



**HAL**  
open science

## A well-balanced scheme for the shallow-water equations with topography or Manning friction

Victor Michel-Dansac, Christophe Berthon, Stéphane Clain, Françoise Foucher

### ► To cite this version:

Victor Michel-Dansac, Christophe Berthon, Stéphane Clain, Françoise Foucher. A well-balanced scheme for the shallow-water equations with topography or Manning friction. *Journal of Computational Physics*, 2017, 335, pp.115-154. 10.1016/j.jcp.2017.01.009 . hal-01247813v2

**HAL Id: hal-01247813**

**<https://hal.science/hal-01247813v2>**

Submitted on 28 Mar 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A well-balanced scheme for the shallow-water equations with topography or Manning friction.

Victor Michel-Dansac<sup>a,\*</sup>, Christophe Berthon<sup>a</sup>, Stéphane Clain<sup>b</sup>, Françoise Foucher<sup>a,c</sup>

<sup>a</sup>*Laboratoire de Mathématiques Jean Leray, CNRS UMR 6629, Université de Nantes, 2 rue de la Houssinière, BP 92208, 44322 Nantes Cedex 3, France*

<sup>b</sup>*Centre of Mathematics, Minho University, Campus de Gualtar - 4710-057 Braga, Portugal*

<sup>c</sup>*École Centrale de Nantes, 1 rue de La Noë, BP 92101 44321 Nantes Cedex 3, France*

---

## Abstract

We consider the shallow-water equations with Manning friction or topography, as well as a combination of both these source terms. The main purpose of this work concerns the derivation of a non-negativity preserving and well-balanced scheme that approximates solutions of the system and preserves the associated steady states, including the moving ones. In addition, the scheme has to deal with vanishing water heights and transitions between wet and dry areas. To address such issues, a particular attention is paid to the study of the steady states related to the friction source term. Then, a Godunov-type scheme is obtained by using a relevant average of the source terms in order to enforce the required well-balance property. An implicit treatment of both topography and friction source terms is also exhibited to improve the scheme while dealing with vanishing water heights. A second-order well-balanced MUSCL extension is designed, as well as an extension for the two-dimensional case. Numerical experiments are performed in order to highlight the properties of the scheme.

*Keywords:* shallow-water equations, Manning friction, Godunov-type schemes, well-balanced schemes, moving steady states

*2000 MSC:* 65M08, 65M12

---

## 1. Introduction

The goal of this paper is to derive a numerical scheme to approximate the solutions of the shallow-water equations with topography and Manning friction. The Manning friction was introduced in [39] (see also [19] for an overview of

---

\*Corresponding author

*Email addresses:* `victor.michel-dansac@univ-nantes.fr` (Victor Michel-Dansac), `christophe.berthon@univ-nantes.fr` (Christophe Berthon), `clain@math.uminho.pt` (Stéphane Clain), `francoise.foucher@ec-nantes.fr` (Françoise Foucher)

other friction models, and [9, 10, 13] for related works). The equations of interest consist in the following system:

$$\begin{cases} \partial_t h + \partial_x q & = 0, \\ \partial_t q + \partial_x \left( \frac{q^2}{h} + \frac{1}{2} g h^2 \right) & = -gh \partial_x Z - \frac{kq|q|}{h^\eta}. \end{cases} \quad (1.1)$$

This system is used to model the flow of water in a one-dimensional channel, with a non-flat bottom that applies a friction force on the water. The variables involved in this model are the non-negative water height  $h(x, t)$  and the depth-averaged discharge of the water  $q(x, t)$ . The known quantities are the shape of the channel bottom  $Z(x)$ , the gravity constant  $g$ , the Manning friction coefficient  $k$  and a parameter  $\eta$ , equal to  $\frac{2}{3}$ . We also define the velocity  $u$  of the water, such that  $q = hu$ .

We define the admissible states space by

$$\Omega = \{W = {}^t(h, q) \in \mathbb{R}^2 ; h \geq 0, q \in \mathbb{R}\}.$$

Let us note that the water height may vanish, which accounts for dry areas. By convention, we impose  $u = 0$  as soon as  $h = 0$ .

In order to shorten the notations, the system (1.1) is rewritten under a simpler form, as follows:

$$\partial_t W + \partial_x f(W) = s(W), \quad W \in \Omega, \quad (1.2)$$

where

$$W = \begin{pmatrix} h \\ q \end{pmatrix}, \quad f(W) = \begin{pmatrix} q \\ \frac{q^2}{h} + \frac{1}{2} g h^2 \end{pmatrix}, \quad s(W) = \begin{pmatrix} 0 \\ -gh \partial_x Z - \frac{kq|q|}{h^\eta} \end{pmatrix}. \quad (1.3)$$

Omitting the source terms, the homogeneous system deriving from (1.1) is a hyperbolic system. Its characteristic velocities are given by  $u - c$  and  $u + c$  (see [10, 25, 34] for instance), where  $c$  is the sound speed, defined as follows:

$$c = \sqrt{gh}. \quad (1.4)$$

The main focus of this paper lies in the study of the solutions of (1.2) that satisfy  $\partial_t W = 0$ . Such solutions are called *steady state solutions*. With vanishing partial derivatives in time in (1.1), we get:

$$\begin{cases} \partial_x q & = 0, \\ \partial_x \left( \frac{q^2}{h} + \frac{1}{2} g h^2 \right) & = -gh \partial_x Z - \frac{kq|q|}{h^\eta}. \end{cases} \quad (1.5)$$

We immediately obtain that, as per the first equation, the discharge  $q$  must be uniform. From now on, we denote this uniform discharge as follows:

$$q = q_0. \quad (1.6)$$

Using  $q_0$  within the second equation of (1.5) yields:

$$\partial_x \left( \frac{q_0^2}{h} + \frac{1}{2}gh^2 \right) = -gh\partial_x Z - \frac{kq_0|q_0|}{h^\eta}. \quad (1.7)$$

The solution of the nonlinear ordinary differential equation (1.7) is the water height  $h$  for a steady state with uniform discharge  $q_0$ . For a general topography and a nonzero friction contribution, if  $q_0 \neq 0$ , we cannot solve this equation to obtain an analytical expression for  $h$ . Thus, we will focus on specific cases; namely, the steady states at rest, and the steady states with topography or friction only.

The most important and most extensively studied steady state of the shallow-water equations is the *lake at rest* steady state (see for instance the non-exhaustive list [2, 8, 10, 14, 23, 29]). The lake at rest is obtained by assuming that the water is at rest, i.e.  $q_0 = 0$ . Thus, this steady state is defined as follows:

$$\begin{cases} q &= 0, \\ h + Z &= \text{cst}. \end{cases} \quad (1.8)$$

The importance of this steady state has led to the derivation of many schemes ensuring that the lake at rest is preserved. For instance, in [3, 29], the authors introduced the notion of a *well-balanced scheme*, that is to say a scheme that preserves some or all the steady states. Next, the well-balance approach was extended to nonlinear systems in [26]. The techniques from [26] were then simplified in [1], to obtain the hydrostatic reconstruction (see also [8, 21, 22, 32, 36, 37, 41]).

Now, let us consider a vanishing friction source term (i.e.  $k = 0$ ) without being at rest (i.e.  $q_0 \neq 0$ ). From (1.7), we get the *moving steady states*, governed by

$$\partial_x \left( \frac{q_0^2}{h} + \frac{1}{2}gh^2 \right) = -gh\partial_x Z. \quad (1.9)$$

The equation (1.9) that defines the moving steady states is known to be harder to satisfy at the discrete level than the lake at rest. Indeed, the lake at rest explicitly yields  $h$  with respect to the known quantity  $Z$ , while the water height is given by the nonlinear ODE (1.9) in the case of the moving steady states. Such steady states have been studied in the past, for instance in [15, 40], where the authors exhibit several properties of the topography steady state solutions. Such steady states have also been studied to provide several properties of numerical schemes, for instance in [28], where the authors define some specific moving steady states and use them to test the accuracy of non-well-balanced schemes. Then, in [26], the author introduced the notion of fully well-balanced schemes, able to preserve every steady state, and designed such a scheme. This approach was extended in [4], where the authors derive a fully well-balanced and entropy-stable Godunov-type scheme, based on the resolution of a Bernoulli-type equation. Since this equation is difficult to solve, the authors of [40] have proposed a fully well-balanced Godunov-type scheme that uses a linearization of

the Bernoulli-type equation. We mention two additional fully well-balanced numerical schemes based on different techniques: in [11], the authors generalize the hydrostatic reconstruction mentioned above, while the authors of [12] propose a high-order approach using a relevant reconstruction. High-order techniques were also developed to ensure a very accurate approximation of the moving steady states (see for instance [43, 44, 49, 50]).

Let us now assume a flat topography and a nonzero discharge. The steady states governed by the friction source term are given as follows:

$$\partial_x \left( \frac{q_0^2}{h} + \frac{1}{2}gh^2 \right) = -\frac{kq_0|q_0|}{h^n}. \quad (1.10)$$

Some research has been devoted to Manning friction terms (for instance, see [19, 9]), but these works do not focus on steady states and the resulting schemes are not well-balanced.

In the present paper, we exhibit smooth steady solutions of (1.10) and propose a well-balanced scheme, to exactly preserve the lake at rest (1.8), the moving steady states (1.9), and the dominant friction steady states (1.10). This scheme will be derived in order to satisfy some essential properties, namely the exact preservation of all the steady states for the system (1.1) with topography or friction and the preservation of the non-negativity of the water height. In addition, the scheme will be designed to model transitions between wet areas, where  $h \neq 0$  and dry areas, where  $h = 0$ .

The paper is organized as follows. First, in Section 2, we study smooth steady states for the shallow-water equations with friction over a flat bottom. Then, in Section 3, we develop a Godunov-type scheme that is well-balanced for the system (1.1). In this section, we first recall some generalities on Godunov-type schemes, then we derive the well-balanced scheme, according to the friction steady states. Afterwards, in Section 4, we extend this scheme to be well-balanced for both the topography and the friction source terms. In addition, in this section, we prove the non-negativity preservation. Then, in Section 5, we propose an implicit treatment of the source terms, designed to deal with transitions between wet and dry areas. The scheme is then extended in Section 6 to be second-order accurate. This extension breaks the well-balance property, but we show a convex combination technique that allows the recovery of this essential property. Finally, in Section 7, we propose three kinds numerical experiments, namely well-balance assessments, dam-break simulations, and two dimensional simulations. The paper is concluded in Section 8.

## 2. Steady states with friction and a flat bottom

This section is devoted to studying the steady states coming from the friction source term, given by (1.5) with a flat topography ( $Z = \text{cst}$ ), or equivalently by (1.10). Recall that, for steady states and after (1.6), the discharge  $q = q_0$  is uniform throughout the domain. Here, we assume  $q_0 \neq 0$ . Indeed, we assume

that the friction contribution vanishes as soon as the discharge is equal to zero. We rewrite (1.10) under the form:

$$q_0^2 \partial_x \frac{1}{h} + \frac{g}{2} \partial_x h^2 = -\frac{kq_0|q_0|}{h^\eta}. \quad (2.1)$$

We are looking for smooth solutions  $h(x) > 0$  of the equation (2.1). As a first step, we integrate this equation. Multiplying by  $h^\eta$  provides:

$$-\frac{q_0^2}{\eta-1} \partial_x h^{\eta-1} + \frac{g}{\eta+2} \partial_x h^{\eta+2} = -kq_0|q_0|. \quad (2.2)$$

Let us consider  $x_0 \in \mathbb{R}$  as an arbitrary reference point, and introduce the initial condition  $h(x_0) = h_0$ . For all  $x \in \mathbb{R}$ , we integrate (2.2) over  $(x_0, x)$ , to get

$$-\frac{q_0^2}{\eta-1} (h^{\eta-1} - h_0^{\eta-1}) + \frac{g}{\eta+2} (h^{\eta+2} - h_0^{\eta+2}) + kq_0|q_0|(x - x_0) = 0. \quad (2.3)$$

To shorten the notations, we rewrite (2.3) under the following form:

$$\xi(h; x, h_0, q_0, x_0) = 0, \quad (2.4)$$

where we have set

$$\begin{aligned} \xi(h; x, h_0, q_0, x_0) = & -\frac{q_0^2}{\eta-1} (h^{\eta-1} - h_0^{\eta-1}) + \frac{g}{\eta+2} (h^{\eta+2} - h_0^{\eta+2}) \\ & + kq_0|q_0|(x - x_0). \end{aligned} \quad (2.5)$$

We consider  $x$  a parameter so that (2.4) is an equation to characterize  $h := h(x)$ . In order to exhibit such solutions, we first study  $\xi$  with a fixed  $x \in \mathbb{R}$ . The derivative of  $\xi$  with respect to  $h$  is given by

$$\frac{\partial \xi}{\partial h}(h; x, h_0, q_0, x_0) = gh^{\eta-2} \left( -\frac{q_0^2}{g} + h^3 \right).$$

Let us define a critical water height  $h_c$ , such that  $\frac{\partial \xi}{\partial h}(h_c; x, h_0, q_0, x_0) = 0$ , as follows:

$$h_c = \left( \frac{q_0^2}{g} \right)^{1/3}. \quad (2.6)$$

We easily obtain that the function  $h \mapsto \xi(h; x, h_0, q_0, x_0)$  is strictly increasing on  $(h_c, +\infty)$ , while it is strictly decreasing on  $(0, h_c)$ . As a consequence, it admits a unique minimum in  $(0, +\infty)$ , and this minimum is reached for  $h = h_c$ .

Now, we evaluate the sign of  $\xi$ . After straightforward computations, the following limits are proven to be satisfied by  $\xi$ :

- $\lim_{h \rightarrow +\infty} \xi(h; x, h_0, q_0, x_0) = +\infty,$

- $\lim_{h \rightarrow 0^+} \xi(h; x, h_0, q_0, x_0) = \xi_\ell(x)$ ,

where we have set

$$\xi_\ell(x) = gh_0^{\eta-1} \left( \frac{h_c^3}{\eta-1} - \frac{h_0^3}{\eta+2} \right) + kq_0|q_0|(x-x_0).$$

Note that  $\xi_\ell$  is finite. Moreover, let us introduce

$$\xi_c(x) := \xi(h_c; x, h_0, q_0, x_0).$$

We note that

$$\xi_c(x) = \xi_\ell(x) - \frac{3gh_c^{\eta+2}}{(\eta-1)(\eta+2)} < \xi_\ell(x).$$

Equipped with these new notations and these properties of  $\xi$ , we state the following result.

**Lemma 1.** *Assume  $h > 0$  and  $q_0 \neq 0$ . Thus,  $h_c > 0$  according to (2.6).*

- (i) *If  $\xi_\ell(x) < 0$ , then the equation (2.4) admits a unique solution  $h^{sub}(x)$ . Moreover, this solution belongs to  $(h_c, +\infty)$ .*
- (ii) *If  $\xi_\ell(x) > 0$  and  $\xi_c(x) > 0$ , then there is no solution to the equation (2.4).*
- (iii) *If  $\xi_\ell(x) > 0$  and  $\xi_c(x) < 0$ , then the equation (2.4) admits two solutions,  $h^{sup}(x)$  and  $h^{sub}(x)$ . One solution,  $h^{sup}(x)$ , stays within  $(0, h_c)$ , while the second one,  $h^{sub}(x)$ , is in  $(h_c, +\infty)$ . If  $\xi_c(x) = 0$ , the equation (2.4) admits  $h_c$  as a unique double solution.*

*Proof.* To establish (i), we use that  $h \mapsto \xi(h; x, h_0, q_0, x_0)$  is strictly increasing on  $(h_c, +\infty)$  and that  $\xi_c(x) < \xi_\ell(x) < 0$ , to deduce that the equation (2.4) admits a unique solution, which belongs to  $(h_c, +\infty)$ .

Next, we turn to proving (ii). If  $\xi_c(x) > 0$ , since  $\xi$  reaches its unique minimum  $\xi_c(x)$  for  $h = h_c$ , it is immediate that there is no solution to the equation (2.4).

Finally, concerning (iii), we assume that  $\xi_\ell(x) > 0$  and  $\xi_c(x) < 0$ . The monotonicity properties of  $\xi$  allow the immediate conclusion that the equation (2.4) admits two solutions. One solution stays within  $(0, h_c)$  and corresponds to a supercritical state, while the second one is in  $(h_c, +\infty)$  and provides a subcritical state. Moreover, if  $\xi_c(x) = 0$ , the equation (2.4) admits a double root,  $h = h_c$ .

The proof is thus achieved.  $\square$

*Remark 1.* Arguing the definition of the Froude number  $Fr = u/c$  (where  $c$  is the sound speed, defined by (1.4)), we obtain for steady states  $Fr = q_0/\sqrt{gh^3}$ . Therefore, it is clear that  $h > h_c$  (resp.  $h < h_c$ ) corresponds to  $Fr < 1$  (resp.  $Fr > 1$ ), i.e. to a subcritical (resp. supercritical) flow. Thus, the solution  $h$  of (2.4) that lies in  $(0, h_c)$  will henceforth be called the *supercritical branch*, while the solution of (2.4) that lies in  $(h_c, +\infty)$  will be called the *subcritical branch*.

*Remark 2.* Note that the case  $\xi_\ell(x) = 0$  implies that  $h = 0$  is a solution of the equation (2.4). However, this equation has been obtained using the positivity of  $h$ : therefore,  $h = 0$  is not a valid solution of (2.4).

Now, we study the solutions  $h$  of (2.4) as functions of  $x$ . As a first step, we evaluate the monotonicity of  $h$ . To this end, we exhibit the derivative of  $h$  with respect to  $x$  by considering (2.2), which writes

$$(-q_0^2 h^{\eta-2} + g h^{\eta+1}) h'(x) = -k q_0 |q_0|.$$

Since  $q_0^2 = g h_c^3$  from (2.6), we immediately obtain

$$g h^{\eta+1} \left( 1 - \frac{h_c^3}{h^3} \right) h'(x) = -k q_0 |q_0|.$$

As a consequence, the sign of  $h'(x)$  coincides with the sign of  $q_0(h_c - h(x))$ . Note that both  $\xi_\ell(x)$  and  $\xi_c(x)$  are linear in  $x$ , and  $\xi_c(x) < \xi_\ell(x)$ . Both of these quantities are thus linear functions of  $x$ , of slope  $k q_0 |q_0|$ , and are strictly monotonic in  $x$ . Therefore, there exists a unique  $x_u$  such that  $\xi_\ell(x) = 0$ , and a unique  $x_c \neq x_u$  such that  $\xi_c(x) = 0$ . Thus, by Lemma 1, if there exists a solution  $h(x)$  of the equation (2.4), then it is either within  $(0, h_c]$  or within  $[h_c, +\infty)$ , for all  $x$  where such a solution exists. Therefore, the sign of  $h'(x)$  is constant, and  $h$  is strictly monotonic. As a consequence,  $h^{sub}$  and  $h^{sup}$  are bijective on their respective domains, since they are continuous and strictly monotonic.

These solutions are displayed on Figure 1. For the sake of simplicity, this figure has been obtained assuming that  $h_0 = h_c$  (and thus  $x_0 = x_c$ ) and that  $q_0 < 0$ .

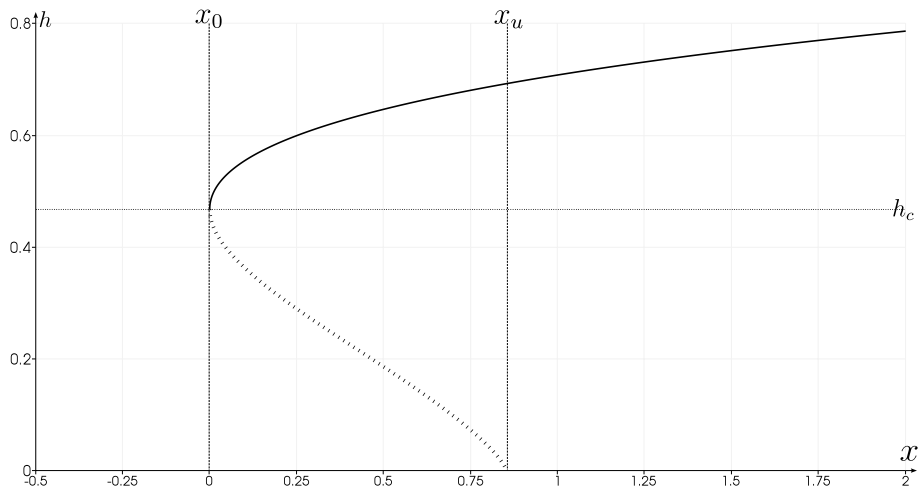


Figure 1: Solutions of (2.4) with  $q_0 = -1$  and  $x_0 = x_c = 0$ , obtained by using Newton's method. The solid line is the *subcritical branch* (the increasing solution) and the dotted line is the *supercritical branch* (the decreasing solution).



As a consequence, we have obtained the general form of steady state solutions for the shallow-water equations (1.1) with friction and flat topography, respectively given by the subcritical and supercritical branches. However, the subcritical and supercritical branches that are solution to (2.4) are parametrized by the choice of the initial conditions  $x_0$  and  $h_0$ .

We quickly discuss, in the context of Figure 1 (i.e.  $q_0 < 0$  and  $x_0 = x_c = 0$ ) without loss of generality, the possibility of extending the smooth steady solution for  $x < x_c$ . Assume that the critical point has been reached, i.e.  $x = x_c = 0$ ,  $q(x, t) = q_0$  and  $h(x, t) = h_c$ . Since  $q_0 < 0$ , we have two distinct steady solutions, the subcritical solution and the supercritical solution (where it exists), for  $x \geq x_c$ . However, if we were to compute a smooth steady solution for  $x < x_c$ , we would still need to have  $q(x, t) = q_0$ , as prescribed by the first steady relation. Lemma 1 states that such a solution to the second steady relation does not exist for  $x < x_c$ . Therefore, there is no way to extend the smooth steady solution for  $x < x_c$ .

### 3. A well-balanced scheme for a the shallow-water equations with friction

Equipped with the steady states, we now derive a numerical scheme able to exactly capture such essential solutions. We begin by recalling the construction of a Godunov-type scheme that uses a two-state approximate Riemann solver [30] (see also [10, 46]). The Godunov-type scheme is then adapted to the shallow-water equations with the friction source term. The purpose of this extension is to preserve the steady states presented in the previous section.

#### 3.1. Generalities on Godunov-type schemes

First, we introduce the discretization of the space and time domain  $\mathbb{R} \times \mathbb{R}^+$ . Let  $\Delta x$  be the space step, assumed to be constant, and  $\Delta t$  the time step. The space discretization consists in cells  $(x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}})$ , for all  $i \in \mathbb{Z}$ . The solution  $W(x, t)$  of (1.2) is approximated by  $W_i^n$  at some time  $t^n$  and in cell  $(x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}})$ . We suppose that this piecewise constant approximation of  $W(x, t)$ , here denoted  $W^\Delta(x, t^n)$ , is known at time  $t^n$  for all cells in the space domain. We now evolve  $W^\Delta(x, t^n)$  in time, which consists in nothing but the solution of the following Riemann problem juxtaposition, defined for all  $i \in \mathbb{Z}$  and all  $x$  between two consecutive cells:

$$\begin{cases} \partial_t W + \partial_x f(W) = s(W), \\ W(x, t^n) = \begin{cases} W_i^n & \text{if } x < x_{i+\frac{1}{2}}, \\ W_{i+1}^n & \text{if } x > x_{i+\frac{1}{2}}. \end{cases} \end{cases} \quad (3.1)$$

From now on, let us emphasize that the consecutive Riemann solutions do not interact as long as  $t$  is small enough. Since an exact solution to (3.1) is difficult to determine, we suggest to consider an *approximate Riemann solver* to

obtain an estimate of the unknown solution. We adopt a two-state approximate Riemann solver, illustrated by [Figure 2](#) and defined as follows:

$$\widetilde{W}\left(\frac{x}{t}; W_L, W_R\right) = \begin{cases} W_L & \text{if } x/t < \lambda_L, \\ W_L^* & \text{if } \lambda_L < x/t < 0, \\ W_R^* & \text{if } 0 < x/t < \lambda_R, \\ W_R & \text{if } x/t > \lambda_R, \end{cases} \quad (3.2)$$

where  $\lambda_L$  and  $\lambda_R$  denote some characteristic wave velocities, and  $W_L^*$  and  $W_R^*$  are the intermediate states, to be detailed later. To ensure that  $\lambda_L < 0$  and  $\lambda_R > 0$ , we choose the following expressions of  $\lambda_L$  and  $\lambda_R$  (see for instance [\[45\]](#) and references therein):

$$\begin{aligned} \lambda_L &= \min(-|u_L| - c_L, -|u_R| - c_R, -\varepsilon_\lambda), \\ \lambda_R &= \max(|u_L| + c_L, |u_R| + c_R, \varepsilon_\lambda), \end{aligned} \quad (3.3)$$

with  $c$  the sound speed, defined by [\(1.4\)](#), and  $\varepsilon_\lambda$  a positive real value to be fixed in the numerical applications.

