2002**, *71*, 553–582. [\[CrossRef\]](#)
13. Käppeli, R.; Mishra, S. Well-balanced schemes for the Euler equations with gravitation. *J. Comput. Phys.* **2014**, *259*, 199–219. [\[CrossRef\]](#)
14. Greenberg, J.M.; LeRoux, A.Y. A well-balanced scheme for the numerical processing of source terms in hyperbolic equations. *SIAM J. Numer. Anal.* **1996**, *33*, 1–16. [\[CrossRef\]](#)
15. Greenberg, J.; Leroux, A.; Baraille, R.; Noussair, A. Analysis and approximation of conservation laws with source terms. *SIAM J. Numer. Anal.* **1997**, *34*, 1980–2007. [\[CrossRef\]](#)
16. LeVeque, R. Balancing source terms and flux gradients in high-resolution Godunov methods: The quasi-steady wave-propagation algorithm. *J. Comput. Phys.* **1998**, *146*, 346–365. [\[CrossRef\]](#)
17. Lukacova-Medvidova, M.; Noelle, S.; Kraft, M. Well-balanced finite volume evolution Galerkin methods for the shallow water equations. *J. Comput. Phys.* **2007**, *221*, 122–147. [\[CrossRef\]](#)
18. Noelle, S.; Pankratz, N.; Puppo, G.; Natvig, J. Well-balanced finite volume schemes of arbitrary order of accuracy for shallow water flows. *J. Comput. Phys.* **2006**, *213*, 474–499. [\[CrossRef\]](#)

19. Noelle, S.; Xing, Y.; Shu, C.W. High-order well-balanced finite volume WENO schemes for shallow water equation with moving water. *J. Comput. Phys.* **2007**, *226*, 29–58. [[CrossRef](#)]
20. Pelanti, M.; Bouchut, F.; Mangeney, A. A Roe-Type Scheme for Two-Phase Shallow Granular Flows over Variable Topography. *ESAIM Math. Model. Numer. Anal.* **2008**, *42*, 851–885. [[CrossRef](#)]
21. Perthame, B.; Simeoni, C. A kinetic scheme for the Saint-Venant system with a source term. *Calcolo* **2001**, *38*, 201–231. [[CrossRef](#)]
22. Perthame, B.; Simeoni, C. Convergence of the upwind interface source method for hyperbolic conservation laws. In *Hyperbolic Problems: Theory, Numerics, Applications*; Springer: Berlin/Heidelberg, Germany, 2003; pp. 61–78.
23. Russo, G.; Khe, A. High order well balanced schemes for systems of balance laws. *Proc. Symp. Appl. Math.* **2009**, *67*, 919–928. [[CrossRef](#)]
24. Tang, H.; Tang, T.; Xu, K. A gas-kinetic scheme for shallow-water equations with source terms. *Z. Angew. Math. Phys.* **2004**, *55*, 365–382. [[CrossRef](#)]
25. Klingenberg, C.; Touma, R.; Koley, U. Well-Balanced Unstaggered Central Schemes for the Euler Equations with Gravitation. *SIAM J. Sci. Comput.* **2016**, *38*. [[CrossRef](#)]
26. Xing, Y.; Shu, C.W. High order well-balanced finite volume WENO schemes and discontinuous Galerkin methods for a class of hyperbolic systems with source terms. *J. Comput. Phys.* **2006**, *214*, 567–598. [[CrossRef](#)]
27. Klingenberg, C.; Puppo, G.; Semplice, M. Arbitrary Order Finite Volume Well-Balanced Schemes for the Euler Equations with Gravity. *SIAM J. Sci. Comput.* **2019**, *41*, A695–A721. [[CrossRef](#)]
28. Kanbar, F.; Touma, R.; Klingenberg, C. Well-balanced Central Schemes for the One and Two-dimensional Euler Systems with Gravity. *Appl. Numer. Math.* **2020**, *156*. [[CrossRef](#)]
29. Cheng, Y.; Kurganov, A. Moving-water equilibria preserving central-upwind schemes for the shallow water equations. *Commun. Math. Sci.* **2016**, *14*, 1643–1663. [[CrossRef](#)]
30. Kurganov, A. Finite-volume schemes for shallow-water equations. *Acta Numer.* **2018**, *27*, 289–351. [[CrossRef](#)]
31. Chertock, A.; Cui, S.; Kurganov, A.; Wu, T. Well-balanced positivity preserving central-upwind scheme for the shallow water system with friction terms. *Int. J. Numer. Methods Fluids* **2015**, *78*. [[CrossRef](#)]
32. Castro, M.J.; de Luna, T.M.; Parés, C. Well-balanced schemes and path-conservative numerical methods. In *Handbook of Numerical Analysis*; Elsevier: Amsterdam, The Netherlands, 2017; Volume 18, pp. 131–175.
