

A SECOND-ORDER WELL-BALANCED POSITIVITY PRESERVING CENTRAL-UPWIND SCHEME FOR THE SAINT-VENANT SYSTEM*

ALEXANDER KURGANOV[†] AND GUERGANA PETROVA[‡]

Abstract. A family of Godunov-type central-upwind schemes for the Saint-Venant system of shallow water equations has been first introduced in [A. Kurganov and D. Levy, *M2AN Math. Model. Numer. Anal.*, 36, 397-425, 2002]. Depending on the reconstruction step, the second-order versions of the schemes there could be made either well-balanced or positivity preserving, but fail to satisfy both properties simultaneously.

Here, we introduce an improved second-order central-upwind scheme which, unlike its forerunners, is capable to both preserve stationary steady states (lake at rest) and to guarantee the positivity of the computed fluid depth. Another novel property of the proposed scheme is its applicability to models with discontinuous bottom topography. We demonstrate these features of the new scheme, as well as its high resolution and robustness in a number of one- and two-dimensional examples.

Key words. Hyperbolic systems of conservation and balance laws, semi-discrete central-upwind schemes, Saint-Venant system of shallow water equations

AMS subject classifications. 65M99, 35L65

1. Introduction

We are interested in developing a simple, accurate, and robust numerical method for the Saint-Venant system of shallow water equations, which was introduced more than 130 years ago in [24] and is still widely used to model flows in rivers and coastal areas. In the one-dimensional (1-D) case, the Saint-Venant system reads:

$$\begin{cases} h_t + (hu)_x = 0, \\ (hu)_t + \left(hu^2 + \frac{1}{2}gh^2\right)_x = -ghB', \end{cases} \quad (1.1)$$

where $B(x)$ represents the bottom elevation, h is the fluid depth above the bottom, u is the velocity, and g is the gravitational constant.

The system (1.1) admits smooth steady-state solutions, satisfying

$$hu = \text{Const}, \quad \frac{u^2}{2} + g(h + B) = \text{Const},$$

as well as nonsmooth steady-state solutions. Both are physically relevant and thus practically important. A good numerical method for the system (1.1) should accurately capture both the steady states and their small perturbations (quasi-steady flows). From practical point of view, one of the most important steady-state solutions is a stationary one (lake at rest):

$$u = 0, \quad h + B = \text{Const}. \quad (1.2)$$

The methods that exactly preserve such solutions are called well-balanced, and we refer the reader to [2, 4, 8, 10, 12, 16, 17, 20, 21, 22, 23, 27, 28], where a variety of high-order well-balanced schemes for the Saint-Venant system can be found. Even though

*Received: June 16, 2006; accepted (in revised version): December 9, 2006. Communicated by Lorenzo Pareschi.

[†]Mathematics Department, Tulane University, New Orleans, LA 70118 (kurganov@math.tulane.edu).

[‡]Department of Mathematics, Texas A&M University, College Station, TX 77843 (gpetrova@math.tamu.edu).

a rigorous stability analysis of these schemes is usually out of reach, they typically produce highly accurate approximation of quasi-steady solutions and nonstationary steady states. A difficulty may occur when dry ($h=0$) or near dry ($h\sim 0$) states are to be captured. In these cases, due to numerical oscillations, h may become negative and the numerical computations will simply break down, since the eigen-values of the Jacobian of (1.1) are $u \pm \sqrt{gh}$ (the problem is especially pronounced when a high-order method is applied). Therefore, another important property of a reliable scheme for (1.1) is the positivity preserving property: all computed values of the fluid depth h should be nonnegative. Note that (near) dry states are as important in practice as the steady states, and thus the positivity preserving property should be the second major requirement when a numerical method for (1.1) is designed. Various positivity preserving numerical methods are available (see, e.g., [2, 12, 21]).