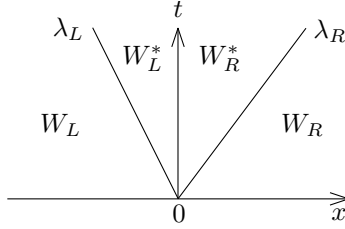


Figure 2: Structure of the chosen approximate Riemann solver.

The approximate solution of the Riemann problem  $W^\Delta(x, t^n + t)$ , depicted on [Figure 3](#), is given for  $t > 0$  and  $x \in (x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}})$  by:

$$\forall i \in \mathbb{Z}, W^\Delta(x, t^n + t) = \begin{cases} W_{i-\frac{1}{2}}^{R,*} & \text{if } x \in [x_{i-\frac{1}{2}}, x_{i-\frac{1}{2}} + \lambda_{i-\frac{1}{2}}^R t], \\ W_i^n & \text{if } x \in [x_{i-\frac{1}{2}} + \lambda_{i-\frac{1}{2}}^R t, x_{i+\frac{1}{2}} + \lambda_{i+\frac{1}{2}}^L t], \\ W_{i+\frac{1}{2}}^{L,*} & \text{if } x \in [x_{i+\frac{1}{2}} + \lambda_{i+\frac{1}{2}}^L t, x_{i+\frac{1}{2}}]. \end{cases} \quad (3.4)$$

Now, we consider a time step  $\Delta t$  satisfying the following Courant-Friedrichs-Lewy stability condition :

$$\Delta t \leq \frac{\Delta x}{2\Lambda}, \text{ where } \Lambda = \max_{i \in \mathbb{Z}} \left( -\lambda_{i+\frac{1}{2}}^L, \lambda_{i+\frac{1}{2}}^R \right), \quad (3.5)$$

so as to ensure there are no interactions between the waves from two consecutive Riemann problems (see [\[25, 45\]](#)). Given such a  $\Delta t$ , we set  $t^{n+1} = t^n + \Delta t$  and

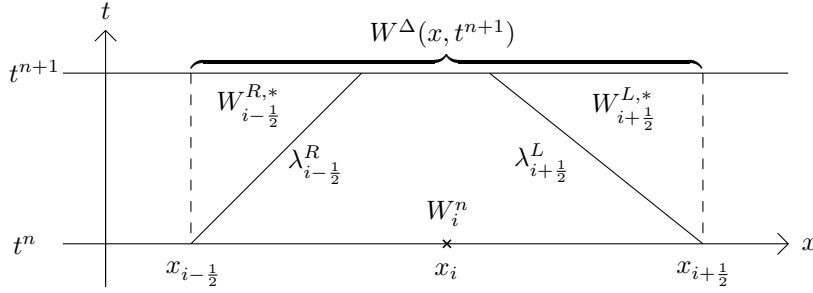


Figure 3: The full Godunov-type scheme using an approximate Riemann solver.

focus on the value of  $W^\Delta$  at time  $t^{n+1}$ . At last, we define  $W_i^{n+1}$  by taking the average of  $W^\Delta(x, t^{n+1})$  over the cell  $(x_{i-1/2}, x_{i+1/2})$ , as follows:

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

Since  $W^\Delta(x, t^{n+1})$  is given by (3.4), we have:

$$\begin{aligned} W_i^{n+1} &= \frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i-1/2} + \lambda_{i-1/2}^R \Delta t} W_{i-1/2}^{R,*} dx + \frac{1}{\Delta x} \int_{x_{i+1/2} + \lambda_{i+1/2}^L \Delta t}^{x_{i+1/2}} W_{i+1/2}^{L,*} dx \\ &\quad + \frac{1}{\Delta x} \int_{x_{i-1/2} + \lambda_{i-1/2}^R \Delta t}^{x_{i+1/2} + \lambda_{i+1/2}^L \Delta t} W_i^n dx. \end{aligned}$$

This relation immediately yields:

$$W_i^{n+1} = W_i^n - \frac{\Delta t}{\Delta x} \left[ \lambda_{i+1/2}^L \left( W_{i+1/2}^{L,*} - W_i^n \right) - \lambda_{i-1/2}^R \left( W_{i-1/2}^{R,*} - W_i^n \right) \right]. \quad (3.6)$$

Note that, in order for  $W_i^{n+1}$  to be fully defined, we only need to give explicit values to the intermediate states  $W_{i+1/2}^{L,*}$  and  $W_{i-1/2}^{R,*}$ . The remainder of this section is devoted to finding suitable intermediate states, which ensure that the scheme is consistent and preserves the steady states.

### 3.2. A well-balanced approximate Riemann solver for the friction source term

We focus on the derivation of a well-balanced approximate Riemann solver for the shallow-water equations with only the friction source term. First, we recall the system under consideration:

$$\begin{cases} \partial_t h + \partial_x q &= 0, \\ \partial_t q + \partial_x \left( \frac{q^2}{h} + \frac{1}{2} g h^2 \right) &= -kq|q|h^{-\eta}. \end{cases} \quad (3.7)$$

We also recall that the steady states for this model are given by:

$$\begin{cases} q = q_0, \\ \partial_x \left( \frac{q_0^2}{h} + \frac{1}{2}gh^2 \right) = -kq|q|h^{-\eta}. \end{cases} \quad (3.8)$$

Note that the friction source term  $S^f = -kq|q|h^{-\eta}$  in the system (3.7) falls under the more generic framework  $S^f = h^\beta \rho(q) \partial_x Y$ , where  $\beta = -\eta$ ,  $\rho(q) = q|q|$  and  $Y = -kx$ . The computations could be performed in this framework, but they would lead to having to solve a nonlinear equation. However, thanks to the specific form of the friction source term, the scheme can be obtained without having to solve this nonlinear equation. As a consequence, we elect to derive the scheme for the specific case of the friction source term instead of a more generic form. These remarks also apply to the topography source term  $S^t = -gh \partial_x Z$ , which is treated in a later section.

Relevant definitions of the intermediate states  $W_L^* = {}^t(h_L^*, q_L^*)$  and  $W_R^* = {}^t(h_R^*, q_R^*)$  need to be obtained in order to fully determine the approximate Riemann solver. Therefore, four equations characterizing the four unknowns  $h_L^*$ ,  $h_R^*$ ,  $q_L^*$  and  $q_R^*$  are required. These four equations will be obtained by considering two essential properties that need to be satisfied by the approximate Riemann solver: consistency with the system (3.7) and well-balance, i.e. preservation of the steady states (3.8).

First, we tackle the issue of the consistency with (3.7), by introducing a necessary consistency condition. From [30], the average over a cell of the approximate Riemann solver  $\widetilde{W}$ , defined by (3.2), has to be equal to the average over the same cell of the exact solution of the Riemann problem  $W_{\mathcal{R}}$ . Therefore, the following equality is imposed:

$$\frac{1}{\Delta x} \int_{-\Delta x/2}^{\Delta x/2} \widetilde{W} \left( \frac{x}{\Delta t}; W_L, W_R \right) dx = \frac{1}{\Delta x} \int_{-\Delta x/2}^{\Delta x/2} W_{\mathcal{R}} \left( \frac{x}{\Delta t}; W_L, W_R \right) dx. \quad (3.9)$$

We introduce the following notations for the sake of simplicity:

$$\begin{aligned} (\lambda_R - \lambda_L)h_{HLL} &= \lambda_R h_R - \lambda_L h_L - [q], \\ (\lambda_R - \lambda_L)q_{HLL} &= \lambda_R q_R - \lambda_L q_L - \left[ \frac{q^2}{h} + \frac{1}{2}gh^2 \right]. \end{aligned}$$

Let us underline that  ${}^t(h_{HLL}, q_{HLL})$  is actually the well-known intermediate state of the HLL approximate Riemann solver introduced in [30]. Also, note that  $h_{HLL} > 0$  for  $\lambda_L$  and  $\lambda_R$  given by (3.3). We also introduce the notation  $(-kq|q|h^{-\eta})_{\mathcal{R}}$  that denotes the second component of  $s(W_{\mathcal{R}})$ , with  $s$  defined by (1.3), to represent the value of the source term of friction for the exact solution of the Riemann problem. Thanks to these notations, (3.9) can be rewritten as

follows, after straightforward computations detailed in [40]:

$$\lambda_R h_R^* - \lambda_L h_L^* = (\lambda_R - \lambda_L) h_{HLL}, \quad (3.10a)$$

$$\begin{aligned} \lambda_R q_R^* - \lambda_L q_L^* &= (\lambda_R - \lambda_L) q_{HLL} \\ &+ \frac{1}{\Delta t} \int_{-\Delta x/2}^{\Delta x/2} \int_0^{\Delta t} (-kq|q|h^{-\eta})_{\mathcal{R}} \left( \frac{x}{t}; W_L, W_R \right) dt dx. \end{aligned} \quad (3.10b)$$

Now, we introduce two parameters,  $\bar{q}$  and  $\overline{h^{-\eta}}$ , to define a consistent approximation of the mean value of the friction source term, as follows:

$$\frac{1}{\Delta x} \frac{1}{\Delta t} \int_{-\Delta x/2}^{\Delta x/2} \int_0^{\Delta t} (-kq|q|h^{-\eta})_{\mathcal{R}} dt dx \simeq -k\bar{q}|\bar{q}|\overline{h^{-\eta}}. \quad (3.11)$$

The parameters  $\bar{q}$  and  $\overline{h^{-\eta}}$  are respectively assumed to be consistent with  $q$  and  $h^{-\eta}$ , in a sense given later. Then, the equation (3.11) is substituted into (3.10b). As a consequence, the four unknowns are imposed to satisfy the following two relations:

$$\begin{aligned} \lambda_R h_R^* - \lambda_L h_L^* &= (\lambda_R - \lambda_L) h_{HLL}, \\ \lambda_R q_R^* - \lambda_L q_L^* &= (\lambda_R - \lambda_L) q_{HLL} - k\bar{q}|\bar{q}|\overline{h^{-\eta}} \Delta x. \end{aligned} \quad (3.12)$$

Concerning the well-balance of the approximate Riemann solver, we exhibit a sufficient condition for the preservation of a solution. The scheme (3.6) provides us with such a condition. Indeed, the solution is obviously stationary, i.e.  $W_i^{n+1} = W_i^n$  for all  $i \in \mathbb{Z}$ , if we have

$$\forall i \in \mathbb{Z}, \quad W_{i+\frac{1}{2}}^{L,*} = W_i^n \quad \text{and} \quad W_{i-\frac{1}{2}}^{R,*} = W_i^n.$$

Thus, in the context of the approximate Riemann solver, the solution will be stationary if  $W_L^* = W_L$  and  $W_R^* = W_R$ . As a consequence, the following well-balance principle will have to be satisfied by the intermediate states:

**Principle (WB).** The intermediate states  $W_L^*$  and  $W_R^*$  are such that  $W_L^* = W_L$  and  $W_R^* = W_R$  as soon as  $W_L$  and  $W_R$  define a steady state.

Here, the pair  $(W_L, W_R)$  is said to define a steady state if they satisfy the identity (3.8). We will then look for intermediate states  $W_L^*$  and  $W_R^*$  that satisfy the (WB) principle.

We begin by determining  $q_L^*$  and  $q_R^*$ . After [4, 40], we choose  $q_L^* = q_R^*$  and we set  $q^* := q_L^* = q_R^*$ . This choice helps guarantee that  $W_L^*$  and  $W_R^*$  satisfy the property (WB). The formulas (3.12) are then rewritten using  $q^*$  as follows:

$$\lambda_R h_R^* - \lambda_L h_L^* = (\lambda_R - \lambda_L) h_{HLL}, \quad (3.13a)$$

$$q^* = q_{HLL} - \frac{k\bar{q}|\bar{q}|\overline{h^{-\eta}} \Delta x}{\lambda_R - \lambda_L}. \quad (3.13b)$$

We only need relevant definitions of  $\bar{q}$  and  $\overline{h^{-\eta}}$  in order to fully determine  $q^*$ . To that end, we specify the discrete steady states satisfied by the system (3.7). Such discrete steady states are nothing but a discretization of the steady relation at the continuous level (3.8). In order to give an expression of these discrete steady states, we assume that  $h_L \neq 0$  and  $h_R \neq 0$ . Now,  $W_L$  and  $W_R$  define a steady state if the following relations hold:

$$\begin{cases} q_L = q_R = q_0, \\ q_0^2 \left[ \frac{1}{h} \right] + \frac{g}{2} [h^2] = -k\bar{q}|\bar{q}|\overline{h^{-\eta}} \Delta x. \end{cases} \quad (3.14a)$$

$$(3.14b)$$

Moreover, since we are considering smooth steady states for just the friction source term, we can rewrite the relation (2.3) between states  $W_L$  and  $W_R$ . Thus,  $W_L$  and  $W_R$  satisfy the following algebraic relation:

$$-\frac{q_0^2}{\eta-1} [h^{\eta-1}] + \frac{g}{\eta+2} [h^{\eta+2}] = -kq_0|q_0| \Delta x. \quad (3.15)$$

We introduce the notation  $\mu_0 = \text{sgn}(q_0)$ . Therefore,  $\mu_0$  represents the direction of the water flow. From (3.15), we obtain the evaluation of  $q_0^2$  as follows:

$$q_0^2 = \frac{g \frac{[h^{\eta+2}]}{\eta+2}}{\frac{[h^{\eta-1}]}{\eta-1} - k\mu_0 \Delta x}. \quad (3.16)$$

Now, to recover the behavior of the friction source term when steady states are involved, we assume that the parameter  $\bar{q}$  is equal to  $q_0$  as soon as  $W_L$  and  $W_R$  define a steady state. Thus, injecting (3.16) into (3.14b) yields the expression of  $\overline{h^{-\eta}}$ . We immediately obtain:

$$\overline{h^{-\eta}} = \frac{[h^2]}{2} \frac{\eta+2}{[h^{\eta+2}]} - \frac{\mu_0}{k \Delta x} \left( \left[ \frac{1}{h} \right] + \frac{[h^2]}{2} \frac{[h^{\eta-1}]}{\eta-1} \frac{\eta+2}{[h^{\eta+2}]} \right). \quad (3.17)$$

**Lemma 2.** *The expression of  $\overline{h^{-\eta}}$  given by (3.17) is consistent with  $h^{-\eta}$ .*

*Proof.* With smooth water heights, we fix, in (3.17),  $h_L = h(x)$  and  $h_R = h(x + \mathcal{O}(\Delta x))$ . Taylor's formula applied to  $h_R$  yields  $h_R = h + \Delta x \partial_x h + \mathcal{O}(\Delta x^2)$ . In order to evaluate the Taylor expansions of  $[h^2]$ ,  $[h^{\eta-1}]$ ,  $[h^{\eta+2}]$  and  $[h^{-1}]$ , we now compute a Taylor expansion, for some  $\beta \in \mathbb{R}$ , of the jump  $[h^\beta]$ :

$$\begin{aligned} [h^\beta] &= h_R^\beta - h_L^\beta = (h + \partial_x h \Delta x + \mathcal{O}(\Delta x^2))^\beta - h^\beta \\ &= h^\beta (1 + \beta h^{-1} \partial_x h \Delta x + \mathcal{O}(\Delta x^2)) - h^\beta \\ &= \beta h^{\beta-1} \partial_x h \Delta x + \mathcal{O}(\Delta x^2). \end{aligned}$$

Using the above evaluation, we have for the first part of the expression of  $\overline{h^{-\eta}}$ :

$$\frac{[h^2]}{2} \frac{\eta+2}{[h^{\eta+2}]} = \frac{h \partial_x h \Delta x + \mathcal{O}(\Delta x^2)}{h^{\eta+1} \partial_x h \Delta x + \mathcal{O}(\Delta x^2)} = h^{-\eta} + \mathcal{O}(\Delta x). \quad (3.18)$$

Moreover, we have, for the second part of  $\overline{h^{-\eta}}$ ,

$$\left[\frac{1}{h}\right] = -h^{-2}\partial_x h \Delta x + \mathcal{O}(\Delta x^2), \quad (3.19)$$

$$\begin{aligned} \frac{[h^2]}{2} \frac{[h^{\eta-1}]}{\eta-1} \frac{\eta+2}{[h^{\eta+2}]} &= \frac{(h\partial_x h \Delta x + \mathcal{O}(\Delta x^2))(h^{\eta-2}\partial_x h \Delta x + \mathcal{O}(\Delta x^2))}{h^{\eta+1}\partial_x h \Delta x + \mathcal{O}(\Delta x^2)} \\ &= h^{-2}\partial_x h \Delta x + \mathcal{O}(\Delta x^2). \end{aligned} \quad (3.20)$$

Combining both equations (3.19) and (3.20) immediately yields

$$-\frac{\mu_0}{k \Delta x} \left( \left[\frac{1}{h}\right] + \frac{[h^2]}{2} \frac{[h^{\eta-1}]}{\eta-1} \frac{\eta+2}{[h^{\eta+2}]} \right) = -\frac{\mu_0}{k \Delta x} \mathcal{O}(\Delta x^2) = \mathcal{O}(\Delta x). \quad (3.21)$$

Using both relations (3.18) and (3.21) gives  $\overline{h^{-\eta}} = h^{-\eta} + \mathcal{O}(\Delta x)$ , which concludes the proof.  $\square$

Concerning  $\bar{q}$ , we choose the following average:

$$\begin{cases} \bar{q} = \frac{2|q_L||q_R|}{|q_L| + |q_R|} \operatorname{sgn}(q_L + q_R) & \text{if } q_L \neq 0 \text{ and } q_R \neq 0; \\ \bar{q} = 0 & \text{if } q_L = 0, q_R = 0 \text{ or } k = 0. \end{cases} \quad (3.22)$$

This average indeed ensures that, if  $q_L = q_R$ , then  $\bar{q} = q_L = q_R$ .

Note that the expression (3.17) of  $\overline{h^{-\eta}}$  contains  $\mu_0$ . This quantity depends on the steady state and has to be determined for non-steady states. To address such an issue, we propose the expression

$$\overline{h^{-\eta}} := \frac{[h^2]}{2} \frac{\eta+2}{[h^{\eta+2}]} - \frac{\bar{\mu}}{k \Delta x} \left( \left[\frac{1}{h}\right] + \frac{[h^2]}{2} \frac{[h^{\eta-1}]}{\eta-1} \frac{\eta+2}{[h^{\eta+2}]} \right), \quad (3.23)$$

where  $\bar{\mu}$  is the sign of the quantity  $\bar{q}$  given by (3.22). Since  $\bar{q} = q_0$  as soon as  $W_L$  and  $W_R$  define a steady state, it is clear that  $\overline{h^{-\eta}}$ , as defined by (3.23), still provides a well-balanced approximate Riemann solver and still satisfies Lemma 2.

*Remark 3.* With  $\overline{h^{-\eta}}$  given by (3.23), the average source term  $-k\bar{q}|\bar{q}|\overline{h^{-\eta}}$ , which satisfies (3.11), rewrites as follows:

$$-k\bar{q}|\bar{q}|\overline{h^{-\eta}} = -k\bar{q}|\bar{q}|\frac{[h^2]}{2} \frac{\eta+2}{[h^{\eta+2}]} + \frac{(\bar{q})^2}{\Delta x} \left( \left[\frac{1}{h}\right] + \frac{[h^2]}{2} \frac{[h^{\eta-1}]}{\eta-1} \frac{\eta+2}{[h^{\eta+2}]} \right). \quad (3.24)$$

As a consequence, when the Manning coefficient  $k$  is almost zero, the second term of this average source term is dominant. From the proof of Lemma 2, we know that this second term is consistent with 0 for a smooth solution. However, this term could become inconsistent within shock waves. To address such an issue, we introduce a cutoff of this inconsistent term. Such a procedure had

already been adopted in [4, 5, 40] in the context of the sole topography source term. Therefore, we modify the expression (3.23) of  $\overline{h^{-\eta}}$ , as follows:

$$\overline{h^{-\eta}} := \overline{h^{-\eta}}(h_L, h_R) = \frac{[h^2]}{2} \frac{\eta + 2}{[h^{\eta+2}]} - \frac{\bar{\mu}}{k \Delta x} [h]_c \left( \frac{-1}{h_L h_R} + \frac{h_L + h_R}{2} \frac{[h^{\eta-1}]}{\eta - 1} \frac{\eta + 2}{[h^{\eta+2}]} \right), \quad (3.25)$$

where  $[h]_c$  is a cutoff of  $[h] = h_R - h_L$ , defined as follows:

$$[h]_c = \begin{cases} h_R - h_L & \text{if } |h_R - h_L| \leq C \Delta x, \\ \text{sgn}(h_R - h_L) C \Delta x & \text{otherwise,} \end{cases} \quad (3.26)$$

with  $C$  a positive constant that does not depend on  $\Delta x$ . Equipped with this cutoff (3.26) within the expression (3.25) of  $\overline{h^{-\eta}}$ , it is clear from the proof of Lemma 2 that the approximate friction source term given by (3.24) becomes consistent, i.e. of order  $\mathcal{O}(\Delta x)$ , when  $k$  tends to 0. Finally, the approximate friction source term reads:

$$-k\bar{q}|\bar{q}|\overline{h^{-\eta}} = -k\bar{q}|\bar{q}|\frac{[h^2]}{2} \frac{\eta + 2}{[h^{\eta+2}]} + \frac{(\bar{q})^2}{\Delta x} [h]_c \left( \frac{-1}{h_L h_R} + \frac{h_L + h_R}{2} \frac{[h^{\eta-1}]}{\eta - 1} \frac{\eta + 2}{[h^{\eta+2}]} \right). \quad (3.27)$$

According to Lemma 2, the expression (3.27) is consistent with the friction source term  $-kq|q|h^{-\eta}$  for all  $k \geq 0$ .

Equipped with the expressions of  $\overline{h^{-\eta}}$  and  $\bar{q}$ , we have fully determined  $q^*$ . Now, we need to complete the system (3.13) to determine the intermediate water heights  $h_L^*$  and  $h_R^*$ . The following relation is chosen so as to ensure that  $W_L^*$  and  $W_R^*$  satisfy the property (WB):

$$\alpha^f (h_R^* - h_L^*) = -k\bar{q}|\bar{q}|\overline{h^{-\eta}} \Delta x, \quad \text{where} \quad \alpha^f = \frac{-\bar{q}^2}{h_L h_R} + \frac{g}{2} (h_L + h_R). \quad (3.28)$$

Therefore,  $h_L^*$  and  $h_R^*$  are easily determined by considering the following linear system, defined by (3.13a) and (3.28):

$$\begin{cases} \lambda_R h_R^* - \lambda_L h_L^* = (\lambda_R - \lambda_L) h_{HLL}, \\ \alpha^f (h_R^* - h_L^*) = -k\bar{q}|\bar{q}|\overline{h^{-\eta}} \Delta x. \end{cases} \quad (3.29)$$

The system (3.29) then provides the following formulas:

$$h_L^* = h_{HLL} + \frac{\lambda_R k \bar{q} |\bar{q}| \overline{h^{-\eta}} \Delta x}{\alpha^f (\lambda_R - \lambda_L)}, \quad (3.30a)$$

$$h_R^* = h_{HLL} + \frac{\lambda_L k \bar{q} |\bar{q}| \overline{h^{-\eta}} \Delta x}{\alpha^f (\lambda_R - \lambda_L)}. \quad (3.30b)$$

*Remark 4.* Note that, in (3.30), a division by  $\alpha^f$  occurs. In rare transcritical cases, the term  $\alpha^f$  could become close to zero, which could lead to a significant error in the numerical solution. However, many simulations have been performed and this problem has not appeared.



### 3.3. Properties satisfied by the approximate Riemann solver

In the previous subsection, we have constructed intermediate states (3.13b) - (3.30) that yield a consistent and well-balanced approximate Riemann solver. However, one can plainly see from (3.30) that the intermediate water heights  $h_L^*$  and  $h_R^*$  can become negative. If that is the case, (3.6) can yield a negative updated water height, which means the updated state would not belong to the admissible space  $\Omega$ . To recover a physically admissible solution, we apply the procedure from [2, 6, 40], which consists in introducing a parameter  $\varepsilon > 0$ , that controls the positivity of  $h_L^*$  and  $h_R^*$ .