33. Franck, E.; Mendoza, L.S. Finite volume scheme with local high order discretization of the hydrostatic equilibrium for the euler equations with external forces. *J. Sci. Comput.* **2016**, *69*, 314–354. [[CrossRef](#)]
34. Grosheintz-Laval, L.; Käppeli, R. Well-balanced finite volume schemes for nearly steady adiabatic flows. *J. Comput. Phys.* **2020**, *423*, 109805. [[CrossRef](#)]
35. Käppeli, R.; Mishra, S. A well-balanced finite volume scheme for the euler equations with gravitation-the exact preservation of hydrostatic equilibrium with arbitrary entropy stratification. *Astron. Astrophys.* **2016**, *587*, A94. [[CrossRef](#)]
36. Castro, M.; Gallardo, J.; López-García, J.; Parés, C. Well-Balanced High Order Extensions of Godunov’s Method for Semilinear Balance Laws. *SIAM J. Numer. Anal.* **2008**, *46*, 1012–1039. [[CrossRef](#)]
37. Castro, M.; Parés, C. Well-Balanced High-Order Finite Volume Methods for Systems of Balance Laws. *J. Sci. Comput.* **2020**, *82*. [[CrossRef](#)]
38. Castro, M.; López-García, J.; Parés, C. High order exactly well-balanced numerical methods for shallow water systems. *J. Comput. Phys.* **2013**, *246*, 242–264. [[CrossRef](#)]
39. Gómez-Bueno, I.; Castro, M.J.; Parés, C. High-order well-balanced methods for systems of balance laws: A control-based approach. *Appl. Math. Comput.* **2021**, *394*, 125820.
40. Gómez-Bueno, I.; Castro, M.; Parés, C. Well-Balanced Reconstruction Operator for Systems of Balance Laws: Numerical Implementation. In *Recent Advances in Numerical Methods for Hyperbolic PDE Systems*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 57–77.
41. Bermudez, A.; Vazquez, M.E. Upwind methods for hyperbolic conservation laws with source terms. *Comput. Fluids* **1994**, *23*, 1049–1071. [[CrossRef](#)]
42. Haier, E.; Lubich, C.; Wanner, G. *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*; Springer: Berlin/Heidelberg, Germany, 2006.
43. van Leer, B. Towards the Ultimate Conservative Difference Scheme V. A Second-order Sequel to Godunov’s Method. *J. Comput. Phys.* **1979**, *32*, 101–136. [[CrossRef](#)]
44. Levy, D.; Puppo, G.; Russo, G. Compact Central WENO Schemes for Multidimensional Conservation Laws. *SIAM J. Sci. Comput.* **2000**, *22*. [[CrossRef](#)]
45. Cravero, I.; Semplice, M. On the accuracy of WENO and CWENO reconstructions of third order on nonuniform meshes. *J. Sci. Comput.* **2015**, *67*. [[CrossRef](#)]
46. Gottlieb, S.; Shu, C.W. Total Variation Diminishing Runge-Kutta Schemes. *Math. Comput.* **1996**, *67*. [[CrossRef](#)]
47. Michel-Dansac, V.; Berthon, C.; Clain, S.; Foucher, F. A well-balanced scheme for the shallow-water equations with topography. *Comput. Math. Appl.* **2015**, *72*. [[CrossRef](#)]
48. Gaburro, E.; Castro, M.J.; Dumbser, M. Well-balanced Arbitrary-Lagrangian-Eulerian finite volume schemes on moving nonconforming meshes for the Euler equations of gas dynamics with gravity. *Mon. Not. R. Astron. Soc.* **2018**, *477*, 2251–2275. [[CrossRef](#)]