In summary, an ideal method for solving (1.1) is a method that satisfies both the well-balanced and positivity preserving properties, and at the same time is sufficiently accurate, efficient, and robust. High-resolution Godunov-type semi-discrete central-upwind schemes, introduced in [13, 14] as a universal Riemann-problem-solver-free method for general multidimensional systems of hyperbolic conservation laws, satisfy the latter three requirements. These schemes have been generalized for systems of balance laws, and in particular for the Saint-Venant system, in [12]. They are well-balanced due to the change from the variables (h, hu) to $(w = h + B, hu)$, where w represents the water surface (the same change of variables was used in [22] in the context of staggered central schemes), and due to the use of a special quadrature for the approximation of the cell averages of the source term. Note that in these schemes the positivity of h can be guaranteed only when a reconstruction of h — not w — is used near dry areas. At the same time, the reconstruction of w is needed in order to preserve the well-balanced property of the schemes. Thus, the central-upwind schemes from [12] may fail to be simultaneously well-balanced and positivity preserving on the entire computational domain. Another disadvantage of these schemes is the fact, that they can be applied only in the case of continuous bottom topography function B .

In the present paper, we introduce a new second-order central-upwind scheme, whose derivation is based on the approach from [12]. The main advantage of the new scheme is its capability to simultaneously guarantee preservation of the stationary steady states and the nonnegativity of the fluid depth h throughout the entire computational domain. We use the same quadrature to approximate the cell averages of the source term and work with the same variables w and hu , as in [12]. The main difference is that we now always use a non-oscillatory conservative piecewise linear reconstruction of w , which is properly corrected near dry areas, without switching to a reconstruction of h there. The correction heavily relies on the fact that we replace the original bottom function B with its continuous piecewise linear (bilinear in the two-dimensional (2-D) case) approximation. This approximant is not only used at the reconstruction step, but is also beneficial for the handling of sharp bottom topographies, even discontinuous ones (see §2.1 and §3.1).

The paper is organized as follows. The description of the new 1-D and 2-D schemes, as well as the proof of their positivity preserving property, are presented in §2 and §3, respectively. In §4, we demonstrate the high resolution and robustness of the new schemes in a series of 1-D and 2-D numerical experiments.

2. One-dimensional scheme

Following the approach in [12, 22], we first rewrite the system (1.1) in an equiva-

lent form in terms of the water surface $w := h + B$ and the discharge hu :

$$\begin{cases} w_t + (hu)_x = 0, \\ (hu)_t + \left[\frac{(hu)^2}{w-B} + \frac{g}{2}(w-B)^2 \right]_x = -g(w-B)B', \end{cases} \quad (2.1)$$

and then we apply the second-order semi-discrete central-upwind scheme from [13] to (2.1).

For simplicity, we introduce a uniform grid $x_\alpha := \alpha \Delta x$, where Δx is a small spatial scale, and we denote by I_j the finite volume cells $I_j := [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$. A central-upwind semi-discretization of (2.1) is the following system of ODEs:

$$\frac{d}{dt} \bar{\mathbf{U}}_j(t) = - \frac{\mathbf{H}_{j+\frac{1}{2}}(t) - \mathbf{H}_{j-\frac{1}{2}}(t)}{\Delta x} + \bar{\mathbf{S}}_j(t), \quad (2.2)$$

where $\bar{\mathbf{U}}_j(t)$ are approximations of the cell averages of the solution:

$$\bar{\mathbf{U}}_j(t) \approx \frac{1}{\Delta x} \int_{I_j} \mathbf{U}(x, t) dx, \quad \mathbf{U} := (w, hu)^T,$$