We now state the expressions of the intermediate states obtained with this procedure, for given  $W_L$  and  $W_R$ . The parameter  $\varepsilon$  is such that

$$\varepsilon = \min(h_L, h_R, h_{HLL}). \quad (3.31)$$

For the sake of simplicity, we introduce the following notation:

$$\bar{S}^f := -k\bar{q}|\bar{q}|\bar{h}^{-\eta}, \quad (3.32)$$

where  $\bar{q}$  and  $\bar{h}^{-\eta}$  are respectively defined by (3.22) and (3.25). Then, the intermediate states  $W_L^* = {}^t(h_L^*, q_L^*)$  and  $W_R^* = {}^t(h_R^*, q_R^*)$  of the approximate Riemann solver (3.2) are given by:

$$\alpha^f = \frac{-\bar{q}^2}{h_L h_R} + \frac{g}{2}(h_L + h_R), \quad (3.33a)$$

$$q_L^* = q_R^* = q^* = q_{HLL} + \frac{\bar{S}^f \Delta x}{\lambda_R - \lambda_L}, \quad (3.33b)$$

$$h_L^* = \min\left(\max\left(h_{HLL} - \frac{\lambda_R \bar{S}^f \Delta x}{\alpha^f (\lambda_R - \lambda_L)}, \varepsilon\right), \left(1 - \frac{\lambda_R}{\lambda_L}\right) h_{HLL} + \frac{\lambda_R}{\lambda_L} \varepsilon\right), \quad (3.33c)$$

$$h_R^* = \min\left(\max\left(h_{HLL} - \frac{\lambda_L \bar{S}^f \Delta x}{\alpha^f (\lambda_R - \lambda_L)}, \varepsilon\right), \left(1 - \frac{\lambda_L}{\lambda_R}\right) h_{HLL} + \frac{\lambda_L}{\lambda_R} \varepsilon\right). \quad (3.33d)$$

The quantities  $q^*$  and  $\alpha^f$  have respectively been defined by (3.13b) and (3.28). Moreover, in (3.33), the positivity correction has been applied, and new expressions (3.33c) and (3.33d) of  $h_L^*$  and  $h_R^*$  have been substituted to (3.30a) and (3.30b). Note that the intermediate states (3.33) are written under the same form as the intermediate states presented in [40].

The following two results sum up the properties we have obtained thus far, for the approximate Riemann solver and the full scheme. For the sake of conciseness, we do not include their proofs here, since they use classical ingredients that have been presented in [40].

**Lemma 3.** *Assume  $\varepsilon > 0$  such that (3.31) is satisfied. Then, intermediate states  $W_L^* = {}^t(h_L^*, q_L^*)$  and  $W_R^* = {}^t(h_R^*, q_R^*)$  given by (3.33) satisfy the following properties:*

- (i) consistency: the quantities  $h_L^*$ ,  $h_R^*$ ,  $q_L^*$  and  $q_R^*$  satisfy the equations (3.13);
- (ii) positivity preservation: if  $h_L > 0$ ,  $h_R > 0$  and  $h_{HLL} > 0$ , then  $h_L^* \geq \varepsilon$  and  $h_R^* \geq \varepsilon$ ;
- (iii) well-balance:  $W_L^* = W_L$  and  $W_R^* = W_R$  satisfy the property (WB).

**Theorem 4.** Consider  $W_i^n \in \Omega^*$  for all  $i \in \mathbb{Z}$ , where  $\Omega^*$  is a restricted admissible states space defined as follows:

$$\Omega^* = \{W = {}^t(h, q) \in \mathbb{R}^2 ; h > 0, q \in \mathbb{R}\}.$$

Assume that the intermediate states  $W_{i+\frac{1}{2}}^{L,*}$  and  $W_{i+\frac{1}{2}}^{R,*}$  are given, for all  $i \in \mathbb{Z}$ , by

$$W_{i+\frac{1}{2}}^{L,*} = \begin{pmatrix} h_L^*(W_i^n, W_{i+1}^n) \\ q^*(W_i^n, W_{i+1}^n) \end{pmatrix} \quad \text{and} \quad W_{i+\frac{1}{2}}^{R,*} = \begin{pmatrix} h_R^*(W_i^n, W_{i+1}^n) \\ q^*(W_i^n, W_{i+1}^n) \end{pmatrix},$$

where  $q^*$ ,  $h_L^*$  and  $h_R^*$  are given by (3.33b), (3.33c) and (3.33d), respectively. Also, assume  $\varepsilon > 0$  given by (3.31). Then the Godunov-type scheme, given by (3.6) under the CFL restriction (3.5), satisfies the following properties:

1. consistency with the shallow-water system (3.7);
2. positivity preservation:  $\forall i \in \mathbb{Z}, W_i^{n+1} \in \Omega^*$ ;
3. well-balance: if  $(W_i^n)_{i \in \mathbb{Z}}$  defines a steady state, then  $\forall i \in \mathbb{Z}, W_i^{n+1} = W_i^n$ .

From the above result, we see that the updated water height never vanishes. This behavior is due to the introduction in (3.33) of the parameter  $\varepsilon$ . Such a behavior is necessary because the average  $\overline{h^{-\eta}}$  has been obtained by considering positive water heights. In a later section, devoted to dry/wet transitions, this average will be extended to be able to deal with dry areas. This extension will allow us to consider  $\varepsilon = 0$  in the definition (3.33), as well as vanishing water heights.

#### 4. The case of both topography and friction source terms

In this section, our goal is the derivation of new intermediate states  $h_L^*$ ,  $h_R^*$  and  $q^*$  to approximate the contributions of both topography and friction source terms. For the sake of simplicity in the notations, we write  $S = S^t + S^f$ , where

$$\begin{aligned} S^t &= -gh\partial_x Z, \\ S^f &= -kq|q|h^{-\eta}, \end{aligned}$$

which corresponds to the case where both topography and friction source terms are involved. The steady states associated to the full source term  $S$  do not admit an algebraic expression. Therefore  $S$  is split between  $S^t$  and  $S^f$ , since the steady states for the individual source terms of topography and friction can both be rewritten using an algebraic expression. In order to complete the scheme, we

need to determine suitable intermediate states in both cases of the topography and the friction. The case of the friction has been treated in the previous section, and the intermediate states are given by (3.33). For the topography, we adopt the approach introduced in [40], which is now recalled for the sake of completeness.

#### 4.1. Approximation of the topography source term

Considering a non-vanishing topography source term but with a vanishing friction, the intermediate states are given by

$$\alpha^t = \frac{-(q^*)^2}{h_L h_R} + \frac{g}{2}(h_L + h_R), \quad (4.1a)$$

$$q_L^* = q_R^* = q^* = q_{HLL} + \frac{\bar{S}^t \Delta x}{\lambda_R - \lambda_L}, \quad (4.1b)$$

$$h_L^* = \min \left( \max \left( h_{HLL} - \frac{\lambda_R \bar{S}^t \Delta x}{\alpha^t (\lambda_R - \lambda_L)}, \varepsilon \right), \left( 1 - \frac{\lambda_R}{\lambda_L} \right) h_{HLL} + \frac{\lambda_R}{\lambda_L} \varepsilon \right), \quad (4.1c)$$

$$h_R^* = \min \left( \max \left( h_{HLL} - \frac{\lambda_L \bar{S}^t \Delta x}{\alpha^t (\lambda_R - \lambda_L)}, \varepsilon \right), \left( 1 - \frac{\lambda_L}{\lambda_R} \right) h_{HLL} + \frac{\lambda_L}{\lambda_R} \varepsilon \right). \quad (4.1d)$$

In (4.1), the approximate topography source term  $\bar{S}^t$  is defined by:

$$\bar{S}^t \Delta x := \bar{S}^t(h_L, h_R, Z_L, Z_R) \Delta x = -2g[Z] \frac{h_L h_R}{h_L + h_R} + \frac{g}{2} \frac{[h]_c^3}{h_L + h_R}, \quad (4.2)$$

where  $[h]_c$  is the cutoff of  $[h] = h_R - h_L$  defined by (3.26). This expression of  $\bar{S}^t$  is consistent with the topography source term  $S^t$ , and it ensures that the scheme is well-balanced (see [4, 5, 40] for more details regarding this expression and how to obtain it). Computations leading to this expression and to the expressions (4.1) have been explained in [40], to which the reader is referred for more details.

#### 4.2. Approximation of the topography and friction source terms

Now, we focus on deriving intermediate states for the shallow-water equations with both topography and friction source terms. We introduce the following discretization of (1.7):

$$q_0^2 \left[ \frac{1}{h} \right] + \frac{g}{2} [h^2] = \bar{S}^t \Delta x + \bar{S}^f \Delta x, \quad (4.3)$$

where the approximate source term  $\bar{S}$  has been split into the topography and friction contributions, by writing  $\bar{S} = \bar{S}^t + \bar{S}^f$ . As a consequence, after (3.33b), we define the following intermediate discharge:

$$q^* = q_{HLL} + \frac{\bar{S}^t \Delta x}{\lambda_R - \lambda_L} + \frac{\bar{S}^f \Delta x}{\lambda_R - \lambda_L}, \quad (4.4)$$

where the expressions of  $\bar{S}^t$  and  $\bar{S}^f$  are given by (4.2) and (3.32), respectively. Concerning the intermediate heights, we first decide to use the following expressions:

$$\begin{aligned}\widetilde{h}_L^* &= h_{HLL} - \frac{\lambda_R \bar{S}^t \Delta x}{\alpha^t (\lambda_R - \lambda_L)} - \frac{\lambda_R \bar{S}^f \Delta x}{\alpha^f (\lambda_R - \lambda_L)}, \\ \widetilde{h}_R^* &= h_{HLL} - \frac{\lambda_L \bar{S}^t \Delta x}{\alpha^t (\lambda_R - \lambda_L)} - \frac{\lambda_L \bar{S}^f \Delta x}{\alpha^f (\lambda_R - \lambda_L)},\end{aligned}\tag{4.5}$$

where the quantities  $\alpha^t$  and  $\alpha^f$  are respectively defined by (4.1a) and (3.33a). We immediately see that the intermediate states (4.4) – (4.5) still yield a consistent scheme. Moreover, it is possible to recover the positivity of the water heights. Let  $\varepsilon \geq 0$  that satisfies (3.31), and consider the same procedure as used for the friction source term, which yields the following intermediate states:

$$\begin{aligned}q^* &= q_{HLL} + \frac{\bar{S}^t \Delta x}{\lambda_R - \lambda_L} + \frac{\bar{S}^f \Delta x}{\lambda_R - \lambda_L}, \\ h_L^* &= \min \left( \max \left( \widetilde{h}_L^*, \varepsilon \right), \left( 1 - \frac{\lambda_R}{\lambda_L} \right) h_{HLL} + \frac{\lambda_R}{\lambda_L} \varepsilon \right), \\ h_R^* &= \min \left( \max \left( \widetilde{h}_R^*, \varepsilon \right), \left( 1 - \frac{\lambda_L}{\lambda_R} \right) h_{HLL} + \frac{\lambda_L}{\lambda_R} \varepsilon \right).\end{aligned}\tag{4.6}$$

From Lemma 3, as soon as  $\varepsilon > 0$ , the intermediate heights  $h_L^*$  and  $h_R^*$  defined by (4.6) are positive.

We note that using the definitions (4.6) and making the friction source term vanish allows the recovery of the intermediate states for topography only. Similarly, if the topography source term vanishes, we recover the intermediate states for friction only. As a consequence, (4.6) yields intermediate states that are well-balanced for the individual source terms of topography or friction.

Let us recall that the steady states relation for the shallow-water system with both topography and friction source terms (1.5) cannot be written under the form of an algebraic relation for all  $Z$ . Therefore, we only manage to preserve the steady states up to the chosen discretization (4.3) of (1.7) (see [20, 33, 38, 51, 52] where a similar approach is used).

We finally study how the approximate source term averages  $\bar{S}^t$  and  $\bar{S}^f$ , as well as the terms  $\bar{S}^t \Delta x / \alpha^t$  and  $\bar{S}^f \Delta x / \alpha^f$ , behave when dealing with vanishing water heights. First, we make the following assumption.

**Assumption.** When the height vanishes, so does the velocity.

From [40], under this assumption, the quantity  $\bar{S}^t \Delta x$  satisfies:

$$\begin{aligned}\bar{S}^t \Delta x &:= \bar{S}^t(h_L, h_R, Z_L, Z_R) \Delta x \\ &= \begin{cases} 0 & \text{if } h_L = 0 \text{ and } h_R = 0, \\ -g[Z] \frac{h_R + h_L}{2} & \text{if } h_L = 0 \text{ or } h_R = 0, \\ -g[Z] \frac{2h_L h_R}{h_L + h_R} + \frac{g}{2} \frac{[h]_c^3}{h_L + h_R} & \text{otherwise,} \end{cases}\end{aligned}\tag{4.7}$$

and the quantity  $\bar{S}^t \Delta x / \alpha^t$  satisfies:

$$\frac{\bar{S}^t \Delta x}{\alpha^t} = \begin{cases} 0 & \text{if } h_L = 0 \text{ and } h_R = 0, \\ -[Z] & \text{if } h_L = 0 \text{ or } h_R = 0, \\ \frac{\bar{S}^t \Delta x}{-\frac{(q^*)^2}{h_L h_R} + \frac{g}{2}(h_L + h_R)} & \text{otherwise.} \end{cases}$$

We then turn to the approximate friction source term. We have to make the following assumption in order to define it for vanishing water heights.

**Assumption.** Using  $\bar{S}^f$  and  $\alpha^f$  respectively defined by (3.32) and (3.33a), we impose that  $\bar{S}^f$  and  $\bar{S}^f / \alpha^f$  are zero as soon as  $h_L$  and/or  $h_R$  vanishes.

It makes sense to consider such a behavior. Indeed, as previously mentioned, the friction contribution is assumed to vanish as soon as the water height vanishes. In order for both quantities  $\bar{S}^f$  and  $\bar{S}^f / \alpha^f$  to satisfy this requirement, we have to impose that they vanish when  $h_L$  and/or  $h_R$  vanishes.

To handle the case where both  $h_L$  and  $h_R$  vanish (and thus  $q_L = q_R = 0$ ), we have to make sure that, in this case,  $q^* = 0$  and  $h_L^* = h_R^* = 0$ . This requirement is met by taking  $\bar{S}^t = 0$  and  $\bar{S}^t / \alpha^t = 0$ , as well as  $\bar{S}^f = 0$  and  $\bar{S}^f / \alpha^f = 0$ , as soon as both  $h_L$  and  $h_R$  are zero.

As a consequence, after (3.32) and the above assumption,  $\bar{S}^f \Delta x$  is given by:

$$\begin{aligned} \bar{S}^f \Delta x &:= \bar{S}^f(h_L, h_R, q_L, q_R) \Delta x \\ &= \begin{cases} 0 & \text{if } h_L = 0 \text{ and/or } h_R = 0, \\ -k\bar{q}|\bar{q}|h^{-\eta} & \text{otherwise,} \end{cases} \end{aligned} \quad (4.8)$$

where  $\bar{q}$  is given by (3.22) and  $h^{-\eta}$  is defined by (3.25). In addition, the quantity  $\bar{S}^f \Delta x / \alpha$  is given by:

$$\frac{\bar{S}^f \Delta x}{\alpha} = \begin{cases} 0 & \text{if } h_L = 0 \text{ and/or } h_R = 0, \\ \frac{-k\bar{q}|\bar{q}|h^{-\eta}}{-\frac{(q^*)^2}{h_L h_R} + \frac{g}{2}(h_L + h_R)} & \text{otherwise.} \end{cases}$$

We can now state the following result, concerning the approximate Riemann solver for both source terms of topography and friction. Since its proof uses classical ingredients, it is not presented here; instead, the reader is referred to [40].

**Lemma 5.** *The intermediate states (4.6) satisfy the following properties:*

- (i) *consistency with the shallow-water equations with topography and friction (1.1);*

(ii) *well-balance*: if  $W_L$  and  $W_R$  define a steady state, i.e. satisfy (4.3), then  $W_L^* = W_L$  and  $W_R^* = W_R$ .

Moreover, we have:

(iii) with  $\varepsilon > 0$ , the positivity is preserved: if  $h_L > 0$ ,  $h_R > 0$  and  $h_{HLL} > 0$ , then  $h_L^* \geq \varepsilon$  and  $h_R^* \geq \varepsilon$ ;

(iv) with  $\varepsilon = 0$ , the non-negativity is preserved: if  $h_L \geq 0$ ,  $h_R \geq 0$  and  $h_{HLL} \geq 0$ , then  $h_L^* \geq 0$  and  $h_R^* \geq 0$ .

This lemma allows us to state the following result, which concerns the full scheme (3.6). Similarly to Lemma 5, we do not prove this result here, and we refer the reader to [40].

**Theorem 6.** Consider  $W_i^n \in \Omega$  for all  $i \in \mathbb{Z}$ . Assume that the intermediate states  $W_{i+\frac{1}{2}}^{L,*}$  and  $W_{i+\frac{1}{2}}^{R,*}$  are given, for all  $i \in \mathbb{Z}$ , by

$$W_{i+\frac{1}{2}}^{L,*} = \left( \begin{array}{c} h_L^*(W_i^n, W_{i+1}^n) \\ q^*(W_i^n, W_{i+1}^n) \end{array} \right) \quad \text{and} \quad W_{i+\frac{1}{2}}^{R,*} = \left( \begin{array}{c} h_R^*(W_i^n, W_{i+1}^n) \\ q^*(W_i^n, W_{i+1}^n) \end{array} \right), \quad (4.9)$$

where  $q^*$ ,  $h_L^*$  and  $h_R^*$  are given by (4.6) with  $\varepsilon = 0$ , i.e. by

$$\begin{aligned} q^* &= q_{HLL} + \frac{\bar{S}^t \Delta x}{\lambda_R - \lambda_L} + \frac{\bar{S}^f \Delta x}{\lambda_R - \lambda_L}, \\ h_L^* &= \min \left( \left( h_{HLL} - \frac{\lambda_R \bar{S}^t \Delta x}{\alpha^t (\lambda_R - \lambda_L)} - \frac{\lambda_R \bar{S}^f \Delta x}{\alpha^f (\lambda_R - \lambda_L)} \right)_+, \left( 1 - \frac{\lambda_R}{\lambda_L} \right) h_{HLL} \right), \\ h_R^* &= \min \left( \left( h_{HLL} - \frac{\lambda_L \bar{S}^t \Delta x}{\alpha^t (\lambda_R - \lambda_L)} - \frac{\lambda_L \bar{S}^f \Delta x}{\alpha^f (\lambda_R - \lambda_L)} \right)_+, \left( 1 - \frac{\lambda_L}{\lambda_R} \right) h_{HLL} \right). \end{aligned}$$

Then the Godunov-type scheme, given by (3.6) under the CFL restriction (3.5), satisfies the following properties:

1. consistency with the shallow-water system (1.1);
2. non-negativity preservation:  $\forall i \in \mathbb{Z}, W_i^{n+1} \in \Omega$ ;
3. well-balance: if  $(W_i^n)_{i \in \mathbb{Z}}$  defines a steady state according to (4.3), then  $\forall i \in \mathbb{Z}, W_i^{n+1} = W_i^n$ . The same result holds for both topography-only and friction-only steady states, respectively given by (1.9) and (1.10).

## 5. Implication of the source terms contribution: wet/dry transitions

The scheme (3.6) - (4.9) allows the simulation of wet/dry transitions. However, several spurious oscillations appear when such experiments are performed. The stiffness of the source terms in the vicinity of wet/dry transitions, especially the friction, induces these spurious oscillations. To correct such a failure, we first rewrite this scheme in order to exhibit the numerical flux function and the source terms contribution. Then, we adopt an explicit scheme for the transport part, and an implicit scheme for the source part.

### 5.1. Reformulation of the scheme

In this subsection, we exhibit the numerical flux function and the numerical source terms. To that end, the scheme (3.6), with intermediate states given by (4.6), is rewritten after straightforward computations as follows (see for instance [30]):

$$W_i^{n+1} = W_i^n - \frac{\Delta t}{\Delta x} \left( f_{i+\frac{1}{2}}^n - f_{i-\frac{1}{2}}^n \right) + \frac{\Delta t}{2} \left( s_{i+\frac{1}{2}}^n + s_{i-\frac{1}{2}}^n \right). \quad (5.1)$$

The quantity  $f_{i+\frac{1}{2}}^n = f(W_i^n, W_{i+1}^n)$  is the numerical flux function evaluated at the interface  $x_{i+\frac{1}{2}}$ , and the quantity  $s_{i+\frac{1}{2}}^n$  is the numerical source term at the interface  $x_{i+\frac{1}{2}}$ . These quantities are approximations of the flux and the source term, respectively, and are defined by:

$$f_{i+\frac{1}{2}}^n = \begin{pmatrix} (f^h)_{i+\frac{1}{2}}^n \\ (f^q)_{i+\frac{1}{2}}^n \end{pmatrix} \quad \text{and} \quad s_{i+\frac{1}{2}}^n = \begin{pmatrix} 0 \\ (S^t)_{i+\frac{1}{2}}^n + (S^f)_{i+\frac{1}{2}}^n \end{pmatrix}. \quad (5.2)$$

The quantities  $(S^t)_{i+\frac{1}{2}}^n$  and  $(S^f)_{i+\frac{1}{2}}^n$  are approximations of the topography and the friction source terms, respectively. Adopting extended notations, they are given by:

$$(S^t)_{i+\frac{1}{2}}^n = \bar{S}^t(h_i^n, h_{i+1}^n, Z_i, Z_{i+1}), \quad (5.3a)$$

$$(S^f)_{i+\frac{1}{2}}^n = \bar{S}^f(h_i^n, h_{i+1}^n, q_i^n, q_{i+1}^n), \quad (5.3b)$$

where  $\bar{S}^t$  and  $\bar{S}^f$  are the approximate source terms already defined by (4.7) and (4.8). The scheme (5.1) then reads:

$$h_i^{n+1} = h_i^n - \frac{\Delta t}{\Delta x} \left( (f^h)_{i+\frac{1}{2}}^n - (f^h)_{i-\frac{1}{2}}^n \right), \quad (5.4a)$$

$$q_i^{n+1} = q_i^n - \frac{\Delta t}{\Delta x} \left( (f^q)_{i+\frac{1}{2}}^n - (f^q)_{i-\frac{1}{2}}^n \right) + \frac{\Delta t}{2} \left( (S^t)_{i+\frac{1}{2}}^n + (S^t)_{i-\frac{1}{2}}^n + (S^f)_{i+\frac{1}{2}}^n + (S^f)_{i-\frac{1}{2}}^n \right), \quad (5.4b)$$

where the approximate fluxes are defined by (5.2) and

$$f_{i+\frac{1}{2}}^n = \frac{1}{2} \left( F(W_i^n) + F(W_{i+1}^n) \right) + \frac{\lambda_{i+\frac{1}{2}}^L}{2} \left( W_{i+\frac{1}{2}}^{L,*} - W_i^n \right) + \frac{\lambda_{i+\frac{1}{2}}^R}{2} \left( W_{i+\frac{1}{2}}^{R,*} - W_{i+1}^n \right). \quad (5.5)$$

For the sake of simplicity, we set the notations:

$$(S^t)_i^n = \frac{1}{2} \left( (S^t)_{i-\frac{1}{2}}^n + (S^t)_{i+\frac{1}{2}}^n \right) \quad \text{and} \quad (S^f)_i^n = \frac{1}{2} \left( (S^f)_{i-\frac{1}{2}}^n + (S^f)_{i+\frac{1}{2}}^n \right), \quad (5.6)$$

such that (5.4b) rewrites:

$$q_i^{n+1} = q_i^n - \frac{\Delta t}{\Delta x} \left( (f^q)_{i+\frac{1}{2}}^n - (f^q)_{i-\frac{1}{2}}^n \right) + \Delta t (S^t)_i^n + \Delta t (S^f)_i^n. \quad (5.7)$$

### 5.2. Implication of the source terms contribution

We now introduce a semi-implicit version of the Godunov-type scheme (5.1). The main idea of this section is to use a splitting method (see for instance [10, 45]) to reduce the impact of the aforementioned instabilities. The splitting strategy we use here is to first consider an explicit treatment of the flux, then an implicit treatment of both source terms. As a consequence, the first step, devoted to the transport part  $\partial_t W + \partial_x f(W) = 0$ , reads as follows:

$$\begin{aligned} h_i^{n+\frac{1}{3}} &= h_i^n - \frac{\Delta t}{\Delta x} \left( (f^h)_{i+\frac{1}{2}}^n - (f^h)_{i-\frac{1}{2}}^n \right), \\ q_i^{n+\frac{1}{3}} &= q_i^n - \frac{\Delta t}{\Delta x} \left( (f^q)_{i+\frac{1}{2}}^n - (f^q)_{i-\frac{1}{2}}^n \right). \end{aligned} \quad (5.8)$$

During the second step, devoted to the topography source term, we approximately solve the initial value problem

$$\begin{cases} \frac{dh}{dt} = 0, \\ \frac{dq}{dt} = S^t(W), \end{cases} \quad \text{with initial data} \quad \begin{cases} h(0) = h_i^{n+\frac{1}{3}}, \\ q(0) = q_i^{n+\frac{1}{3}}, \end{cases}$$

where  $S^t(W) = -gh\partial_t Z$ . We suggest the following implicit scheme to approximate  $h_i^{n+\frac{2}{3}}$  and  $q_i^{n+\frac{2}{3}}$ :

$$\begin{cases} h_i^{n+\frac{2}{3}} = h_i^{n+\frac{1}{3}}, \\ q_i^{n+\frac{2}{3}} = q_i^{n+\frac{1}{3}} + \Delta t (S^t)_i^{n+\frac{2}{3}}, \end{cases} \quad (5.9a)$$

$$(5.9b)$$

where, as per (5.6),  $(S^t)_i^{n+\frac{2}{3}} = \left( (S^t)_{i+\frac{1}{2}}^{n+\frac{2}{3}} + (S^t)_{i-\frac{1}{2}}^{n+\frac{2}{3}} \right) / 2$ , with, after (5.3a),

$$(S^t)_{i+\frac{1}{2}}^{n+\frac{2}{3}} = \bar{S}^t \left( h_i^{n+\frac{2}{3}}, h_{i+1}^{n+\frac{2}{3}}, Z_i, Z_{i+1} \right).$$

Let us underline that, during the second step, the evaluation of the implicit unknowns is immediate since the topography source term involves known quantities.