$\bar{\mathbf{S}}_j$ is an appropriate discretization of the cell averages of the source term:

$$\bar{\mathbf{S}}_j(t) \approx \frac{1}{\Delta x} \int_{I_j} \mathbf{S}(\mathbf{U}(x, t), B(x)) dx, \quad \mathbf{S} := (0, -g(w-B)B')^T,$$

and the central-upwind numerical fluxes $\mathbf{H}_{j+\frac{1}{2}}$ are given by:

$$\begin{aligned} \mathbf{H}_{j+\frac{1}{2}}(t) &= \frac{a_{j+\frac{1}{2}}^+ \mathbf{F}(\mathbf{U}_{j+\frac{1}{2}}^-, B(x_{j+\frac{1}{2}})) - a_{j+\frac{1}{2}}^- \mathbf{F}(\mathbf{U}_{j+\frac{1}{2}}^+, B(x_{j+\frac{1}{2}}))}{a_{j+\frac{1}{2}}^+ - a_{j+\frac{1}{2}}^-} \\ &\quad + \frac{a_{j+\frac{1}{2}}^+ a_{j+\frac{1}{2}}^-}{a_{j+\frac{1}{2}}^+ - a_{j+\frac{1}{2}}^-} [\mathbf{U}_{j+\frac{1}{2}}^+ - \mathbf{U}_{j+\frac{1}{2}}^-]. \end{aligned} \quad (2.3)$$

Here,

$$\mathbf{F}(\mathbf{U}, B) := (hu, \frac{(hu)^2}{w-B} + \frac{g}{2}(w-B)^2)^T,$$

and $\mathbf{U}_{j+\frac{1}{2}}^\pm$ are the right/left point values at $x = x_{j+\frac{1}{2}}$ of the piecewise linear reconstruction $\tilde{\mathbf{U}}$,

$$\tilde{\mathbf{U}}(x) := \bar{\mathbf{U}}_j + (\mathbf{U}_x)_j(x - x_j), \quad x_{j-\frac{1}{2}} < x < x_{j+\frac{1}{2}}, \quad (2.4)$$

of \mathbf{U} at time t , that is:

$$\mathbf{U}_{j+\frac{1}{2}}^\pm := \tilde{\mathbf{U}}(x_{j+\frac{1}{2}} \pm 0) = \bar{\mathbf{U}}_{j+\frac{1}{2} \pm \frac{1}{2}} \mp \frac{\Delta x}{2} (\mathbf{U}_x)_{j+\frac{1}{2} \pm \frac{1}{2}}. \quad (2.5)$$

The numerical derivatives $(\mathbf{U}_x)_j$ are (at least) first-order, component-wise approximations of $\mathbf{U}_x(x_j, t)$, computed using a nonlinear limiter needed to ensure a non-oscillatory nature of the reconstruction (2.4). The right- and left-sided local speeds

$a_{j+\frac{1}{2}}^\pm$ in (2.3) are obtained from the largest and the smallest eigenvalues of the Jacobian $\frac{\partial \mathbf{F}}{\partial \mathbf{U}}$ (see §2.2 for details). Note that the quantities $\bar{\mathbf{U}}_j$, $\mathbf{U}_{j+\frac{1}{2}}^\pm$, $(\mathbf{U}_x)_j$, and $a_{j+\frac{1}{2}}^\pm$ in (2.3)–(2.5) depend on t , but from now on we, for simplicity, will suppress this dependence in our notation.

The central-upwind semi-discretization of (2.1) results in the system of ODEs (2.2), which should be solved by a stable ODE solver of an appropriate order. In our numerical experiments, we have used the third-order strong stability preserving Runge-Kutta (SSP-RK) ODE solver (see [7] for details).

The central-upwind scheme (2.2)–(2.3),(2.5) with an appropriate discretization of $\bar{\mathbf{S}}_j(t)$, which ensures that the method is well-balanced, was proposed in [12]. The idea of its construction is very simple: to design a quadrature formula for $\int_{I_j} \mathbf{S} dx$ in such a way that it is of second order and equals $-(\mathbf{H}_{j+\frac{1}{2}}(t) - \mathbf{H}_{j-\frac{1}{2}}(t))$ when both the numerical source and fluxes are applied to stationary steady-state solutions ($w = \text{Const}$, $hu = 0$). Such a discretization of the second (nonzero) component of the source \mathbf{S} has been derived in [12] and is:

$$\bar{\mathbf{S}}_j^{(2)}(t) \approx -g \frac{B(x_{j+\frac{1}{2}}) - B(x_{j-\frac{1}{2}})}{\Delta x} \cdot \frac{(w_{j+\frac{1}{2}}^- - B(x_{j+\frac{1}{2}})) + (w_{j-\frac{1}{2}}^+ - B(x_{j-\frac{1}{2}}))}{2}. \quad (2.6)$$

The resulting semi-discrete scheme (2.2)–(2.3),(2.5)–(2.6) is well balanced, but unfortunately, it fails to satisfy another extremely important property that a reliable numerical method for the Saint-Venant system should possess — the ability to preserve the positivity of the fluid depth h . This property is crucial in computing (almost) dry states ($h \sim 0$) appearing in many different applications. To overcome this difficulty, another central-upwind scheme was derived in [12], which has the positivity preserving property, but is not well-balanced, and an adaptive strategy of switching from the well-balanced scheme to the positivity preserving one at/near (almost) dry states was suggested.

The main result of this paper is the development of a new central-upwind scheme, which is both well-balanced and positivity preserving. It also can be applied to Saint-Venant systems with a discontinuous bottom topography B , which cannot be numerically solved by the schemes from [12]. This is achieved by including the following two ingredients into the scheme (2.2)–(2.3),(2.5)–(2.6):

- Replacement of the bottom topography function B with its piecewise linear continuous approximation (see §2.1).
- A special conservative correction of the piecewise linear reconstruction \tilde{w} (the first component in (2.4)), which ensures the positivity preserving property of the resulting central-upwind scheme without sacrificing its well-balanced feature (see §2.2).

2.1. Piecewise linear approximation of the bottom. Before making the first evolution step, we replace the bottom topography function B with its continuous piecewise linear approximation \tilde{B} (see Figure 2.1), consisting of the linear pieces that connect the points $(x_{j+\frac{1}{2}}, B_{j+\frac{1}{2}})$:

$$\tilde{B}(x) = B_{j-\frac{1}{2}} + \left(B_{j+\frac{1}{2}} - B_{j-\frac{1}{2}} \right) \cdot \frac{x - x_{j-\frac{1}{2}}}{\Delta x}, \quad x_{j-\frac{1}{2}} \leq x \leq x_{j+\frac{1}{2}}, \quad (2.7)$$

where

$$B_{j+\frac{1}{2}} := \frac{B(x_{j+\frac{1}{2}}+0) + B(x_{j+\frac{1}{2}}-0)}{2}, \quad (2.8)$$

which reduces to $B_{j+\frac{1}{2}} = B(x_{j+\frac{1}{2}})$ if B is continuous at $x = x_{j+\frac{1}{2}}$.

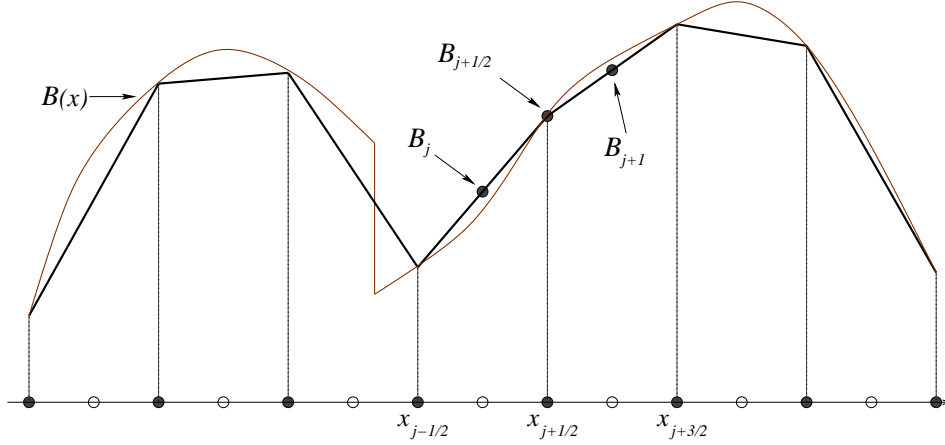


FIG. 2.1. Bottom topography function B and its piecewise linear approximation.