The third and last step concerns the friction and consists in solving

$$\begin{cases} \frac{dh}{dt} = 0, \\ \frac{dq}{dt} = -kq|q|(h_i^{n+1})^{-\eta}, \end{cases} \quad \text{with initial data} \quad \begin{cases} h(0) = h_i^{n+\frac{2}{3}}, \\ q(0) = q_i^{n+\frac{2}{3}}. \end{cases} \quad (5.10)$$

This system can be solved to find an analytic expression of the solution. For  $t \in [0, \Delta t]$ , the exact solution of the above system reads as follows:

$$\begin{cases} h(t) = h(0), \\ q(t) = \frac{h(0)^\eta q(0)}{h(0)^\eta + kt|q(0)|}. \end{cases} \quad (5.11)$$



Note that the analytic expression (5.11) guarantees that, for all  $t \in (0, \Delta t]$ , the sign of  $q(t)$  stays the same as the sign of  $q(0)$ , and that  $|q(t)| < |q(0)|$ . This behavior is consistent with the fact that friction should only slow down the movement of the fluid, rather than changing its direction. Then, evaluating (5.11) at  $t = \Delta t$  and plugging the initial data yields

$$\begin{cases} h_i^{n+1} = h_i^{n+\frac{2}{3}}, & (5.12a) \\ q_i^{n+1} = \frac{(h_i^{n+1})^\eta q_i^{n+\frac{2}{3}}}{(h_i^{n+1})^\eta + k \Delta t |q_i^{n+\frac{2}{3}}|}. & (5.12b) \end{cases}$$

Let us note that the updated state  ${}^t(h_i^{n+1}, q_i^{n+1})$  retains its non-negativity preservation property, since the computations for  $h_i^{n+1}$  have not been modified and [Theorem 6](#) applies. Furthermore, the expression (5.12b) is well-posed for  $h_i^{n+1} = 0$  and, when that is the case, forces the discharge  $q_i^{n+1}$  to vanish. However, the well-balance property is lost. Indeed, if  $W_{i-1}^n, W_i^n$  and  $W_{i+1}^n$  define a steady state, we do not necessarily recover  $q_i^{n+1} = q_i^n$ . As a consequence, we decide to consider an approximation of  $(h_i^{n+1})^\eta$ , denoted by  $(\bar{h}^\eta)_i^{n+1}$ , in (5.12b). This approximation is determined in order to ensure the required well-balance of the scheme.

In order to obtain such an expression of  $(\bar{h}^\eta)_i^{n+1}$ , we momentarily suppose that  $W_{i-1}^n, W_i^n$  and  $W_{i+1}^n$  define a steady state. In this case, we need to ensure that  $W_i^{n+1} = W_i^n$ . Since the scheme (5.1) is well-balanced, the equation (5.7) yields  $q_i^{n+1} = q_i^n$ , and can be rewritten as:

$$(S^f)_i^n = \frac{1}{\Delta x} \left( (f^q)_{i+\frac{1}{2}}^n - (f^q)_{i-\frac{1}{2}}^n \right) - (S^t)_i^n. \quad (5.13)$$

Moreover, we have  $h_i^{n+\frac{2}{3}} = h_i^n$ . Therefore, by the definition (5.3a) of  $(S^t)_i^n$ , we have  $(S^t)_i^{n+\frac{2}{3}} = (S^t)_i^n$ . The evaluation of  $q_i^{n+\frac{2}{3}}$  is then obtained from (5.9b) as follows:

$$q_i^{n+\frac{2}{3}} = q_i^n - \frac{\Delta t}{\Delta x} \left( (f^q)_{i+\frac{1}{2}}^n - (f^q)_{i-\frac{1}{2}}^n \right) + \Delta t (S^t)_i^n. \quad (5.14)$$

Since  $W_{i-1}^n, W_i^n$  and  $W_{i+1}^n$  define a steady state, the relations (5.13) and (5.14) yield:

$$q_i^{n+\frac{2}{3}} = q_i^n - \Delta t (S^f)_i^n.$$

Thus, using  $(\bar{h}^\eta)_i^{n+1}$ , (5.12b) reads

$$q_i^n = \frac{(\bar{h}^\eta)_i^{n+1} (q_i^n - \Delta t (S^f)_i^n)}{(\bar{h}^\eta)_i^{n+1} + k \Delta t |q_i^n - \Delta t (S^f)_i^n|}. \quad (5.15)$$

We are now able to determine the expression of  $(\bar{h}^\eta)_i^{n+1}$  that ensures the well-balance of the scheme. We set  $\mu_i^n = \text{sgn } q_i^n$  and  $\mu_i^{n+\frac{2}{3}} = \text{sgn } q_i^{n+\frac{2}{3}}$ . With these notations, we get from relation (5.15) the following expression of  $(\bar{h}^\eta)_i^{n+1}$ :

$$(\bar{h}^\eta)_i^{n+1} = -\frac{k(q_i^n)^2 \mu_i^{n+\frac{2}{3}}}{(S^f)_i^n} + k \Delta t \mu_i^{n+\frac{2}{3}} q_i^n. \quad (5.16)$$

Now, recall that the numerical source term  $(\bar{S}^f)_i^n$  is defined by (5.6). As a consequence, we get the following expression for  $(\bar{S}^f)_i^n$ :

$$(S^f)_i^n = \frac{1}{2} \left( -k\bar{q}_{i-\frac{1}{2}}^n |\bar{q}_{i-\frac{1}{2}}^n| (\bar{h}^{-\eta})_{i-\frac{1}{2}}^n - k\bar{q}_{i+\frac{1}{2}}^n |\bar{q}_{i+\frac{1}{2}}^n| (\bar{h}^{-\eta})_{i+\frac{1}{2}}^n \right),$$

where the averages  $\bar{q}_{i\pm\frac{1}{2}}^n$  and  $(\bar{h}^{-\eta})_{i\pm\frac{1}{2}}^n$  are given with clear notations by (3.22) and (3.25), respectively. However, recall that the only requirement that has to be fulfilled by the average  $\bar{q}$  is to recover the steady discharge  $q_0$  as soon as a steady state was reached. In the current context,  $W_{i-1}^n$ ,  $W_i^n$  and  $W_{i+1}^n$  define a steady state; hence, we have  $q_{i-1}^n = q_i^n = q_{i+1}^n = q_0$ . A relevant choice is therefore to take  $\bar{q}_{i\pm\frac{1}{2}}^n = q_i^n$ , which yields the following formula:

$$(S^f)_i^n = \frac{1}{2} \left( -kq_i^n |q_i^n| (\bar{h}^{-\eta})_{i-\frac{1}{2}}^{n+1} - kq_i^n |q_i^n| (\bar{h}^{-\eta})_{i+\frac{1}{2}}^{n+1} \right). \quad (5.17)$$

In (5.17), we have substituted  $(\bar{h}^{-\eta})_{i\pm\frac{1}{2}}^n$  with  $(\bar{h}^{-\eta})_{i\pm\frac{1}{2}}^{n+1}$ . This substitution has no effect on the well-balance property, and considering the updated water height makes the scheme contain more implicit terms.

With this simplification in the source term approximation, we get the following expression for  $(\bar{h}^\eta)_i^{n+1}$ , from (5.16) and (5.17):

$$(\bar{h}^\eta)_i^{n+1} = \frac{2\mu_i^{n+\frac{2}{3}} \mu_i^n}{(\bar{h}^{-\eta})_{i-\frac{1}{2}}^{n+1} + (\bar{h}^{-\eta})_{i+\frac{1}{2}}^{n+1}} + k \Delta t \mu_i^{n+\frac{2}{3}} q_i^n.$$

Arguing the formula (3.25), the above expression can then be rewritten as

$$(\bar{h}^\eta)_i^{n+1} = \frac{2k\mu_i^{n+\frac{2}{3}} \Delta x}{k\mu_i^n \Delta x \left( \beta_{i-\frac{1}{2}}^{n+1} + \beta_{i+\frac{1}{2}}^{n+1} \right) - \left( \gamma_{i-\frac{1}{2}}^{n+1} + \gamma_{i+\frac{1}{2}}^{n+1} \right)} + k \Delta t \mu_i^{n+\frac{2}{3}} q_i^n, \quad (5.18)$$

where we have set

$$\beta_{i+\frac{1}{2}}^{n+1} = \frac{\eta + 2}{2} \frac{(h_{i+1}^{n+1})^2 - (h_i^{n+1})^2}{(h_{i+1}^{n+1})^{\eta+2} - (h_i^{n+1})^{\eta+2}}, \text{ and}$$

$$\gamma_{i+\frac{1}{2}}^{n+1} = \frac{1}{h_{i+1}^{n+1}} - \frac{1}{h_i^{n+1}} + \beta_{i+\frac{1}{2}}^{n+1} \frac{(h_{i+1}^{n+1})^{\eta-1} - (h_i^{n+1})^{\eta-1}}{\eta - 1}.$$

Computations within the expression of  $(\bar{h}^\eta)_i^{n+1}$  show that it tends to 0 as soon as  $h_{i-1}^{n+1}$ ,  $h_i^{n+1}$  or  $h_{i+1}^{n+1}$  tends to 0, which is the correct behavior when dealing with wet/dry transitions. We have therefore devised a way to consider the source terms contribution in an implicit way, while still retaining the well-balance property of the scheme. We can thus state the following result.

**Theorem 7.** *The scheme (5.8) - (5.9) - (5.12) - (5.18) is positivity-preserving and well-balanced. Indeed, if  $(W_i^n)_{i \in \mathbb{Z}}$  defines a steady state according to (4.3), then  $\forall i \in \mathbb{Z}$ ,  $W_i^{n+1} = W_i^n$ . The same result holds for both topography-only and friction-only steady states, respectively given by (1.9) and (1.10).*

*Proof.* We begin this proof by recalling that, from (5.8), (5.9) and (5.12), we have  $h_i^{n+1} = h_i^{n+\frac{1}{3}}$ . Therefore, since  $h_i^{n+\frac{1}{3}}$  is given by the full scheme (3.6), which is positivity-preserving, the scheme under consideration is positivity-preserving. We now have to show that it is well-balanced. To that end, assume that  $W_{i-1}^n$ ,  $W_i^n$  and  $W_{i+1}^n$  define a steady state, according to (4.3), (1.9) or (1.10). Using a similar chain of arguments as above, we immediately recover that  $h_i^{n+1} = h_i^n$ . To complete the proof, we now have to show that  $q_i^{n+1} = q_i^n$ . The updated discharge  $q_i^{n+1}$  is given by (5.12), with  $q_i^{n+\frac{2}{3}}$  defined by (5.9) and (5.8). Since  $(\bar{h}^n)_i^{n+1}$  is given by (5.18) and has been chosen to ensure  $q_i^{n+1} = q_i^n$ , the proof is concluded.  $\square$

## 6. Second-order MUSCL extension

We devote this section to a second-order extension based on a MUSCL technique (for instance, see [47, 48, 34, 35, 45]), to improve the space accuracy of the scheme. The MUSCL procedure involves a piecewise linear reconstruction, instead of piecewise constant, in the Godunov-type scheme. In addition, we use Heun's method to increase the time accuracy of the scheme.

### 6.1. The MUSCL reconstruction

The variables to be reconstructed are  $h$ ,  $q$  and  $h + Z$ . Thus, note that the approximation of  $Z$  turns out to be time-dependent. For the sake of simplicity and conciseness in the notations, consider  $w \in \{h, q, h + Z\}$ . The reconstruction procedure consists in replacing the constant state  $w_i^n$  with a linear approximation, given in each cell  $(x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}})$  by

$$w_i^n(x) = w_i^n + (x - x_i) \sigma_i^n,$$

where  $\sigma_i^n$  is the limited slope of the linear reconstruction. A limiter is applied in order to improve the stability of the scheme. Here, we have chosen the classical minmod limiter (the reader is referred for instance to [35] for more details regarding the use of slope limiters and a wider range of limiters). The limited slope  $\sigma_i^n$  is thus given by

$$\sigma_i^n = \text{minmod} \left( \frac{w_{i+1}^n - w_i^n}{\Delta x}, \frac{w_i^n - w_{i-1}^n}{\Delta x} \right),$$

where the minmod function is defined by

$$\text{minmod}(a, b) = \begin{cases} a & \text{if } |a| < |b| \text{ and } ab > 0, \\ b & \text{if } |a| > |b| \text{ and } ab > 0, \\ 0 & \text{if } ab \leq 0. \end{cases}$$

Thus, at the inner interfaces of the cell  $(x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}})$ , the reconstructed variables are given by

$$\begin{aligned} w_i^- &:= w_i^n \left( x_i - \frac{\Delta x}{2} \right) = w_i^n - \frac{\Delta x}{2} \sigma_i^n, \\ w_i^+ &:= w_i^n \left( x_i + \frac{\Delta x}{2} \right) = w_i^n + \frac{\Delta x}{2} \sigma_i^n. \end{aligned} \quad (6.1)$$

Since  $w \in \{h, q, h+Z\}$ , the reconstructed value of  $Z$  at the interfaces is computed from the reconstructed values of  $h+Z$  and  $h$  at the interfaces.

Therefore, the updated states are given by the following three step scheme.

- Transport step:

$$\begin{aligned} h_i^{n+\frac{1}{3}} &= h_i^n - \frac{\Delta t}{\Delta x} \left( (f^h)_{i+\frac{1}{2}}^n - (f^h)_{i-\frac{1}{2}}^n \right), \\ q_i^{n+\frac{1}{3}} &= q_i^n - \frac{\Delta t}{\Delta x} \left( (f^q)_{i+\frac{1}{2}}^n - (f^q)_{i-\frac{1}{2}}^n \right), \end{aligned} \quad (6.2)$$

where  $(f^h)_{i+\frac{1}{2}}^n := f^h(W_i^+, W_{i+1}^-)$  and  $(f^q)_{i+\frac{1}{2}}^n := f^q(W_i^+, W_{i+1}^-)$  according to (5.5). The water height  $h_i^{n+\frac{1}{3}}$  is then reconstructed within the cell  $(x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}})$ , following (6.1), to get  $h_i^-$  and  $h_i^+$ .

- Topography step:

$$\begin{aligned} h_i^{n+\frac{2}{3}} &= h_i^{n+\frac{1}{3}}, \\ q_i^{n+\frac{2}{3}} &= q_i^{n+\frac{1}{3}} + \Delta t (S^t)_i^{n+\frac{2}{3}}, \\ (S^t)_i^{n+\frac{2}{3}} &= \frac{1}{2} \left( \bar{S}^t(h_{i-1}^+, h_i^-, Z_{i-1}^+, Z_i^-) + \bar{S}^t(h_i^+, h_{i+1}^-, Z_i^+, Z_{i+1}^-) \right), \end{aligned} \quad (6.3)$$

where  $\bar{S}^t$  is defined by (4.7).

- Friction step:

$$\begin{aligned} h_i^{n+1} &= h_i^{n+\frac{2}{3}}, \\ q_i^{n+1} &= \frac{((\bar{h}^\eta)_i^{n+1})^\eta q_i^{n+\frac{2}{3}}}{((\bar{h}^\eta)_i^{n+1})^\eta + k \Delta t |q_i^{n+\frac{2}{3}}|}, \\ (\bar{h}^\eta)_i^{n+1} &= \frac{2\mu_i^{n+\frac{2}{3}} \mu_i^n}{\left(\bar{h}^\eta\right)_{i-\frac{1}{2}}^{n+1} + \left(\bar{h}^\eta\right)_{i+\frac{1}{2}}^{n+1}} + k \Delta t \mu_i^{n+\frac{2}{3}} q_i^n, \end{aligned} \quad (6.4)$$

where  $(\bar{h}^\eta)_{i-\frac{1}{2}}^{n+1} := \bar{h}^\eta(h_i^+, h_{i+1}^-)$ , according to (3.25).

Finally, the scheme's time accuracy is improved by the use of the classical Heun's method (see [27]).

### 6.2. MOOD technique to recover the well-balance property

At this level, the designed MUSCL scheme is not well-balanced. Indeed, if  $W_i^n$  and  $W_{i+1}^n$  define a steady state, the reconstruction step (6.1) will provide reconstructed values  $W_i^+$  and  $W_{i+1}^-$  that no longer define a steady state. Thus, the scheme will not be able to exactly preserve the initial steady state, but will provide a second-order approximation. To restore this essential property, we use a MOOD-like technique (see [17] for an overview of such techniques, and [7] for more recent applications). We suggest to introduce a convex combination between the reconstructed state and the non-reconstructed one (see [31] for related work). As a consequence, we adopt the following reconstruction:

$$\begin{aligned} w_i^- &= (1 - \theta_i^n) w_i^n + \theta_i^n \left( w_i^n - \frac{\Delta x}{2} \sigma_i^n \right) = w_i^n - \frac{\Delta x}{2} \sigma_i^n \theta_i^n, \\ w_i^+ &= (1 - \theta_i^n) w_i^n + \theta_i^n \left( w_i^n + \frac{\Delta x}{2} \sigma_i^n \right) = w_i^n + \frac{\Delta x}{2} \sigma_i^n \theta_i^n, \end{aligned} \quad (6.5)$$

where  $0 \leq \theta_i^n \leq 1$  is the parameter of the convex combination. If  $\theta_i^n = 1$ , the full MUSCL scheme is recovered. If  $\theta_i^n = 0$ , the states are not reconstructed and the first-order well-balanced scheme is used.

Now, we propose a suitable process to define the parameter  $\theta_i^n$ . To that end, we first define

$$\Delta \psi_{i+\frac{1}{2}}^n = \frac{(q_{i+1}^n)^2}{h_{i+1}^n} - \frac{(q_i^n)^2}{h_i^n} + \frac{g}{2} ((h_{i+1}^n)^2 - (h_i^n)^2) - \Delta x \bar{S}_{i+\frac{1}{2}}^t - \Delta x \bar{S}_{i+\frac{1}{2}}^f,$$

where  $\bar{S}_{i+\frac{1}{2}}^t = \bar{S}^t(h_i^n, h_{i+1}^n, Z_i, Z_{i+1})$  and  $\bar{S}_{i+\frac{1}{2}}^f = \bar{S}^f(h_i^n, h_{i+1}^n, q_i, q_{i+1})$ , with  $\bar{S}^t$  and  $\bar{S}^f$  respectively defined by (4.7) and (4.8). Note that, from (4.3),  $\Delta \psi_{i+\frac{1}{2}}^n$  vanishes when  $W_i^n$  and  $W_{i+1}^n$  define a steady state. Thus, from this quantity, we define a function to evaluate the deviation with respect to the equilibrium, as follows:

$$\varphi_i^n = \left\| \begin{pmatrix} q_i^n - q_{i-1}^n \\ \Delta \psi_{i-\frac{1}{2}}^n \end{pmatrix} \right\|_2 + \left\| \begin{pmatrix} q_{i+1}^n - q_i^n \\ \Delta \psi_{i+\frac{1}{2}}^n \end{pmatrix} \right\|_2. \quad (6.6)$$

Since  $\Delta \psi_{i+\frac{1}{2}}^n$  vanishes when  $W_i^n$  and  $W_{i+1}^n$  define a steady state, this newly introduced quantity  $\varphi_i^n$  vanishes when  $W_{i-1}^n$ ,  $W_i^n$  and  $W_{i+1}^n$  define a steady state. Let  $M > m > 0$ . We define the parameter of the convex combination  $\theta_i^n$  as follows:

$$\theta_i^n = \begin{cases} 0 & \text{if } \varphi_i^n < m \Delta x \\ \frac{\varphi_i^n - m \Delta x}{M \Delta x - m \Delta x} & \text{if } m \Delta x \leq \varphi_i^n \leq M \Delta x \\ 1 & \text{if } \varphi_i^n > M \Delta x. \end{cases} \quad (6.7)$$

This definition of  $\theta_i^n$  enforces the use of the MUSCL scheme if the states are far from defining a steady state, i.e.  $\varphi_i^n$  is large enough; the first-order well-balanced scheme is used if the equilibrium error  $\varphi_i^n$  is small enough. In addition, the closer the states are to the equilibrium, the more the convex combination will favor the first-order well-balanced scheme.

## 7. Numerical experiments

This last section is devoted to numerical tests, in one and two dimensions. We start by recalling the different schemes we shall test:

- the *explicit* scheme is (3.6) – (4.9);
- the *implicit* scheme is (5.8) – (5.9) – (5.12) – (5.18);
- the *MUSCL* scheme is (6.2) – (6.3) – (6.4) – (6.5).

In order to assess the properties of these schemes, we present four sets of numerical experiments. The first set assesses the well-balance of the scheme, by considering steady states at rest and moving steady states with friction, over flat and non-flat topographies. Then, the goal of the second set is to study the influence of the parameters  $C$ ,  $m$  and  $M$  on the numerical solution. Afterwards, the focus of the third set is the simulation of wet and dry dam-breaks, over a possibly complicated topography, and with a nonzero Manning coefficient  $k$ . Finally, the fourth set of experiments concerns simulations in two dimensions.

In order to provide reference solutions when no exact solution is available, we elect to use the second-order hydrostatic reconstruction scheme (HR) presented in [1], equipped with the HLL flux and an implicit treatment of the friction source term. Note that this scheme preserves the steady states at rest, even with a nonzero friction contribution.

For the numerical experiments involving friction, a suitable value of the Manning coefficient  $k$  has to be chosen. For instance, the reader is referred to [16] (page 109), where multiple values of  $k$  are given for different types of channel beds. Nevertheless, we deliberately impose stronger Manning coefficients than in reality (up to 10 times). This choice is made to ensure that the friction source term is preponderant compared to the topography source term, in order to enhance the effects of the friction.

An important step in these numerical experiments is the choice of the constant  $C$ , introduced in (3.26) to ensure the consistency of the approximate topography source term, and the choice of the constants  $m$  and  $M$ , introduced in (6.7) to define a steady state detector for the MUSCL scheme. In this article, these constants are chosen heuristically; a deeper study of the stability and of the order of accuracy of the scheme could provide several bounds for these constants. For each set of experiments, we will provide an insight regarding the choice of  $C$ ,  $m$  and  $M$ .

The other constants are chosen according to Table 1.

Constant	Equation	Value
$g$	(1.1)	$g = 9.81 \text{ m.s}^{-2}$
$\varepsilon_\lambda$	(3.3)	$\varepsilon_\lambda = 10^{-10} \text{ m.s}^{-1}$
$\varepsilon$	(3.33)	$\varepsilon = 0 \text{ m}$

Table 1: Values of the constants within the numerical experiments.

Moreover, we will compute errors to make sure the schemes satisfy the required properties. The  $L^1$ ,  $L^2$  and  $L^\infty$  errors for a bounded function  $w$  are computed as follows:

$$L^1 : \frac{1}{N} \sum_{i=1}^N |w_i - w_i^{ex}| ; \quad L^2 : \sqrt{\frac{1}{N} \sum_{i=1}^N (w_i - w_i^{ex})^2} ; \quad L^\infty : \max_{1 \leq i \leq N} |w_i - w_i^{ex}|, \quad (7.1)$$

where  $w_i$  and  $w_i^{ex}$  are respectively the approximate and the exact solution at the point  $x_i$  and at the final physical time  $t_{end}$ , and  $N$  is the number of discretization cells. Finally, we recall that the CFL condition (3.5) gives the time step  $\Delta t$  at each iteration, as follows:

$$\Delta t \leq \frac{\Delta x}{2\Lambda}, \quad \text{where } \Lambda = \max_{i \in \mathbb{Z}} \left( -\lambda_{i+\frac{1}{2}}^L, \lambda_{i+\frac{1}{2}}^R \right).$$

### 7.1. Verification of the well-balance

In this first set of experiments, we assess the well-balance of the scheme, i.e. its ability to exactly preserve and capture steady states. Recall that steady states are given by the equation (1.5), which prescribes a uniform discharge over the space domain, denoted  $q_0$ . First, steady states for the friction source term only are studied, that is to say we impose  $q_0 \neq 0$  and  $\partial_x Z = 0$ . Then, we consider steady states for both friction and topography, which are either analytic solutions in specific cases or steady states obtained by approximately solving (1.5).

#### 7.1.1. Steady states for the friction source term

We first focus on the preservation of the friction-only steady states. Thus, we assume  $q_0 \neq 0$  and a flat topography, i.e.  $\partial_x Z = 0$ . The smooth steady states are then given according to (1.10), or equivalently (2.2). The space domain for these experiments is  $[0, 1]$ , and its discretization is made of 200 cells (i.e.  $\Delta x = 5 \cdot 10^{-3}$ ).

In Section 2, the water height for a smooth steady state was obtained by considering a zero of the nonlinear function  $\xi$  defined by (2.5). We consider the subcritical branch of the steady state obtained by assuming  $q_0 = -1$ ,  $x_0 = -\Delta x$  and  $h_0 = h_c$ , with  $h_c$  defined by (2.6). The Manning coefficient  $k$  is chosen equal to 1.

The first experiment concerns the preservation of this steady state. We take  $q(x) = q_0$  and the exact height as initial conditions, and we impose the exact solution at the points  $-\Delta x$  and  $1 + \Delta x$  as inhomogeneous Dirichlet boundary conditions. We compute the approximate solution until the final time  $t_{end} = 1s$ .

Moreover, the parameters are taken as follows:  $C = 10^{-3}$ ,  $m = 10^{-12}$  and  $M = 10^{-11}$ . Such a value for  $C$  was chosen to ensure that the term  $[h]_c$  present in the approximate source terms was small enough. The justification of the choice of the parameters  $m$  and  $M$  is the same as in the case of the steady states are rest.

The results are presented on [Figure 4](#), and the errors to the steady state are displayed in [Table 2](#).

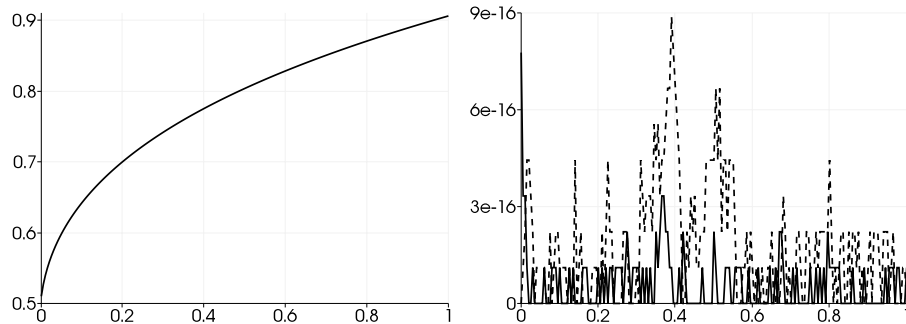


Figure 4: Left panel: initial height for the friction steady state. Right panel: height (solid line) and discharge (dashed line) errors to the steady state after 1s, with the explicit scheme.

	$h$			$q$		
	$L^1$	$L^2$	$L^\infty$	$L^1$	$L^2$	$L^\infty$
explicit	6.00e-17	1.09e-16	7.77e-16	1.81e-16	2.56e-16	8.88e-16
implicit	6.49e-17	1.15e-16	7.77e-16	2.91e-16	3.82e-16	9.99e-16
MUSCL	8.49e-17	1.33e-16	6.66e-16	2.42e-16	3.30e-16	8.88e-16

Table 2: Height and discharge errors with the three schemes for the friction steady state.

From [Figure 4](#) and [Table 2](#), we observe that the friction-only steady state is indeed preserved up to the machine precision by the three schemes, which is not the case when considering non-well-balanced schemes (see for instance [\[40\]](#) in the context of the shallow-water equations with topography only).

The second experiment is based on the previous steady state, but with a perturbation, as shown on [Figure 5](#). With  $h_{ex}$  the exact height, this perturbation is defined by choosing the initial water height as follows:

$$h(x) = \begin{cases} h_{ex}(x) + 0.2 & \text{if } x \in \left[\frac{3}{7}, \frac{4}{7}\right]; \\ h_{ex}(x) & \text{otherwise.} \end{cases}$$

The initial discharge is unperturbed, and taken equal to  $q_0 = -1$  throughout the domain. The boundary conditions consist in the unperturbed exact solution. We use the three schemes and 100 discretization cells for the numerical simulation on the domain  $[0, 1]$ .

Moreover, we take  $C = 10^{-3}$ ,  $m = 10^{-12}$  and  $M = 10^{-1}$ . For this perturbed initial data, we have raised the value of the parameter  $M$  so that the scheme exactly converges towards the unperturbed steady solution. Indeed, the well-balanced scheme starts being used as soon as the steady state deviation, defined



by (6.6), is lower than  $M\Delta x$ . Therefore, the value  $M = 10^{-1}$  ensures that the well-balanced scheme is used soon enough to allow the exact recovery of the unperturbed steady state.

The computations are carried out until the final time  $t_{end} = 9$ s. Indeed, such a final time allows the perturbation to be dissipated and a steady state to be reached. In fact, this steady state is the original, unperturbed steady state. The results of the explicit scheme are presented on Figure 5 and an error comparison with the unperturbed steady state is provided in Table 3.

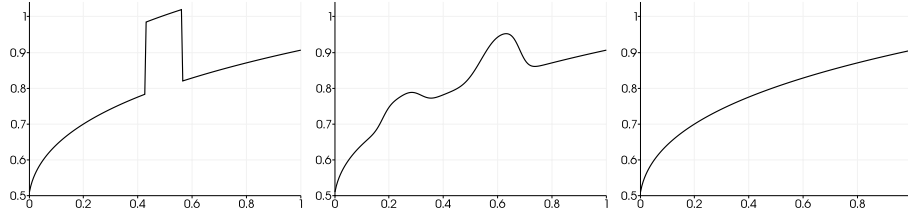


Figure 5: From left to right: water height for  $t = 0$ s,  $t = 0.06$ s and  $t = 9$ s, with the explicit scheme.

	$h$			$q$		
	$L^1$	$L^2$	$L^\infty$	$L^1$	$L^2$	$L^\infty$
explicit	7.42e-15	7.48e-15	8.55e-15	7.06e-15	8.32e-15	1.48e-14
implicit	7.06e-15	7.09e-15	8.10e-15	7.61e-15	8.81e-15	1.51e-14
MUSCL	8.40e-15	8.47e-15	1.04e-14	6.56e-15	7.52e-15	1.73e-14

Table 3: Height and discharge errors with the three schemes for the perturbed friction steady state.

Figure 5 shows that the perturbation is eventually dissipated and we recover the unperturbed steady state. This assertion is confirmed by the error analysis presented in Table 3, which shows that all three schemes recover the unperturbed steady state up to the machine precision.

*Remark 5.* We have carried out similar experiments with a supercritical branch instead of the subcritical branch, and we obtained the same results, i.e. the steady state solution was exactly recovered. For the sake of conciseness, we do not present them here.

### 7.1.2. Steady states with both friction and topography

We end the sanity check of the well-balance property with numerical experiments consisting in the preservation of steady states involving both topography and friction (for experiments with the sole topography source term, including the preservation of steady states at rest, the reader is referred to [40]). Hence, we take  $k \neq 0$ ,  $\partial_x Z \neq 0$ , and  $q_0 \neq 0$ . Steady states are therefore given by the full

equation (1.5). Recall that this equation cannot be rewritten under an algebraic form. Thus, to find a general steady state solution, we have to numerically solve the equation (1.5).

We begin by considering two very specific cases, where the height or the free surface are constant, in order to obtain an analytical solution of (1.5).

First, assume  $h(x) = h_0$ , i.e. the water height is uniform and equal to  $h_0$  in the whole space domain. Thus, (1.7) rewrites

$$gh_0\partial_x Z + kq_0|q_0|h_0^\eta = 0,$$

and we immediately obtain a relation characterizing the slope of the topography:

$$\partial_x Z = -\frac{kq_0|q_0|}{gh_0^{\eta+1}}. \quad (7.2)$$

This specific case is tested numerically by taking  $h_0 = q_0 = 1$ ,  $k = 10$ ,  $Z(0) = 0$  and the slope of  $Z$  given by (7.2). The space domain is  $[0, 1]$  and is discretized using 100 cells. The initial and boundary conditions are the exact solution. The computations are carried out with all three schemes, and we choose  $t_{end} = 1$ s. We take  $C = +\infty$ ,  $m = 10^{-12}$  and  $M = 10^{-11}$ , for reasons similar to the ones invoked for the first experiment with friction only. Table 4 shows that the topography and friction steady state with constant height is indeed exactly preserved by the three schemes.

	$h$			$q$		
	$L^1$	$L^2$	$L^\infty$	$L^1$	$L^2$	$L^\infty$
explicit	9.99e-18	3.68e-17	2.22e-16	9.99e-17	1.84e-16	6.66e-16
implicit	2.22e-17	5.21e-17	2.22e-16	9.99e-17	1.84e-16	6.66e-16
MUSCL	0	0	0	2.55e-17	6.75e-17	2.22e-16

Table 4: Height and discharge errors for the topography and friction steady state with constant height.

To build another exact solution of (1.5), we assume  $h + Z = H_0$  instead of  $h = h_0$ . We therefore have a constant free surface  $H_0$  over the whole space domain  $[0, 1]$ . Note that (1.7) rewrites for smooth solutions with positive water height as follows:

$$-\frac{q_0^2}{h^2}\partial_x h + gh\partial_x(h + Z) + \frac{kq_0|q_0|}{h^\eta} = 0.$$

Since the free surface  $h + Z$  is constant, the above relation rewrites

$$h^{\eta-2}\partial_x h = k\mu_0, \quad (7.3)$$

where  $\mu_0 = \text{sgn}(q_0)$ . For  $x \in [0, 1]$ , integrating (7.3) over  $[0, x]$  provides an expression of  $h$  that depends on the initial condition  $h_0 := h(0)$ , namely:

$$h^{\eta-1}(x) = h_0^{\eta-1} + (\eta - 1)k\mu_0 x. \quad (7.4)$$

Note that (7.4) defines an expression of the height that may become non-positive for some  $x$ . Thus, we choose the constants  $k$ ,  $q_0$  and  $h_0$  such that  $h$  is positive over the whole domain  $[0, 1]$ . In the simulation, we set  $k = q_0 = h_0 = 1$  and  $H_0 = 1$ . Equipped with the water height, we define the topography by setting  $Z(x) = H_0 - h(x)$ . We discretize the space domain with 100 cells and use the three schemes to carry out the simulation until the time  $t_{end} = 1$ s. We again take  $C = +\infty$ ,  $m = 10^{-12}$  and  $M = 10^{-11}$ . An error comparison of the three schemes is displayed in Table 5, which shows that the steady state is preserved up to the machine precision by the explicit, implicit and MUSCL schemes.

	$h$			$q$		
	$L^1$	$L^2$	$L^\infty$	$L^1$	$L^2$	$L^\infty$
explicit	4.23e-15	4.48e-15	6.66e-15	3.06e-14	3.07e-14	3.60e-14
implicit	1.31e-16	2.00e-16	6.66e-16	1.45e-15	1.83e-15	4.88e-15
MUSCL	8.88e-17	1.57e-16	4.44e-16	4.46e-16	6.81e-16	2.22e-15

Table 5: Height and discharge errors for the topography and friction steady state with constant free surface.

Finally, we derive a steady state for the shallow-water equations with topography and friction, without considering a constant height or free surface. Thus, we approximately solve the discretization (4.3) of the full equilibrium relation (1.7). First, we set  $k = 0.01$  and choose  $[0, 1]$  to be the space domain. Moreover, the topography is given by

$$Z(x) = \frac{1}{2} \frac{e^{\cos(4\pi x)} - e^{-1}}{e^1 - e^{-1}}. \quad (7.5)$$

We set  $q(x) = q_0$  throughout the domain. The equation (4.3) is then approximately solved using Newton's method, imposing  $h(0) = 0.3$ . This procedure allows us to define the water height over the whole domain. This steady state  ${}^t(h(x), q(x))$  is then chosen as the initial and boundary conditions for this experiment. We take 100 discretization cells, and we once again set  $C = +\infty$ ,  $m = 10^{-12}$  and  $M = 10^{-11}$ . The numerical simulation of this experiment is done using all three schemes, and runs until a final physical time  $t_{end} = 1$ s. The results of the explicit scheme are presented on Figure 6 and the errors to the steady state are displayed in Table 6.

Figure 6 shows that the explicit scheme exactly preserves this topography and friction steady state. Moreover, Table 6 shows that both other schemes also exactly preserve such a steady state.

The very last experiment in the assessment of the well-balance property focuses on a perturbation of the aforementioned steady state. We denote by  $h_{ex}(x)$  the water height of the previous topography and friction steady state.

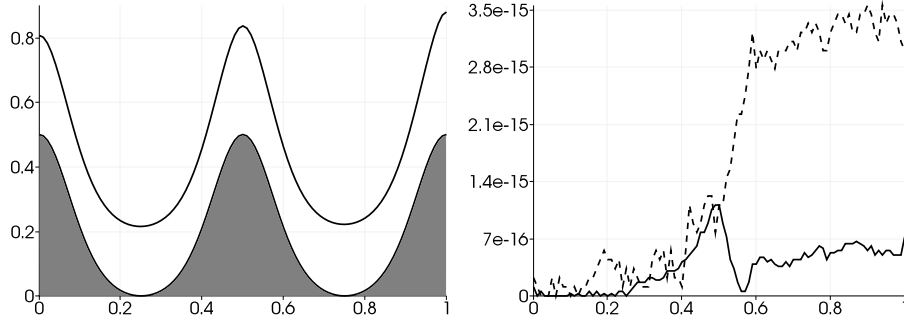


Figure 6: Left panel: initial height for the topography and friction steady state. Right panel: height (solid line) and discharge (dashed line) errors to the steady state with the explicit scheme.

	$h$			$q$		
	$L^1$	$L^2$	$L^\infty$	$L^1$	$L^2$	$L^\infty$
explicit	3.54e-16	4.49e-16	1.11e-15	1.66e-15	2.16e-15	3.55e-15
implicit	5.73e-16	8.99e-16	3.00e-15	5.07e-16	6.29e-16	1.55e-15
MUSCL	3.52e-17	5.40e-17	1.67e-16	3.75e-16	4.85e-16	1.33e-15

Table 6: Height and discharge errors with the three schemes for the topography and friction steady state.

The initial water height of this last experiment is then given by

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

and the discharge is defined by

$$q(x) = \begin{cases} q_0 + \frac{1}{2} & \text{if } x \in \left[\frac{2}{7}, \frac{3}{7}\right] \cup \left[\frac{4}{7}, \frac{5}{7}\right], \\ q_0 & \text{otherwise.} \end{cases}$$

The unperturbed steady state is prescribed as the boundary conditions. For this numerical experiment, the domain  $[0, 1]$  is discretized with 100 cells, and the simulation runs until the perturbation has been dissipated and the unperturbed steady state has been recovered. The final physical time we choose for these conditions to be met is  $t_{end} = 2s$ . Moreover, since we seek the capture of this perturbed steady state, we set the parameters as  $C = +\infty$ ,  $m = 10^{-13}$  and  $M = 0.05$ . The evolution of the perturbation with the explicit scheme is presented on [Figure 7](#). Then, in [Table 7](#), we present the errors to the original unperturbed steady state when the physical time is elapsed.

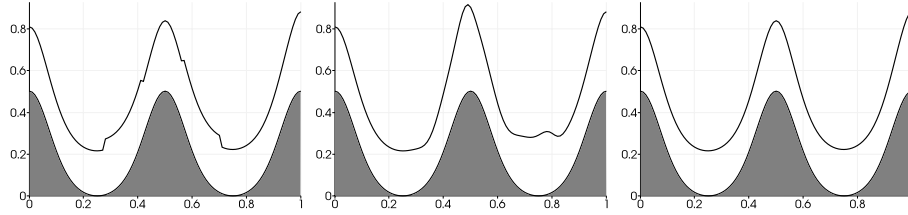


Figure 7: From left to right: water height for  $t = 0\text{s}$ ,  $t = 0.015\text{s}$  and  $t = 2\text{s}$ , with the explicit scheme.

	$h$			$q$		
	$L^1$	$L^2$	$L^\infty$	$L^1$	$L^2$	$L^\infty$
explicit	5.10e-16	7.09e-16	2.44e-15	2.20e-15	2.92e-15	5.33e-15
implicit	4.85e-16	7.97e-16	2.72e-15	9.73e-16	1.15e-15	2.78e-15
MUSCL	6.46e-16	9.40e-16	3.50e-15	1.71e-15	2.23e-15	4.22e-15

Table 7: Height and discharge errors with the three schemes for the perturbed topography and friction steady state.

Figure 7 and Table 7 show that the scheme indeed allows to recover the original unperturbed steady state. This experiment emphasizes the ability of the three schemes to exactly capture a steady state, even after a perturbation.

## 7.2. Influence of the parameters $C$ , $m$ and $M$

This second set of experiments is devoted to the study of the influence of the parameters present in the scheme:  $C$ ,  $m$  and  $M$ . Our goal is to emphasize the need for a relevant value of these parameters in numerical simulations.

We first perform three dam-break experiments whose goals are to show the dependence in  $C$  of the numerical solution, when considering the topography or the friction source term. Then, the dependence in  $m$  and  $M$ , related to the order of the numerical scheme, is studied.

### 7.2.1. Influence of the parameter $C$ for the topography

Recall that the approximate topography source term  $\bar{S}^t$ , defined by (4.7), depends on the parameter  $C$ . Namely, this parameter is used to ensure the consistency of the approximate source term  $\bar{S}^t$  with the actual source term  $S^t$ , by ensuring that the absolute value of the water height jump  $[h]$  is no larger than  $C\Delta x$ . The purpose of the first set of experiments is to highlight the influence of this parameter  $C$ .

#### Shock waves over a flat topography

The first experiment we consider concerns the propagation of shock waves over a flat topography. To that end, we consider the domain  $[0, 1]$  with a flat

topography (i.e.  $Z \equiv 0$ ), homogeneous Neumann boundary conditions, and we take the following dam-break initial data:

$$h(0, x) = 1 \quad \text{and} \quad q(0, x) = \begin{cases} 7.5 & \text{if } x < 0.5, \\ -7 & \text{if } x \geq 0.5. \end{cases}$$

Recall that the cutoff involving  $C$  had been introduced to ensure the consistency of  $\bar{S}^t$ , especially for flat topographies. As a consequence, for this experiment, the value of  $C$  should be instrumental in getting the correct shock waves and an accurate approximation of the intermediate state.

To carry out this experiment, we set a vanishing friction contribution (i.e.  $k = 0$ ), we use 250 discretization cells, and we take the final time  $t_{end} = 0.1s$ . This experiment is performed for  $C = 10$  and  $C = 1000$ . The results of the implicit scheme are presented on [Figure 8](#), where we display the exact solution (see for instance [\[35, 45\]](#) for more details) as well as the numerical approximation.

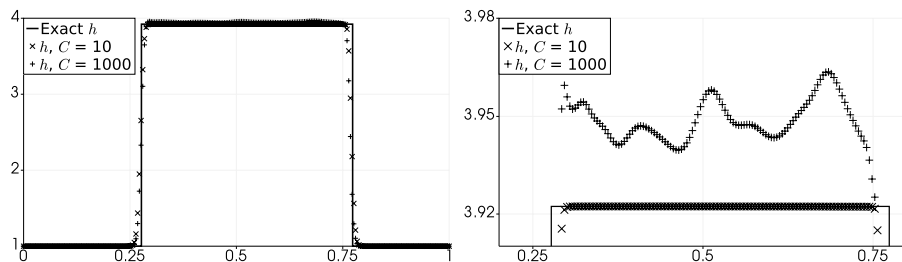


Figure 8: Dam-break creating two shock waves over a flat bottom. Left panel: whole domain depicted at  $t = 0.1s$ . Right panel: zoom on the intermediate state of the dam-break problem.

The left panel of [Figure 8](#) shows that, for both values of  $C$ , the approximate shock waves are located at a consistent position and seem to have the correct amplitude. However, on the right panel of [Figure 8](#), we note, on the one hand, that the intermediate state obtained with  $C = 1000$  presents spurious oscillations, whose amplitude does not decrease when  $\Delta x$  decreases. On the other hand, with  $C = 10$ , the implicit scheme provides a good approximation of the intermediate state.

#### *Dam-break on an emerging bottom*

The second dam-break experiment we perform in order to study the influence of the parameter  $C$  presents an emerging bottom. To that end, we modify the flow at rest with emerging bottom experiment from [\[24\]](#), to add a wave perturbing the water at rest. On the space domain  $[0, 15]$ , we consider the topography function  $Z(x) = (0.2 - 0.05(x - 10)^2)_+$  and the following dam-break initial conditions:

$$h(0, x) + Z(x) = \begin{cases} 0.2 & \text{if } x < 5, \\ 0.15 & \text{if } x \geq 5, \end{cases} \quad \text{and} \quad q(0, x) = 0.$$

We prescribe homogeneous Neumann boundary conditions, and we consider 4000 discretization cells. We once again take a vanishing friction contribution, i.e.  $k = 0$ . The initial condition, as well as the reference solution, computed with the HR scheme at  $t_{end} = 5\text{s}$ , are displayed on [Figure 9](#).

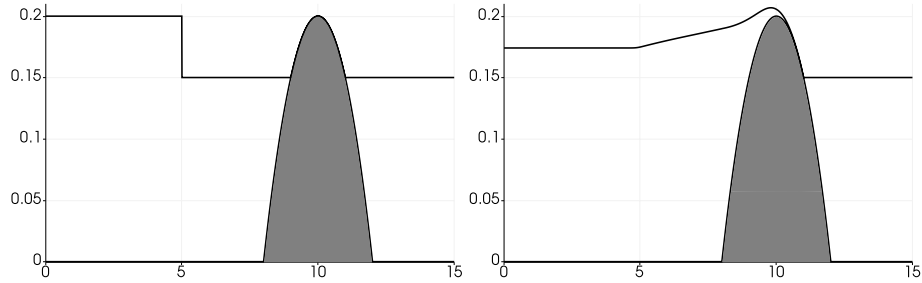


Figure 9: Dam-break on an emerging bottom. Left panel: Initial free surface. Right panel: Reference free surface obtained with the HR scheme. On both panels, the gray area is the topography.

In order to study the influence of the parameter  $C$ , we consider the implicit scheme, and we carry out the simulation with  $C = 1$  and  $C = 10$ . The results are displayed on [Figure 10](#). Indeed, with  $C = 10$ , large spurious oscillations appear, a wave is reflected to the left of the domain, and an inconsistent dry area appears on the bump. However, with  $C = 1$ , we only get small oscillations near  $x = 9.75$ ; the amplitude of these oscillations decreases as  $\Delta x$  decreases.

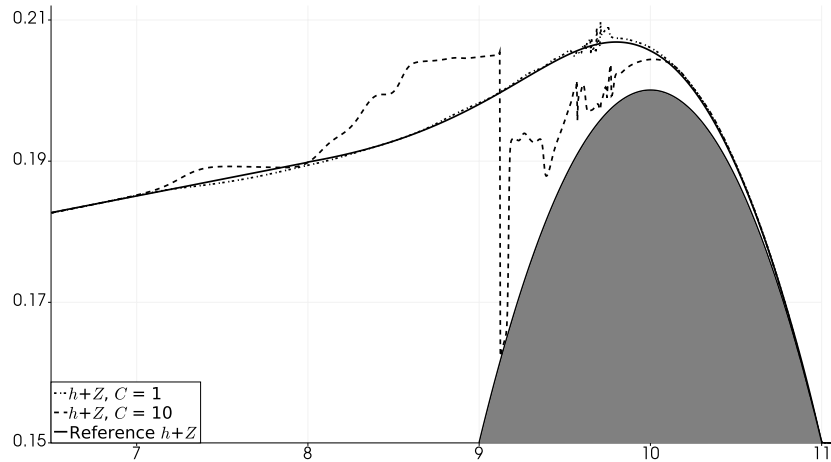


Figure 10: Dam-break on an emerging bottom: zoomed comparison between the HR scheme, the implicit scheme with  $C = 1$ , and the implicit scheme with  $C = 10$ . The gray area is still the topography.

On [Figure 10](#), we once again observe that the consistency is ensured by the

cutoff procedure. Indeed, with  $C = 10$ , large spurious oscillations appear, a wave is reflected to the left of the domain, and an inconsistent dry area appears on the bump. However, with  $C = 1$ , we only get small oscillations near  $x = 9.75$ ; the amplitude of these oscillations correctly decreases as  $\Delta x$  decreases.

### 7.2.2. Influence of the parameter $C$ for the friction

Next, we turn to the study of the influence of the parameter  $C$  on the approximate friction source term  $\bar{S}^f = -k\bar{q}|\bar{q}|\bar{h}^{-\eta}$ , defined by (3.27). Recall that the purpose of this parameter is to ensure the consistency of  $\bar{S}^f$  with the friction source term  $S^f$ , especially when dealing with discontinuities and a small Manning coefficient  $k$ .

In order to highlight the relevance of this parameter, we consider the following numerical experiment. We consider a dam-break problem on a flat topography (i.e.  $Z \equiv 0$ ) with a very small Manning coefficient (we take  $k = 10^{-9}$ ). As a consequence, the first term of the expression (3.27) of  $\bar{S}^f$  is negligible, and the presence of the cutoff (3.26) makes its second term consistent. The initial conditions of the dam-break problem are taken as follows:

$$h(0, x) = \begin{cases} 20 & \text{if } x < 0.5, \\ 30 & \text{if } x \geq 0.5. \end{cases} \quad \text{and} \quad q(0, x) = \begin{cases} 250 & \text{if } x < 0.5, \\ -200 & \text{if } x \geq 0.5. \end{cases}$$

and the space domain is  $[0, 1]$ , with homogeneous Neumann boundary conditions. Equipped with such initial conditions, we expect two large shock waves to propagate in the domain.

This experiment is carried out with the explicit scheme, until the final physical time  $t_{end} = 0.0175\text{s}$ . Note that we know the exact solution  ${}^t(h^{exact}, q^{exact})$  to this dam-break problem for  $k = 0$  (see for instance [35]). As a consequence, in order to quantify the error introduced by the inconsistent term and hopefully corrected by the use of  $C$ , we consider the following error:

$$\Delta q := \left| \sum_{i \in [1, N]} (q_i^{n_{end}} \Delta x) - \int_0^1 q^{exact}(t_{end}, x) dx \right|,$$

where  $N$  is the number of cells and where  $t^{n_{end}} = t_{end}$ .

To that end, we compute the errors  $\Delta q$ , and the associated orders of accuracy, for the following three configurations:

- (a)  $C = 0$ ;
- (b)  $C = 1$ ;
- (c)  $C = +\infty$ .

The errors and orders of accuracy are collected in [Table 8](#).

- (a) For  $C = 0$ , we get a consistent scheme. The values of the error presented in [Table 8](#) are of the same order of magnitude as  $k$ , which was to be expected since the approximate solution is compared to the exact solution obtained with  $k = 0$ . This is verified by taking, for instance,  $k = 10^{-11}$ : in this case,  $\Delta q \simeq 1.2 \times 10^{-13}$  for all  $N$ . However, if a steady state solution were considered, the value  $C = 0$  would not allow its exact preservation.



N	$\Delta q, C = 0$		$\Delta q, C = 1$		$\Delta q, C = +\infty$	
200	2.44e-10	—	1.64e-04	—	2.21e-01	—
400	2.41e-10	—	8.03e-05	1.03	2.12e-01	0.06
800	2.40e-10	—	3.98e-05	1.01	2.08e-01	0.03
1600	2.39e-10	—	1.98e-05	1.01	2.06e-01	0.01

Table 8: Dam-break experiment with large shocks: influence of the parameter  $C$ .

- (b) For  $C = 1$ , since the error is of the order of magnitude of  $\Delta x$ , we also get a consistent scheme. Indeed, in this case, the formerly inconsistent term of (3.27) becomes a perturbation of size  $\mathcal{O}(\Delta x)$ .
- (c) For  $C = +\infty$ , the scheme is inconsistent: indeed, the error  $\Delta q$  does decrease with  $\Delta x$ , but the approximate solution does not converge towards the correct solution. However, this value of  $C$  allows steady state solutions to be exactly preserved.

This experiment therefore highlights the necessity of the cutoff introduced in (3.27) in order to recover the consistency of the scheme when considering discontinuous solutions and a negligible friction contribution, while ensuring that the scheme is still well-balanced.

### 7.2.3. Influence of the parameters $m$ and $M$

The goal of this next experiment is to emphasize the dependence of the numerical solution in  $m$  and  $M$ . Recall that these parameters were introduced during the MOOD procedure to ensure the well-balance of the scheme. Namely, they were used to define, in (6.7), the parameter  $\theta_i^n$  of the convex combination between the non-reconstructed states and the reconstructed ones.

In order to study the influence of these parameters, we consider an experiment devoted to computing the order of accuracy of the scheme. To set up this experiment, we compute a steady solution with both topography and friction contributions. We set  $k = 0.01$  and a topography given by (7.5) over the space domain  $[0.4, 0.6]$ . We take the steady discharge  $q(x) = q_0 = 1$  and we approximately solve (4.3) with  $h(0) = 1$  to finally obtain the steady water height  $h(x)$  on the whole space domain. This steady state  ${}^t(h(x), q(x))$  is then chosen as the initial and boundary conditions for this experiment.

However, this steady state is exactly preserved by the well-balanced scheme. In order to study the order of accuracy of the scheme, we introduce the following smooth perturbation, of amplitude  $10^{-6}$ , of the steady solution on  $[0.45, 0.55]$ :

$$ash_p(x) = h(x) + 10^{-6} \times \omega \left( \frac{2}{0.05}(x - 0.5) \right),$$

where we have set

$$\omega(z) = \begin{cases} \left( \frac{2 - |z|}{2} \right)^4 (1 + 2|z|) & \text{if } |z| < 2, \\ 0 & \text{otherwise.} \end{cases}$$

The initial data of the experiment is the perturbed steady state  ${}^t(h_p(x), q(x))$ . Note that such initial data is smooth (of class  $\mathcal{C}^4$ ) on  $[0.4, 0.6]$ , and it is a steady solution on  $[0.4, 0.45] \cup [0.55, 0.6]$ . As a consequence, on the one hand, a well-balanced scheme will exactly preserve its steady part, and provide a first-order approximation of the unsteady part. On the other hand, a second-order scheme will provide a second-order approximation of the solution everywhere. Therefore, the advantage of using the aforementioned MOOD procedure lies in ensuring both the exact preservation of the steady part and the second-order approximation of the unsteady part.

This experiment is carried out until the final physical time  $t_{end} = 10^{-3}$ . Such a final time has been chosen to ensure that the perturbation does not reach the boundaries of the domain. We use the MUSCL scheme for several values of  $m$  and  $M$ , in order to study the order of accuracy for each pair  $(m, M)$  under consideration. To that end, we first compute a reference solution using 20480 cells. Then, to obtain the order of accuracy, we begin by using (7.1) to compute the  $L^1$ ,  $L^2$  and  $L^\infty$  errors. With  $e_N$  representing either of these errors for  $N$  cells, the order of accuracy  $\rho$  is computed as follows:

$$\rho = \frac{\ln(e_N) - \ln(e_{2N})}{\ln 2}. \quad (7.6)$$

Equipped with this definition, we compute the  $L^1$ ,  $L^2$  and  $L^\infty$  discharge errors and the associated orders of accuracy for the following pairs  $(m, M)$ :

- (a)  $m = 10^{-12}$  and  $M = 10^{-11}$ ;
- (b)  $m = 10^{-9}$  and  $M = 10^{-6}$ ;
- (c)  $m = 10^{-5}$  and  $M = 10^{-2}$ .

The results are collected in [Table 9](#), [Table 10](#) and [Table 11](#), as well as [Figure 11](#).

N	$L^1$		$L^2$		$L^\infty$	
80	3.34e-07	—	3.57e-07	—	4.60e-07	—
160	8.39e-08	1.99	8.97e-08	1.99	1.16e-07	1.99
320	2.11e-08	1.99	2.25e-08	1.99	2.90e-08	1.99
640	5.27e-09	2.00	5.63e-09	2.00	7.26e-09	2.00
1280	1.31e-09	2.01	1.39e-09	2.01	1.79e-09	2.02

Table 9: Discharge errors and order of accuracy for the pair (a).

From [Table 9](#), [Table 10](#), [Table 11](#) and [Figure 11](#), we note that the scheme is second-order accurate when using the pairs (a) and (b), and first-order accurate for the pair (c). These observations can be explained as follows: for the first pair, the second-order reconstruction is used almost everywhere, which leads to the scheme being second-order accurate. For the third pair, the second-order reconstruction is never used, and the scheme is first-order accurate. It is interesting to remark that, for the coarsest grid, the first-order well-balanced scheme is more accurate than the second-order scheme. The second pair leads to a balance between the second-order piecewise linear reconstruction and the

N	$L^1$		$L^2$		$L^\infty$	
80	2.73e-07	—	3.39e-07	—	5.68e-07	—
160	3.62e-08	2.91	4.55e-08	2.90	8.00e-08	2.83
320	5.13e-09	2.82	6.38e-09	2.84	1.10e-08	2.86
640	9.39e-10	2.45	1.13e-09	2.50	1.84e-09	2.58
1280	2.08e-10	2.18	2.43e-10	2.21	4.22e-10	2.12

Table 10: Discharge errors and order of accuracy for the pair (b).

N	$L^1$		$L^2$		$L^\infty$	
80	2.00e-07	—	3.33e-07	—	7.74e-07	—
160	9.98e-08	1.01	1.67e-07	1.00	3.92e-07	0.98
320	4.96e-08	1.01	8.31e-08	1.01	1.96e-07	1.00
640	2.45e-08	1.02	4.10e-08	1.02	9.68e-08	1.02
1280	1.18e-08	1.05	1.99e-08	1.05	4.69e-08	1.05

Table 11: Discharge errors and order of accuracy for the pair (c).

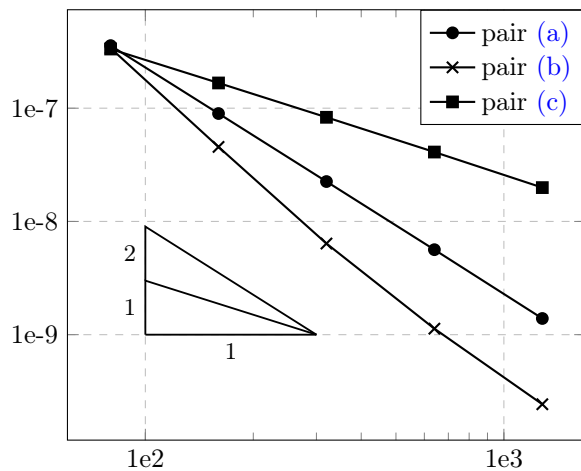


Figure 11: Error plots in  $L^2$ -norm for the three  $(m, M)$  pairs.

first-order piecewise constant values. This balance ensures that the second-order scheme is used where the solution is unsteady and that the well-balanced scheme is used where the solution is steady, thus leading to a scheme that is more accurate than in the other two cases.

This experiment emphasizes that the choice of  $m$  and  $M$  is of crucial importance in order to ensure a good quality of the numerical approximation.

### 7.3. Dam-break experiments

Now, we focus on several dam-break experiments, whose goal is to provide a validation of the scheme on unsteady cases. We consider a dam located at  $x = x_D$  that instantly breaks at  $t = 0$ s, thus releasing the water it was holding. Such a dam-break experiment has initial data  $W_L = {}^t(h_L(x), 0)$  if  $x \leq x_D$  and  $W_R = {}^t(h_R(x), 0)$  if  $x > x_D$ . The topography function  $Z$  and the Manning coefficient  $k$  are different in each experiment.

In addition, the parameters  $m$  and  $M$  are chosen to ensure the preservation of the situations of water at rest located away from the waves, while still retaining a more accurate approximation within the waves. More details on such water at rest situations are given in the next paragraph. Finally, the parameter  $C$  has been chosen heuristically, and it is reported in each experiment.

#### 7.3.1. Wet dam-break

We first carry out a wet dam-break experiment. Here, the topography is given by  $Z(x) = \cos^2(2\pi x)/2$  and the Manning coefficient is  $k = 1$ . The space domain is  $[0, 1]$  and the dam is located at  $x_D = 0.5$ . We prescribe homogeneous Neumann boundary conditions at both boundaries. The initial data is  $h_L(x) + Z(x) = 2$  and  $h_R(x) + Z(x) = 1$ . The results, obtained using 100 discretization cells and  $t_{end} = 0.05$ s, are displayed on [Figure 12](#). For this simulation, we choose  $C = 10$ ,  $m = 10^{-6}$  and  $M = 10^{-1}$ . The goal of this experiment is to compare the three schemes to a reference solution, obtained using the HR scheme with  $10^4$  discretization cells.

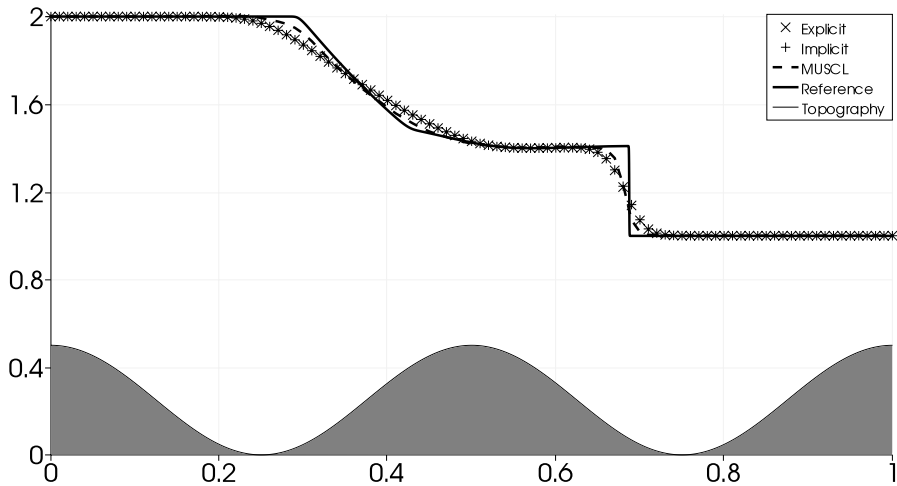


Figure 12: Free surface observed at the final physical time with the three schemes, for the wet dam-break experiment.

The three schemes yield a correct approximation of the reference solution. The explicit and the implicit schemes yield very close solutions, while the MUSCL scheme provides a better approximation of the reference solution.

Note that we are dealing with a lake at rest configuration in the regions far from the waves. Indeed, the free surface unaffected by the rarefaction wave or the shock wave remains unperturbed. This means that  $h(x) + Z(x) = 2$  and  $q(x) = 0$  for all  $x$  before the head of the rarefaction wave, and  $h(x) + Z(x) = 1$  and  $q(x) = 0$  for all  $x$  after the shock wave.

Such a property is satisfied by the explicit and implicit schemes. The MUSCL scheme, equipped with the well-balance correction (6.5), also exactly preserves this configuration. In addition, the MUSCL scheme provides a better approximation of the rarefaction wave and the shock wave.

Therefore, such an experiment highlights the interest of using a well-balanced scheme for such simulations, even if the whole domain does not involve a steady state. Moreover, it showcases the relevance of the convex combination procedure involved in the MUSCL scheme, to preserve steady states while still ensuring a better approximation of the exact solution.

In order to emphasize even more the importance of using a well-balanced scheme in such situations, we compare the implicit scheme with a non-well-balanced one. The non-well-balanced scheme we use is the HLL one (see [30]), coupled with a classical splitting treatment of the topography and friction source terms (see [10] for instance). The discharge obtained by both these schemes is compared on Figure 13 with the discharge of the reference solution.

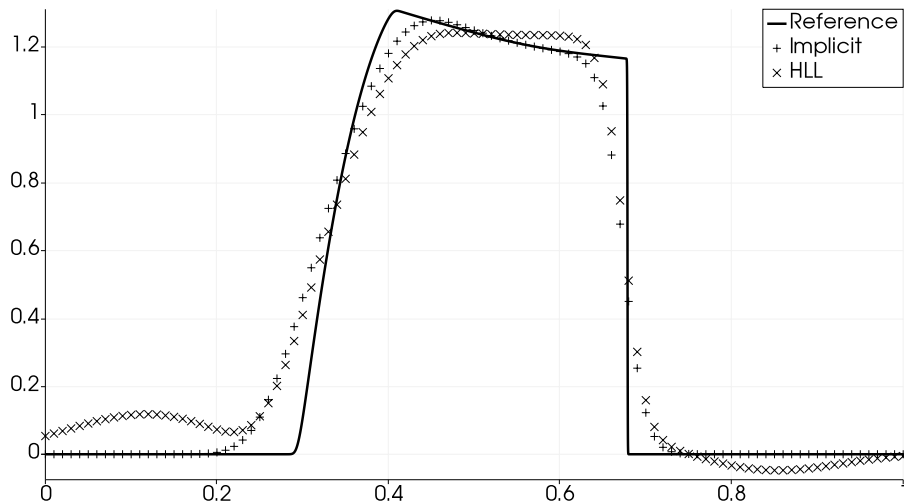


Figure 13: Discharge for the wet dam-break experiment. Results of the implicit scheme and the HLL scheme with 100 discretization cells, compared with a reference solution computed with the MUSCL scheme using 10000 discretization cells.

On Figure 13, we observe that the implicit scheme exactly preserves the lake at rest configurations ahead of the waves. Indeed, the discharge vanishes in those areas. However, the HLL scheme does not preserve this steady state, and the discharge shows perturbations near the boundaries of the domain. In

addition, we note that the intermediate state between the waves is correctly approximated by the implicit scheme, while the HLL scheme provides a poor approximation, which proves the importance of using a well-balanced scheme for such simulations.

### 7.3.2. Dry dam-break over a flat topography: explicit scheme disqualification

The goal of this experiment is to show the spurious oscillations induced by the explicit scheme in the presence of the friction, and the improvement obtained by using the semi-implication of the friction in the implicit scheme.

The topography function is  $Z(x) = 0$ , and we take  $k = 5$ . The initial data is  $h_L(x) = 1.5$  and  $h_R(x) = 0$ , and the space domain is  $[-1, 1]$ , for a dam located at  $x_D = 0$ . Hence, the wet/dry front is located within the interval  $[0, 1]$ . [Figure 14](#) displays the discharge computed by the explicit and the implicit schemes on this interval. To obtain these results, we used 200 cells, the final physical time  $t_{end} = 0.03s$ , and we took  $C = 1$ . Finally, we prescribe homogeneous Neumann boundary conditions on both boundaries, and we compute the reference solutions using the HR scheme with  $2 \times 10^4$  cells.

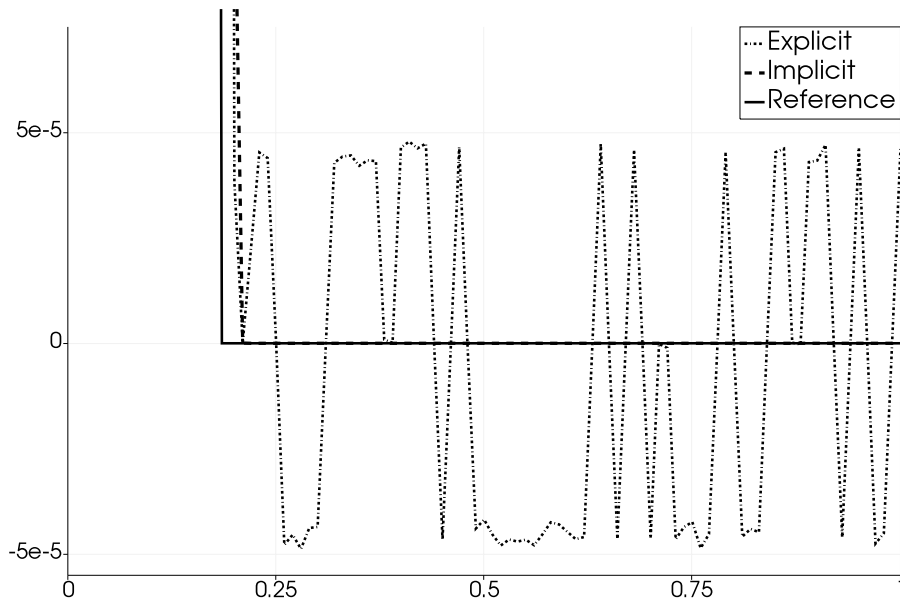


Figure 14: Dry dam-break: instabilities induced by the explicit scheme. The discharge is represented for the explicit scheme (dotted line), the implicit scheme (dashed line) and the reference solution (solid line). Oscillations in the discharge produced by the explicit scheme are clearly visible.

From [Figure 14](#), on the one hand, we note that the discharge obtained with the explicit scheme presents several non-physical oscillations. According to the reference solution, it should stay equal to zero. On the other hand, we remark that no oscillations appear when using the implicit scheme. As a consequence,

the explicit scheme is not suitable to perform the simulation of a dry area using the same CFL condition (3.5) as the implicit scheme. In the remainder of this section, the implicit scheme will be used for numerical experiments as soon they involve dry areas.

### 7.3.3. Dry dam-break over a non-flat topography

Now, we consider a dry dam-break. The topography is still given by  $Z(x) = \cos^2(2\pi x)/2$  and the Manning coefficient is now  $k = 10$ . The initial data is  $h_L(x) + Z(x) = 2$  and  $h_R(x) = 0$ . The space domain is  $[0, 1]$  and the dam is located at  $x_D = 0.5$ . Moreover, the boundaries at  $x = 0$  and  $x = 1$  are equipped with homogeneous Neumann boundary conditions. We use 100 discretization cells and display the results on Figure 15 at time  $t_{end} = 0.05s$ . We choose the constants  $C = 7.5$ ,  $m = 10^{-6}$  and  $M = 10^{-1}$ . We use the implicit and MUSCL schemes. Indeed, the explicit scheme present instabilities when dealing with dry/wet transitions. We also provide a reference solution for comparison purposes, obtained with the HR scheme using  $10^4$  discretization cells.

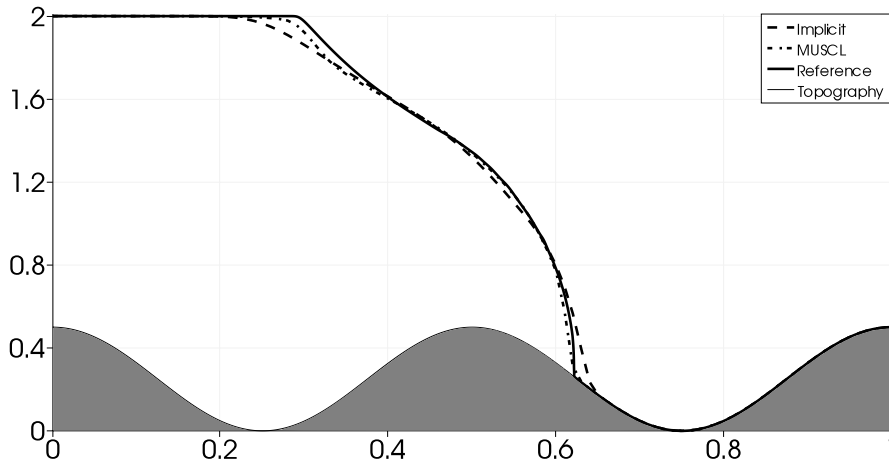


Figure 15: Free surface observed at the final physical time with the implicit and MUSCL schemes, for the dry dam-break experiment.

The reference solution is once again correctly approximated by the two schemes, and the MUSCL scheme provides a better approximation of the solution than the implicit scheme. We reiterate the same remark concerning the exact preservation of the lake at rest before the head of the rarefaction wave.

### 7.3.4. Dry dam-break with two bumps

This last dry dam-break experiment concerns a more complicated topography, with two bumps. The space domain is  $[0, 5]$  and we choose to use  $10^4$  discretization cells with the MUSCL scheme to have a relevant simulation. The

two boundaries at  $x = 0\text{m}$  and  $x = 5\text{m}$  are solid walls. The topography is defined by

$$Z(x) = \frac{1}{2} \left( 1 - \frac{(x - 5/2)^2}{1/25} \right)_+ + 2 \left( 1 - \frac{(x - 4)^2}{1/25} \right)_+,$$

and it indeed consists in two quadratic bumps, a smaller one followed by a larger one. The dam is located at  $x_D = 0.7\text{m}$ , breaks at  $t = 0\text{s}$ , and contains an initial water height  $h_L = 6\text{m}$ . The domain  $x > x_D$  contains no water, i.e.  $h_R = 0$ . We take a Manning coefficient  $k$  equal to 1. Figure 16 shows that the initial water height is significantly larger than the bumps, in order to impact them with a high enough velocity. We choose  $C = 0.1$ ,  $m = 10^{-5}$  and  $M = 10^{-1}$ .

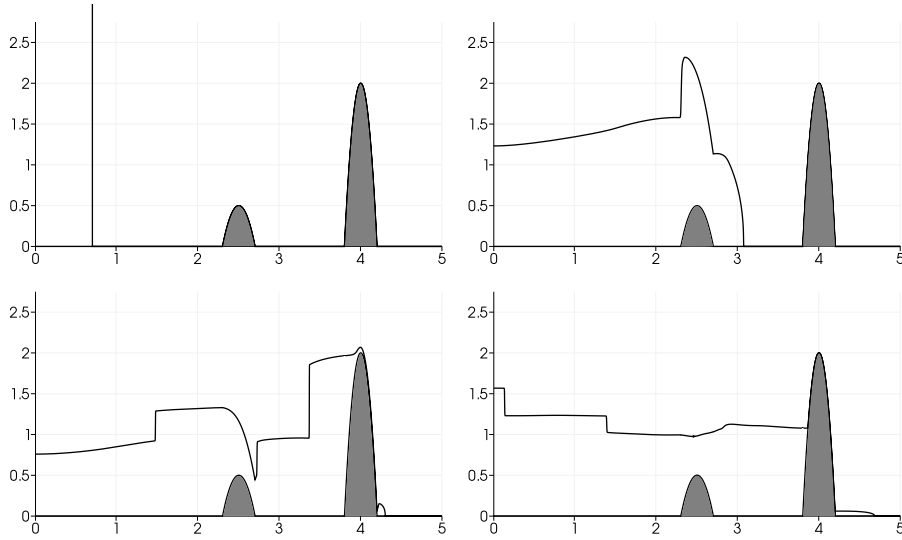


Figure 16: Free surface for the double bump test case at different times: from left to right and top to bottom, the solution is observed at times 0s, 0.38s, 0.74s and 1.70s.

#### 7.4. 2D simulations

This subsection is devoted to experiments in two space dimensions. In this context of two-dimensional geometries, the shallow-water equations with topography and Manning friction read as follows:

$$\begin{cases} \partial_t h + \partial_x p + \partial_y q & = 0, \\ \partial_t p + \partial_x \left( \frac{p^2}{h} + \frac{1}{2} g h^2 \right) + \partial_y \left( \frac{p q}{h} \right) & = -g h \partial_x Z - k p D h^{-\eta}, \\ \partial_t q + \partial_x \left( \frac{p q}{h} \right) + \partial_y \left( \frac{q^2}{h} + \frac{1}{2} g h^2 \right) & = -g h \partial_y Z - k q D h^{-\eta}, \end{cases} \quad (7.7)$$



where  $p$  is the discharge in the  $x$  direction,  $q$  is the discharge in the  $y$  direction, and  $D$  is the Euclidean norm of the discharge vector, i.e.  $D = \sqrt{p^2 + q^2}$ . We rewrite (7.7) under the following condensed form:

$$\partial_t W + \partial_x f(W) + \partial_y g(W) = s(W),$$

where the definitions of  $W$ ,  $f$ ,  $g$  and  $s$  can easily be inferred from (7.7).

The numerical approximation involves a uniform Cartesian grid, made of square cells with sides  $\Delta x = \Delta y$ . For the 2D simulations, we will consider 2D variants of the implicit and MUSCL scheme, by applying the 1D strategy detailed in Section 5 and Section 6 in both directions, horizontal and vertical, of each square cell. This process involves exactly solving the initial value problem (5.10), related to the contribution of the friction only. However, this initial value problem is modified in two dimensions. Indeed, we have to solve

$$\begin{cases} \frac{dh}{dt} = 0, \\ \frac{dp}{dt} = -k(h_{i,j}^{n+1})^{-\eta} p \sqrt{p^2 + q^2}, \\ \frac{dq}{dt} = -k(h_{i,j}^{n+1})^{-\eta} q \sqrt{p^2 + q^2}, \end{cases} \quad \text{with initial data} \quad \begin{cases} h(0) = h_{i,j}^{n+\frac{2}{3}}, \\ p(0) = p_{i,j}^{n+\frac{2}{3}}, \\ q(0) = q_{i,j}^{n+\frac{2}{3}}. \end{cases} \quad (7.8)$$

We immediately see that  $h_{i,j}^{n+1} = h_{i,j}^{n+\frac{2}{3}}$ . For the sake of simplicity in the notations, we set  $\tau = -k(h_{i,j}^{n+1})^{-\eta} < 0$ ,  $p_n = p_{i,j}^{n+\frac{2}{3}}$  and  $q_n = q_{i,j}^{n+\frac{2}{3}}$ . Thus, (7.8) rewrites as the following initial value problem:

$$\begin{cases} p' = \tau p \sqrt{p^2 + q^2}, \\ q' = \tau q \sqrt{p^2 + q^2}, \end{cases} \quad \text{with initial data} \quad \begin{cases} p(0) = p_n, \\ q(0) = q_n. \end{cases}$$

The exact solution to this system is:

$$p(t) = \frac{p_n h_{i,j}^{n+1}}{h_{i,j}^{n+1} + kt \sqrt{p_n^2 + q_n^2}} \quad \text{and} \quad q(t) = \frac{q_n h_{i,j}^{n+1}}{h_{i,j}^{n+1} + kt \sqrt{p_n^2 + q_n^2}}.$$

Therefore, the friction step of the implicit procedure yields the following updated discharges:

$$p_{i,j}^{n+1} = \frac{p_{i,j}^{n+\frac{2}{3}} (h_{i,j}^{n+1})^\eta}{(h_{i,j}^{n+1})^\eta + k \Delta t \sqrt{(p_{i,j}^{n+\frac{2}{3}})^2 + (q_{i,j}^{n+\frac{2}{3}})^2}},$$

$$q_{i,j}^{n+1} = \frac{q_{i,j}^{n+\frac{2}{3}} (h_{i,j}^{n+1})^\eta}{(h_{i,j}^{n+1})^\eta + k \Delta t \sqrt{(p_{i,j}^{n+\frac{2}{3}})^2 + (q_{i,j}^{n+\frac{2}{3}})^2}}.$$

The quantity  $(h_{i,j}^{n+1})^\eta$  is then replaced with the average given by (5.18), taken in the  $x$  direction for  $p_{i,j}^{n+1}$  and in the  $y$  direction for  $q_{i,j}^{n+1}$ .

Such a 2D scheme will be able to exactly preserve the 1D steady states taken in the  $x$  direction or the  $y$  direction. Moreover, the splitting procedure allows simulating dry/wet interfaces. We use this 2D scheme for the simulations of three relevant situations. First, we focus on assessing the order of accuracy of the implicit and MUSCL schemes. Then, we turn to the simulation of a dry dam-break with two bumps, which is a 2D version of the experiment presented [Section 7.3.4](#). The next experiment is a partial dam-break, and the last experiment represents an urban topography.

#### 7.4.1. Order of accuracy assessment

The purpose of this first experiment is to assess the order of accuracy of the proposed schemes. Namely, the orders of accuracy of the implicit and the MUSCL schemes will be computed. The implicit scheme should be first-order accurate, while the MUSCL scheme should be second-order accurate.

To address this issue, we need a smooth exact solution of the 2D shallow-water equations with topography and friction [\(7.7\)](#). We elect to use a 2D steady state solution, which will not be exactly preserved by the well-balanced scheme, and that will allow computing the order of accuracy.

On the space domain  $[-0.3, 0.3] \times [0.4, 1]$ , we consider the following topography function:

$$Z(x, y) = \frac{2k\|r\| - 1}{2g\|r\|^2},$$

where  $r = {}^t(x, y)$ . Then, to define a steady solution of [\(7.7\)](#), the water height and discharge are given by:

$$h(x, y, t) = 1, \quad p(x, y, t) = \frac{x}{\|r\|^2}, \quad q(x, y, t) = \frac{y}{\|r\|^2}.$$

For the experiment, we take  $k = 10$ , and we prescribe the exact solution as both the initial and the boundary conditions. The final physical time is  $t_{end} = 0.1\text{s}$ , and we set  $C = +\infty$ ,  $m = 0.05$  and  $M = 1$ . The parameter  $C$  can here be infinite since the exact solution is very smooth and the topography is non-flat, thus making the cutoff procedure irrelevant. In addition, the other two parameters are chosen to ensure that the MUSCL scheme is second-order accurate everywhere.

Recall that the order of accuracy is defined, in the one-dimensional case, by [\(7.6\)](#). In our 2D case, the order of accuracy  $\rho$  is computed as follows:

$$\rho = \frac{\ln(e_N) - \ln(e_{4N})}{\ln 2}.$$

Equipped with this definition, we compute the  $L^1$  and  $L^\infty$  errors and the associated orders of accuracy of both schemes. These results are collected in [Table 12](#) and [Table 13](#) for the implicit scheme, and in [Table 14](#) and [Table 15](#) for the MUSCL scheme.

From [Table 12](#), [Table 13](#), [Table 14](#) and [Table 15](#), we conclude that both schemes have the expected order of accuracy. Indeed, the implicit scheme is first-order accurate, while the MUSCL scheme is second-order accurate.

N	$h$		$p$		$q$	
900	1.33e-02	—	2.54e-02	—	3.50e-02	—
3600	6.82e-03	0.97	1.36e-02	0.90	1.83e-02	0.94
14400	3.44e-03	0.99	7.02e-03	0.95	9.36e-03	0.96
57600	1.73e-03	0.99	3.57e-03	0.97	4.75e-03	0.98

Table 12:  $L^1$  errors for the friction and topography 2D steady state using the implicit scheme.

N	$h$		$p$		$q$	
900	4.01e-02	—	7.04e-02	—	7.02e-02	—
3600	2.27e-02	0.82	3.89e-02	0.86	3.65e-02	0.94
14400	1.28e-02	0.83	2.00e-02	0.96	1.85e-02	0.98
57600	7.03e-03	0.86	1.01e-02	0.99	9.22e-03	1.00

Table 13:  $L^\infty$  errors for the friction and topography 2D steady state using the implicit scheme.

N	$h$		$p$		$q$	
900	6.89e-04	—	1.43e-03	—	1.04e-03	—
3600	1.91e-04	1.85	3.90e-04	1.87	2.95e-04	1.81
14400	5.11e-05	1.90	1.03e-04	1.92	8.05e-05	1.87
57600	1.33e-05	1.94	2.67e-05	1.95	2.12e-05	1.92

Table 14:  $L^1$  errors for the friction and topography 2D steady state using the MUSCL scheme.

N	$h$		$p$		$q$	
900	2.38e-03	—	4.13e-03	—	2.23e-03	—
3600	8.05e-04	1.56	1.14e-03	1.86	7.58e-04	1.56
14400	2.49e-04	1.69	3.04e-04	1.91	2.44e-04	1.63
57600	7.36e-05	1.76	7.89e-05	1.94	7.25e-05	1.75

Table 15:  $L^\infty$  errors for the friction and topography 2D steady state using the MUSCL scheme.

#### 7.4.2. 2D dry dam-break with two bumps

The first 2D experiment is a dry dam-break with a topography presenting two bumps. This experiment is the 2D analogue of the 1D experiment presented in [Section 7.3.4](#). It is heavily inspired from an experiment presented in [\[8\]](#), which did not include the friction source term. The Manning coefficient is  $k = 0.1$ ,

and the topography function is given by

$$Z(x, y) = \frac{1}{2} \left( 1 - 25 \left( \left( x - \frac{5}{2} \right)^2 + \left( y - \frac{1}{2} \right)^2 \right) \right)_+ + 2 \left( 1 - 25 \left( (x - 4)^2 + \left( y - \frac{1}{2} \right)^2 \right) \right)_+.$$

The space domain is  $[0, 5] \times [0, 1]$ . The initial discharge is zero in both directions, i.e.  $p(x, y, 0) = q(x, y, 0) = 0$ , and the initial water height is given by

$$h(x, y, 0) = \begin{cases} 6 & \text{if } x < 0.7, \\ 0 & \text{otherwise.} \end{cases}$$

In addition, we prescribe wall boundary conditions, that is to say we impose  $p(0, y, t) = p(5, y, t) = 0$  and  $q(x, 0, t) = q(x, 1, t) = 0$  for all  $x, y$  and  $t$ , and homogeneous Neumann boundary conditions for the other variables.

The simulation runs until a physical time  $t_{end} = 1.35s$  with the MUSCL scheme, using  $C = 1$ ,  $m = 10^{-5}$  and  $M = 25$ . The choice of these constants is motivated by the same reasons as invoked in [Section 7.3.4](#). We take 288000 discretization cells (1200 in the  $x$  direction and 240 in the  $y$  direction). The results are presented on [Figure 17](#) and [Figure 18](#).

This experiment has been carried out to make sure that the numerical scheme still behaves correctly in two space dimensions and in the presence of dry/ wet transition. We recover a numerical solution involving the friction source term, which can be compared to the numerical solution without friction presented in [\[8\]](#). In addition, this 2D experiment is similar to the 1D double bump experiment we presented on [Figure 16](#). Indeed, the behavior of the water before it comes into contact with the first bump should be the same in both experiments. As expected, we obtain similar results in 1D and 2D.

#### 7.4.3. Partial dam-break

Next, we focus on a partial dam-break (see for instance [\[42, 18\]](#)). This experiment concerns a dam that has partially broken, leaving a corridor where the water flows. We consider the space domain  $[-100, 100] \times [-100, 100]$ . The dam is located at the middle of the domain, in the  $x$  direction, and is 10m wide and 100m high. It breaks in the middle, over a length of 80m. Thus, we take the following topography function:

$$Z(x, y) = \begin{cases} 1 & \text{if } x \leq -5, \\ 0 & \text{if } x \leq 5, \\ 0.1(5 - x) & \text{if } -5 < x < 5 \text{ and } -40 < y < 40, \\ 100 & \text{if } -5 < x < 5 \text{ and } y \in [-100, -40] \cup [40, 100]. \end{cases}$$

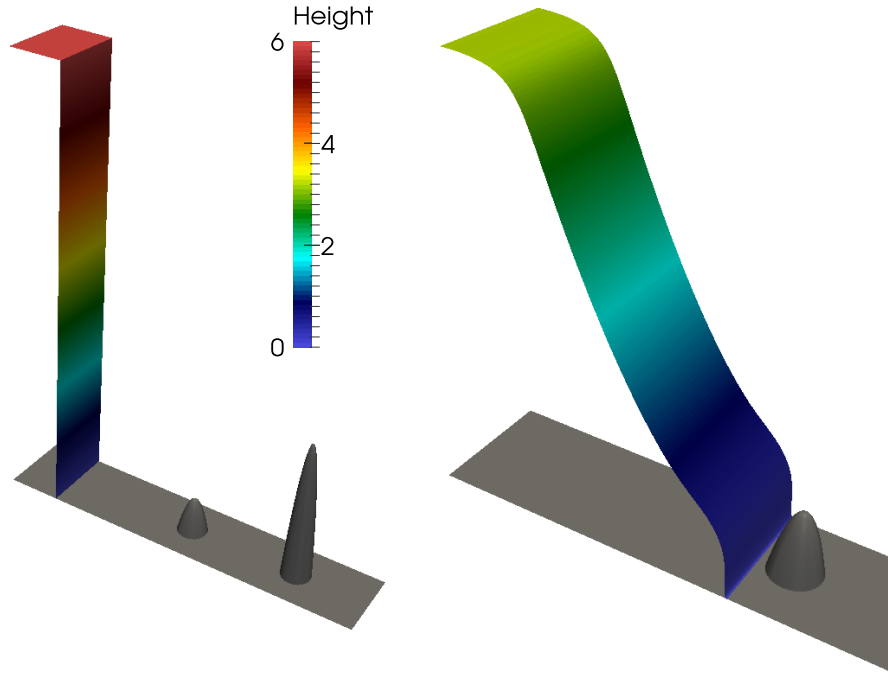


Figure 17: Left panel: initial condition of the 2D dam-break over a double bump experiment. Note that the same color scale for the water height is used in Figure 17 and Figure 18, and the solid gray color represents the topography. Right panel: approximate solution at  $t = 0.15$ s, just before the water hits the first bump. Note the shape of the front of the water, due to the nonzero bottom friction.

As in the previous dam-break experiment, we take  $p(x, y, 0) = 0$  and  $q(x, y, 0) = 0$ . Moreover, we define the initial water height by:

$$h(x, y, 0) = \begin{cases} 10 & \text{if } x < -5, \\ 5 & \text{otherwise.} \end{cases}$$

From such initial data, we present two simulations. The first one is the case without friction, i.e.  $k = 0$ , and for the second one we choose a nonzero Manning coefficient  $k = 2$ . In both cases, homogeneous Neumann boundary conditions are prescribed on all the boundaries. We choose  $t_{end} = 7$ s, and we set  $C = 0.5$ ,  $m = 10^{-10}$  and  $M = 0.5$ .

The justification of the parameter choice is similar to the one for the previous dam-break experiments, with one difference. Since the current experiment does not involve any water at rest with a non-flat topography, it is relevant to use the MUSCL scheme as much as possible, as evidenced by the relatively low value of  $m$ .

The numerical simulation of this experiment is carried out with the implicit and the MUSCL schemes on 202500 cells (450 in each direction). The results are

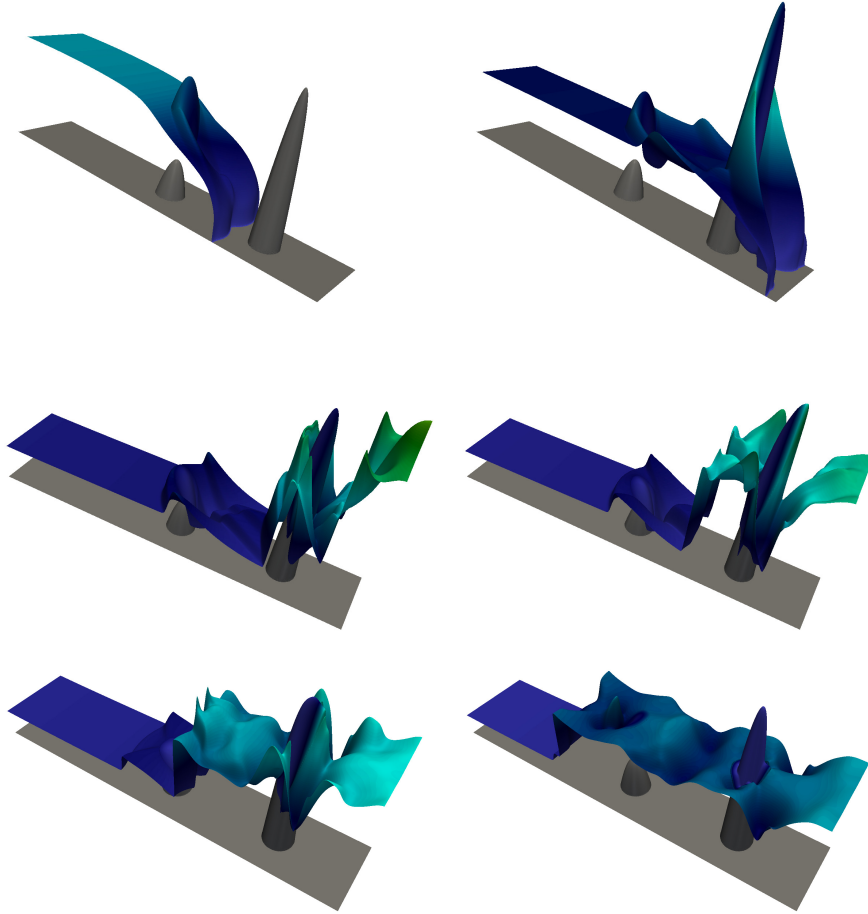


Figure 18: From left to right and top to bottom: approximate solution of the 2D dam-break over a double bump experiment, displayed at times  $t = 0.3\text{s}$ ,  $t = 0.45\text{s}$ ,  $t = 0.75\text{s}$ ,  $t = 0.9\text{s}$ ,  $t = 1.05\text{s}$  and  $t = 1.35\text{s}$ .

displayed on [Figure 19](#) without friction, and on [Figure 20](#) with friction, where the display of the free surface is clipped when  $Z > 10$ , in order to have a clearer visualization of the numerical results.

[Figure 19](#) shows that, when no friction is present, vortices appear at both tips of the dam. However, as evidenced by [Figure 20](#), these vortices no longer exist when the friction is nonzero. Both behaviors are exhibited with the implicit and MUSCL schemes. In addition, the tip of the left shock wave has traveled further on [Figure 19](#) than on [Figure 20](#). This behavior is consistent with the fact that the friction tends to slow down the movement of the water.

Now, to conclude the analysis of this experiment, we introduce  $(x_T, y_T)$  and  $(x_B, y_B)$ . They are, respectively, the points where the top and bottom vortices

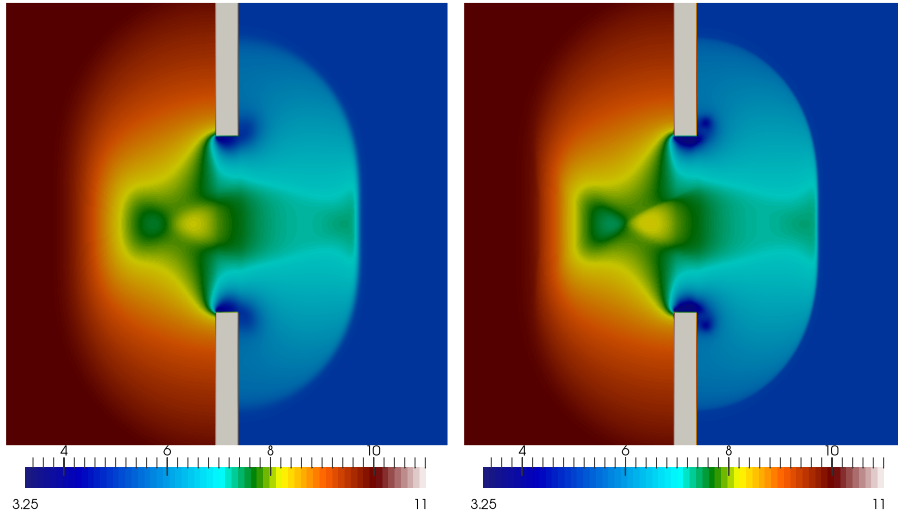


Figure 19: Approximate free surface for the partial dam-break experiment without friction. Left panel: implicit scheme; right panel: MUSCL scheme.

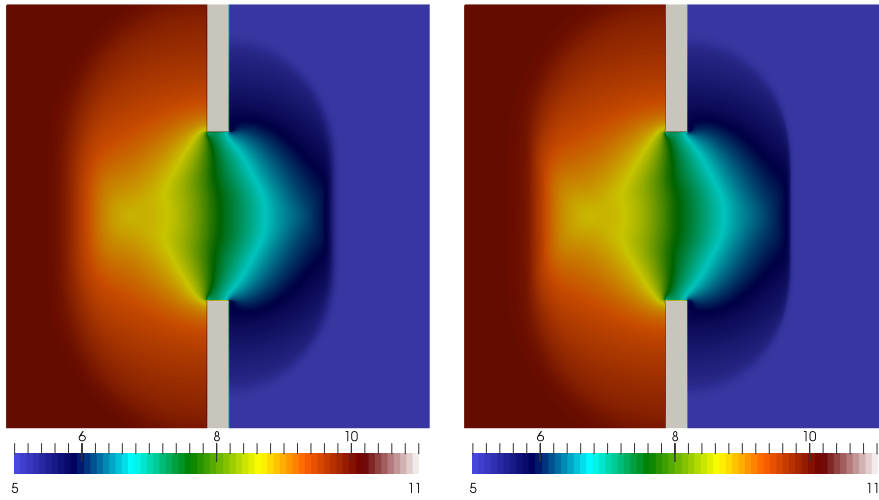


Figure 20: Approximate free surface for the partial dam-break experiment with nonzero friction. Left panel: implicit scheme; right panel: MUSCL scheme.

are the deepest when  $k = 0$ . In [Table 16](#), we present the height at these two points, obtained with the implicit and MUSCL schemes, and considering  $k = 0$  or  $k = 2$ . In addition, we display in this table the approximate position  $x_S$  of the left shock wave in the four cases under consideration, where it crosses the line  $y = 0$ .

	$h_T$	$h_B$	$x_S$
implicit, $k = 0$	2.81	2.81	59.3
implicit, $k = 2$	6.08	6.08	53.9
MUSCL, $k = 0$	2.33	2.30	59.8
MUSCL, $k = 2$	5.86	5.86	52.5

Table 16: Height at the positions of the two vortices and position of the left shock wave. We have set  $h_T = h(x_T, y_T, t_{end})$  and  $h_B = h(x_B, y_B, t_{end})$ .

#### 7.4.4. Urban topography

The last 2D experiment is a simulation of a city being hit by a wave. We consider the space domain  $[0, 1000] \times [0, 1000]$ . The topography consists in an upwards slope leading to a flat surface, upon which buildings are placed. Disregarding the buildings, the bottom has the following topography:

$$Z(x, y) = \begin{cases} x/50 & \text{if } x < 500, \\ 10 & \text{otherwise.} \end{cases}$$

The 100 meters high buildings occupy the flat part of the topography, i.e. buildings are only present for  $x > 500$ . [Figure 21](#) displays the shapes and the positions of the buildings, for a uniform Cartesian mesh of  $10^6$  cells (1000 in each direction).

The initial conditions are  $W(x, y, 0) = 0$  for all  $x$  and  $y$  in the space domain. Indeed, the boundary conditions help create the flood and the wave that hits the city. We prescribe homogeneous Neumann boundary conditions for each boundary of the domain, except the left boundary, where a time-dependent boundary condition the  $x$ -discharge  $p$  is applied, as follows:

$$\begin{cases} p(0, y, t) = 15 & \text{if } t < 350, \\ \partial_x p(0, y, t) = 0 & \text{otherwise.} \end{cases} \quad (7.9)$$

Such a boundary condition creates water that fills the sloping part of the topography and creates a wave that hits the city. To illustrate the situation, [Figure 21](#) shows the free surface at time  $t = 300$ s, when the first buildings are about to be hit by the wave.

We consider a nonzero Manning coefficient  $k = 1$  and a final time  $t_{end} = 850$ s. The simulation is run using the MUSCL scheme, and we take  $C = 10^{-2}$ ,  $m = 10^{-5}$  and  $M = 1$  for the same reasons as invoked before.

The left panel of [Figure 21](#) shows the wave created by the Dirichlet boundary condition arriving in the city. Because of the friction, this wave presents a rather steep front. On the left panel of [Figure 22](#), the wave has hit the first buildings located at the south of the city. Note that the space between the first two columns of buildings is still dry. Also note that, as per (7.9), the boundary condition imposed on the  $x$ -discharge  $q_x$  is now a homogeneous Neumann boundary condition, and no more water is injected into the domain. The right



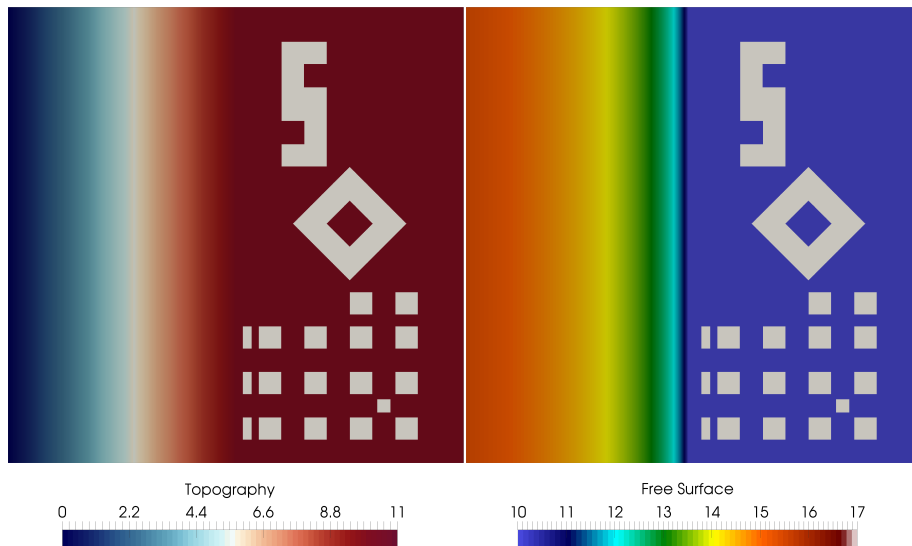


Figure 21: Wave on an urban topography simulation. Left panel: topography of the city. The buildings are actually 100 meters high, and are represented in white in this figure. One can see the upwards slope on the left, leading to the city itself. Right panel: free surface at  $t = 300$ s. The wave is present to the left of the figure. Note that the same free surface color scale will be used in the next figures.

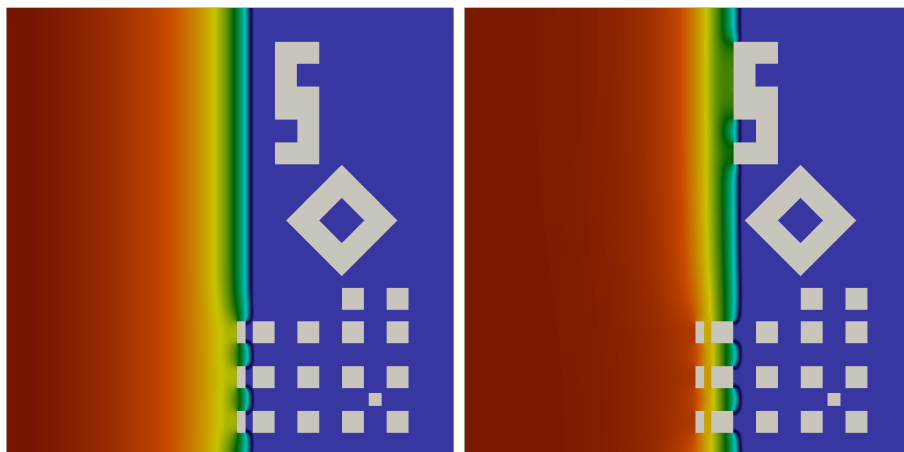


Figure 22: Free surface for the urban topography simulation at  $t = 355$ s (left panel) and  $t = 410$ s (right panel).

panel of [Figure 22](#) displays the wave about to hit the square building located at the middle of the city. As expected, between the southern buildings, the wave is slowed down.

On the left panel of [Figure 23](#), the wave has reflected on the southwestern

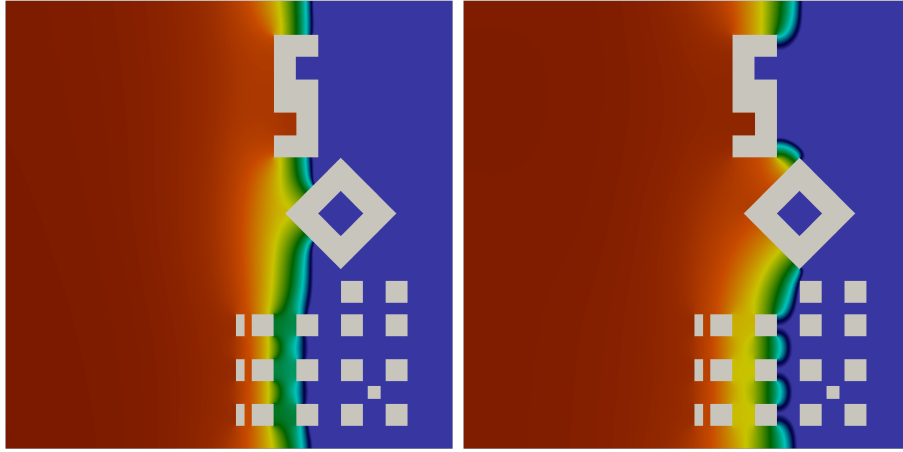


Figure 23: Free surface for the urban topography simulation at  $t = 465\text{s}$  (left panel) and  $t = 520\text{s}$  (right panel).

side of the square building, and it has thus moved faster towards the buildings to the south. On the right panel of [Figure 23](#), the waves reflected from the square building are moving south and north. Moreover, the back of the “S”-shaped building will soon be flooded.

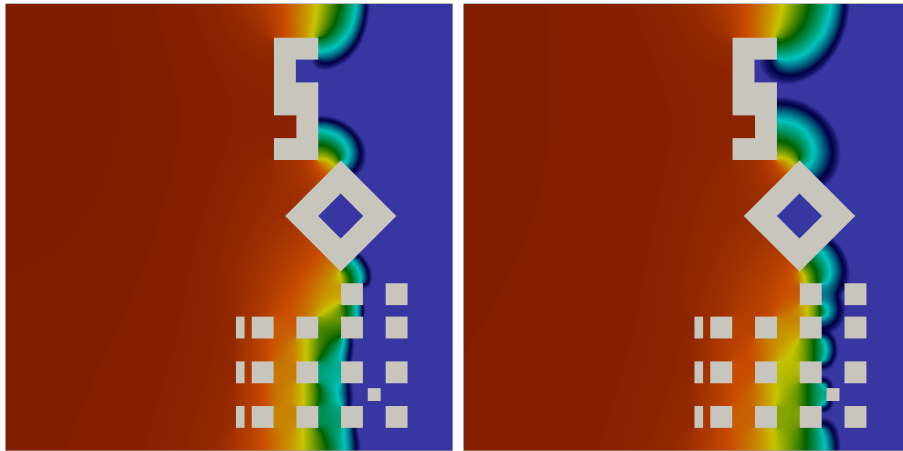


Figure 24: Free surface for the urban topography simulation at  $t = 575\text{s}$  (left panel) and  $t = 630\text{s}$  (right panel).

The flooding of the back of the “S”-shaped building is happening on [Figure 24](#), with only a small area still dry on the right panel. In addition, on the right panel of [Figure 24](#), the wave has almost hit the small square building on the bottom right of the city.

[Figure 25](#) and [Figure 26](#) display the final phases of the flooding of the city.

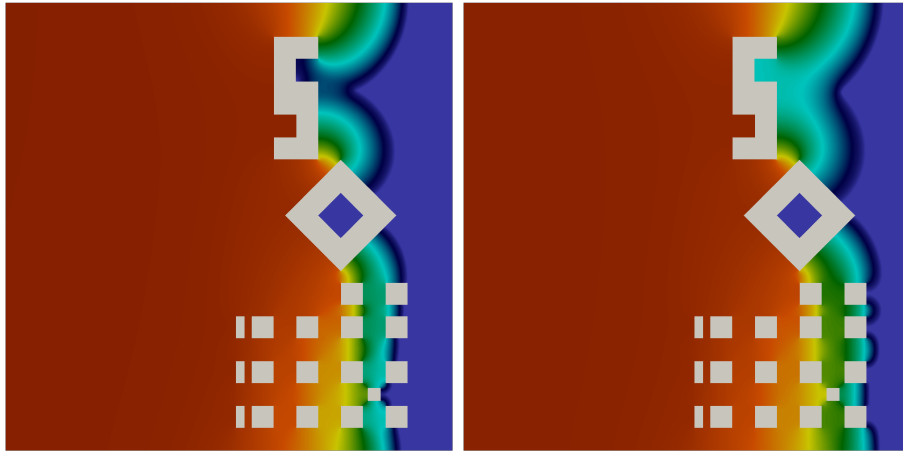


Figure 25: Free surface for the urban topography simulation at  $t = 685$ s (left panel) and  $t = 740$ s (right panel).

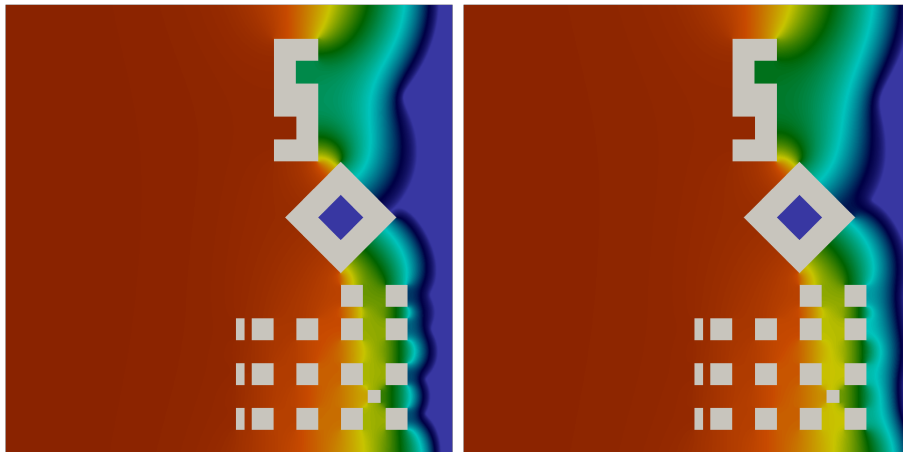


Figure 26: Free surface for the urban topography simulation at  $t = 795$ s (left panel) and  $t = 850$ s (right panel).

Note that the southern buildings are mostly uniformly flooded and that the inner courtyard of the square building is still dry. Moreover, the water at the back of the “S”-shaped building is less deep than at other points of the same vertical line.

## 8. Conclusion

In this work, we first studied smooth steady state solutions for the shallow-water equations with friction only. Then, we have proposed a Godunov-type scheme for the same equations that is well-balanced. This scheme has then been

extended to add the topography contribution as well as the positivity preservation. Afterwards, we have proposed a semi-implicit scheme to deal with dry/wet interfaces, followed by a well-balanced second-order MUSCL extension. Finally, we have proposed several one-dimensional and two-dimensional numerical experiments to assess the efficiency and robustness of the explicit, semi-implicit and MUSCL scheme.

One issue that remains is the division by  $\alpha^f$  occurring when computing the intermediate water heights (see [Remark 4](#)). Further work could be undertaken in order to try and provide a satisfactory answer to this issue.

### Acknowledgments

C. Berthon, F. Foucher and V. Michel-Dansac would like to thank the ANR-12-IS01-0004-01 GEONUM for financial support. S. Clain acknowledges the *Fundação para a Ciência e a Tecnologia* for the funding of projects FCT-ANR/MAT-NAN/0122/2012 and UID/MAT/00013/2013.

The authors acknowledge the anonymous referees for very relevant remarks which greatly improved the scheme.

### References

- [1] E. Audusse, F. Bouchut, M.-O. Bristeau, R. Klein, and B. Perthame. A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows. *SIAM J. Sci. Comput.*, 25(6):2050–2065, 2004.
- [2] E. Audusse, C. Chalons, and P. Ung. A simple well-balanced and positive numerical scheme for the shallow-water system. *Commun. Math. Sci.*, 13(5):1317–1332, 2015.
- [3] A. Bermudez and M. E. Vazquez. Upwind methods for hyperbolic conservation laws with source terms. *Comput. & Fluids*, 23(8):1049–1071, 1994.
- [4] C. Berthon and C. Chalons. A fully well-balanced, positive and entropy-satisfying Godunov-type method for the shallow-water equations. *Math. Comp.*, 85(299):1281–1307, 2016.
- [5] C. Berthon, C. Chalons, S. Cornet, and G. Sperone. Fully well-balanced, positive and simple approximate Riemann solver for shallow water equations. *Bull. Braz. Math. Soc. (N.S.)*, 47(1):117–130, 2016.
- [6] C. Berthon, A. Crestetto, and F. Foucher. A Well-Balanced Finite Volume Scheme for a Mixed Hyperbolic/Parabolic System to Model Chemotaxis. *J. Sci. Comput.*, 67(2):618–643, 2016.
- [7] C. Berthon and V. Desveaux. An entropy preserving MOOD scheme for the Euler equations. *Int. J. Finite Vol.*, 11, 2014.

- [8] C. Berthon and F. Foucher. Efficient well-balanced hydrostatic upwind schemes for shallow-water equations. *J. Comput. Phys.*, 231(15):4993–5015, 2012.
- [9] C. Berthon, F. Marche, and R. Turpault. An efficient scheme on wet/dry transitions for shallow water equations with friction. *Comput. & Fluids*, 48:192–201, 2011.
- [10] F. Bouchut. *Nonlinear stability of finite volume methods for hyperbolic conservation laws and well-balanced schemes for sources*. Frontiers in Mathematics. Birkhäuser Verlag, Basel, 2004.
- [11] M. J. Castro, A. Pardo Milanés, and C. Parés. Well-balanced numerical schemes based on a generalized hydrostatic reconstruction technique. *Math. Models Methods Appl. Sci.*, 17(12):2055–2113, 2007.
- [12] M. J. Castro Díaz, J. A. López-García, and C. Parés. High order exactly well-balanced numerical methods for shallow water systems. *J. Comput. Phys.*, 246:242–264, 2013.
- [13] C. Chalons, F. Coquel, E. Godlewski, P.-A. Raviart, and N. Seguin. Godunov-type schemes for hyperbolic systems with parameter-dependent source. The case of Euler system with friction. *Math. Models Methods Appl. Sci.*, 20(11):2109–2166, 2010.
- [14] A. Chertock, S. Cui, A. Kurganov, and T. Wu. Well-balanced positivity preserving central-upwind scheme for the shallow water system with friction terms. *Internat. J. Numer. Methods Fluids*, 78(6):355–383, 2015.
- [15] A. Chinnayya, A.-Y. LeRoux, and N. Seguin. A well-balanced numerical scheme for the approximation of the shallow-water equations with topography: the resonance phenomenon. *Int. J. Finite Vol.*, 1(1):33, 2004.
- [16] V. T. Chow. *Open-channel hydraulics*. McGraw-Hill civil engineering series. McGraw-Hill, 1959.
- [17] S. Clain, S. Diot, and R. Loubère. A high-order finite volume method for systems of conservation laws—Multi-dimensional Optimal Order Detection (MOOD). *J. Comput. Phys.*, 230(10):4028–4050, 2011.
- [18] S. Clain and J. Figueiredo. The MOOD method for the non-conservative shallow-water system. working paper or preprint, oct 2014.
- [19] O. Delestre, C. Lucas, P.-A. Ksinant, F. Darboux, C. Laguerre, T.-N.-T. Vo, F. James, and S. Cordier. SWASHES: a compilation of shallow water analytic solutions for hydraulic and environmental studies. *Internat. J. Numer. Methods Fluids*, 72(3):269–300, 2013.

- [20] V. Desveaux, M. Zenk, C. Berthon, and C. Klingenberg. A well-balanced scheme for the Euler equation with a gravitational potential. In *Finite volumes for complex applications. VII. Methods and theoretical aspects*, volume 77 of *Springer Proc. Math. Stat.*, pages 217–226. Springer, Cham, 2014.
- [21] E. D. Fernández-Nieto, D. Bresch, and J. Monnier. A consistent intermediate wave speed for a well-balanced HLLC solver. *C. R. Math. Acad. Sci. Paris*, 346(13-14):795–800, 2008.
- [22] U. S. Fjordholm, S. Mishra, and E. Tadmor. Well-balanced and energy stable schemes for the shallow water equations with discontinuous topography. *J. Comput. Phys.*, 230(14):5587–5609, 2011.
- [23] J. M. Gallardo, M. Castro, C. Parés, and J. M. González-Vida. On a well-balanced high-order finite volume scheme for the shallow water equations with bottom topography and dry areas. In *Hyperbolic problems: theory, numerics, applications*, pages 259–270. Springer, Berlin, 2008.
- [24] T. Gallouët, J.-M. Hérard, and N. Seguin. Some approximate Godunov schemes to compute shallow-water equations with topography. *Comput. & Fluids*, 32(4):479–513, 2003.
- [25] E. Godlewski and P.-A. Raviart. *Numerical approximation of hyperbolic systems of conservation laws*, volume 118 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1996.
- [26] L. Gosse. A well-balanced flux-vector splitting scheme designed for hyperbolic systems of conservation laws with source terms. *Comput. Math. Appl.*, 39(9-10):135–159, 2000.
- [27] S. Gottlieb, C.-W. Shu, and E. Tadmor. Strong stability-preserving high-order time discretization methods. *SIAM Rev.*, 43(1):89–112, 2001.
- [28] N. Goutal and F. Maurel. Proceedings of the 2<sup>nd</sup> Workshop on Dam-Break Wave Simulation. Technical report, Groupe Hydraulique Fluviale, Département Laboratoire National d’Hydraulique, Electricité de France, 1997.
- [29] J. M. Greenberg and A.-Y. LeRoux. A well-balanced scheme for the numerical processing of source terms in hyperbolic equations. *SIAM J. Numer. Anal.*, 33(1):1–16, 1996.
- [30] A. Harten, P. D. Lax, and B. van Leer. On upstream differencing and Godunov-type schemes for hyperbolic conservation laws. *SIAM Rev.*, 25(1):35–61, 1983.
- [31] X. Y. Hu, N. A. Adams, and C.-W. Shu. Positivity-preserving method for high-order conservative schemes solving compressible Euler equations. *J. Comput. Phys.*, 242:169–180, 2013.

- [32] S. Jin. A steady-state capturing method for hyperbolic systems with geometrical source terms. *M2AN Math. Model. Numer. Anal.*, 35(4):631–645, 2001.
- [33] R. Käppeli and S. Mishra. Well-balanced schemes for the Euler equations with gravitation. *J. Comput. Phys.*, 259:199–219, 2014.
- [34] R. J. LeVeque. *Numerical methods for conservation laws*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 1992.
- [35] R. J. LeVeque. *Finite volume methods for hyperbolic problems*. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 2002.
- [36] Q. Liang and F. Marche. Numerical resolution of well-balanced shallow water equations with complex source terms. *Adv. Water Resour.*, 32(6):873–884, 2009.
- [37] M. Lukáčová-Medvidová, S. Noelle, and M. Kraft. Well-balanced finite volume evolution Galerkin methods for the shallow water equations. *J. Comput. Phys.*, 221(1):122–147, 2007.
- [38] J. Luo, K. Xu, and N. Liu. A well-balanced symplecticity-preserving gas-kinetic scheme for hydrodynamic equations under gravitational field. *SIAM J. Sci. Comput.*, 33(5):2356–2381, 2011.
- [39] R. Manning. On the flow of water in open channels and pipes. *Transactions of the Institution of Civil Engineers of Ireland*, 20:161–207, 1890.
- [40] V. Michel-Dansac, C. Berthon, S. Clain, and F. Foucher. A well-balanced scheme for the shallow-water equations with topography. *Comput. Math. Appl.*, 72(3):568–593, 2016.
- [41] R. Natalini, M. Ribot, and M. Twarogowska. A well-balanced numerical scheme for a one dimensional quasilinear hyperbolic model of chemotaxis. *Commun. Math. Sci.*, 12(1):13–39, 2014.
- [42] I. K. Nikolos and A. I. Delis. An unstructured node-centered finite volume scheme for shallow water flows with wet-dry fronts over complex topography. *Comput. Methods Appl. Mech. Engrg.*, 198(47-48):3723–3750, 2009.
- [43] S. Noelle, Y. Xing, and C.-W. Shu. High-order well-balanced finite volume WENO schemes for shallow water equation with moving water. *J. Comput. Phys.*, 226(1):29–58, 2007.
- [44] G. Russo and A. Khe. High order well balanced schemes for systems of balance laws. In *Hyperbolic problems: theory, numerics and applications*, volume 67 of *Proc. Sympos. Appl. Math.*, pages 919–928. Amer. Math. Soc., Providence, RI, 2009.

- [45] E. F. Toro. *Riemann solvers and numerical methods for fluid dynamics. A practical introduction*. Springer-Verlag, Berlin, third edition, 2009.
- [46] E. F. Toro, M. Spruce, and W. Speares. Restoration of the contact surface in the HLL-Riemann solver. *Shock Waves*, 4(1):25–34, 1994.
- [47] B. van Leer. Towards the Ultimate Conservative Difference Scheme, V. A Second Order Sequel to Godunov’s Method. *J. Com. Phys.*, 32:101–136, 1979.
- [48] B. van Leer. On the relation between the upwind-differencing schemes of Godunov, Engquist-Osher and Roe. *SIAM J. Sci. Statist. Comput.*, 5(1):1–20, 1984.
- [49] Y. Xing. Exactly well-balanced discontinuous Galerkin methods for the shallow water equations with moving water equilibrium. *J. Comput. Phys.*, 257(part A):536–553, 2014.
- [50] Y. Xing, C.-W. Shu, and S. Noelle. On the advantage of well-balanced schemes for moving-water equilibria of the shallow water equations. *J. Sci. Comput.*, 48(1-3):339–349, 2011.
- [51] K. Xu. A well-balanced gas-kinetic scheme for the shallow-water equations with source terms. *J. Comput. Phys.*, 178(2):533–562, 2002.
- [52] K. Xu, J. Luo, and S. Chen. A well-balanced kinetic scheme for gas dynamic equations under gravitational field. *Adv. Appl. Math. Mech.*, 2(2):200–210, 2010.