We would like to point out that since \tilde{B} is a piecewise linear function, its point value at $x = x_j$ coincides with its cell average over the cell I_j and is also equal to the average of the values of \tilde{B} at the endpoints of I_j , namely:

$$B_j := \tilde{B}(x_j) = \frac{1}{\Delta x} \int_{I_j} \tilde{B}(x) dx = \frac{B_{j+\frac{1}{2}} + B_{j-\frac{1}{2}}}{2}. \quad (2.9)$$

Equation (2.9) is important for the analysis of the new scheme and plays an essential role in the proof of the positivity preserving property of the scheme (see Theorem 2.1). Notice that if one takes B_j to be the value of the bottom topography function at $x = x_j$, that is, if $B_j = B(x_j)$, as it was done in [12], equation (2.9) would not hold. We would also like to point out that the condition (2.9) was imposed in [22] in order to derive a well-balanced staggered central schemes.

REMARK 2.1. Using the notation introduced in this section, we can rewrite the quadrature (2.6) as:

$$\bar{\mathbf{S}}_j^{(2)}(t) \approx -g(\bar{w}_j - B_j) \frac{B_{j+\frac{1}{2}} - B_{j-\frac{1}{2}}}{\Delta x}, \quad (2.10)$$

and the numerical fluxes (2.3) as:

$$\begin{aligned} \mathbf{H}_{j+\frac{1}{2}}(t) &= \frac{a_{j+\frac{1}{2}}^+ \mathbf{F}(\mathbf{U}_{j+\frac{1}{2}}^-, B_{j+\frac{1}{2}}) - a_{j+\frac{1}{2}}^- \mathbf{F}(\mathbf{U}_{j+\frac{1}{2}}^+, B_{j+\frac{1}{2}})}{a_{j+\frac{1}{2}}^+ - a_{j+\frac{1}{2}}^-} \\ &\quad + \frac{a_{j+\frac{1}{2}}^+ a_{j+\frac{1}{2}}^-}{a_{j+\frac{1}{2}}^+ - a_{j+\frac{1}{2}}^-} \left[\mathbf{U}_{j+\frac{1}{2}}^+ - \mathbf{U}_{j+\frac{1}{2}}^- \right], \end{aligned} \quad (2.11)$$

where $B_{j+\frac{1}{2}}$ is given by (2.8).

REMARK 2.2. Replacing B with \tilde{B} does not affect the (formal) order of the central-upwind scheme since the piecewise linear interpolant (2.7) is second order accurate for smooth B .

2.2. Positivity preserving reconstruction for w . In this section, we discuss the details of the piecewise linear reconstruction $\tilde{\mathbf{U}} \equiv (\tilde{w}, \tilde{hu})$, where, according to (2.4)–(2.5), one needs to evaluate the numerical derivatives $(\mathbf{U}_x)_j \equiv ((w_x)_j, (hu)_x)_j$. It is a well-known fact that the non-oscillatory property and nonlinear stability of the resulting scheme hinges on the non-oscillatory property of the reconstruction. This is accomplished if the numerical derivatives are computed using a nonlinear limiter. A library of such limiters is available (see, e.g., [6, 11, 15, 17, 18, 19, 25]), and one can compute the numerical derivatives using one's favorite limiter. In our numerical experiments, we have used the generalized minmod limiter [15, 18, 19, 25]:

$$(\mathbf{U}_x)_j = \text{minmod} \left(\theta \frac{\bar{\mathbf{U}}_j - \bar{\mathbf{U}}_{j-1}}{\Delta x}, \frac{\bar{\mathbf{U}}_{j+1} - \bar{\mathbf{U}}_{j-1}}{2\Delta x}, \theta \frac{\bar{\mathbf{U}}_{j+1} - \bar{\mathbf{U}}_j}{\Delta x} \right), \quad \theta \in [1, 2], \quad (2.12)$$

where the minmod function is defined as:

$$\text{minmod}(z_1, z_2, \dots) := \begin{cases} \min_j \{z_j\}, & \text{if } z_j > 0 \quad \forall j, \\ \max_j \{z_j\}, & \text{if } z_j < 0 \quad \forall j, \\ 0, & \text{otherwise,} \end{cases} \quad (2.13)$$

and the parameter θ can be used to control the amount of numerical viscosity present in the resulting scheme. Let us recall that larger values of θ correspond to less dissipative but, in general, more oscillatory reconstructions.

Unfortunately, the use of a nonlinear limiter cannot guarantee positivity of the point values $h_{j+\frac{1}{2}}^\pm$, which are obtained from the reconstructed values of $w_{j+\frac{1}{2}}^\pm$ by:

$$h_{j+\frac{1}{2}}^\pm := w_{j+\frac{1}{2}}^\pm - B_{j+\frac{1}{2}}, \quad (2.14)$$

since not only $w_{j+\frac{1}{2}}^\pm$, but even the cell average value \bar{w}_j may be smaller than either $B_{j+\frac{1}{2}}$ or $B_{j-\frac{1}{2}}$. To illustrate this possibility, a typical situation of an almost dry state is shown in Figure 2.2, where, for example, at the j -th cell, the cell average \bar{w}_j is clearly smaller than $B_{j+\frac{1}{2}}$ and thus even the least oscillatory, piecewise constant reconstruction ($w_{j-\frac{1}{2}}^+ = w_{j+\frac{1}{2}}^- = \bar{w}_j$) would produce a negative $h_{j+\frac{1}{2}}^-$. Therefore, we have to correct the basic piecewise linear reconstruction (2.4)–(2.5), (2.12)–(2.13) to ensure that $h_{j+\frac{1}{2}}^\pm \geq 0, \forall j$, provided $\bar{h}_j := \bar{w}_j - B_j \geq 0, \forall j$. In fact, we need to correct the reconstruction only for \tilde{w} and only in the following two cases:

$$\begin{aligned} & \text{if } w_{j+\frac{1}{2}}^- < B_{j+\frac{1}{2}}, \text{ then take } (w_x)_j := \frac{B_{j+\frac{1}{2}} - \bar{w}_j}{\Delta x/2}, \\ & \implies w_{j+\frac{1}{2}}^- = B_{j+\frac{1}{2}}, \quad w_{j-\frac{1}{2}}^+ = 2\bar{w}_j - B_{j+\frac{1}{2}}; \end{aligned} \quad (2.15)$$

$$\begin{aligned} & \text{if } w_{j-\frac{1}{2}}^+ < B_{j-\frac{1}{2}}, \text{ then take } (w_x)_j := \frac{\bar{w}_j - B_{j-\frac{1}{2}}}{\Delta x/2}, \\ & \implies w_{j+\frac{1}{2}}^- = 2\bar{w}_j - B_{j-\frac{1}{2}}, \quad w_{j-\frac{1}{2}}^+ = B_{j-\frac{1}{2}}. \end{aligned} \quad (2.16)$$

It is obvious that this correction procedure guarantees that the resulting reconstruction \tilde{w} will remain conservative and will stay above the piecewise linear approximant of the bottom topography function \tilde{B} (see Figure 2.2). Therefore, the corrected values of $h_{j+\frac{1}{2}}^\pm$, computed from (2.14) with the corrected values of the reconstruction \tilde{w} , will be nonnegative (this feature of the modified reconstruction is used in §2.3, where we prove the positivity preserving property of our new central-upwind scheme).

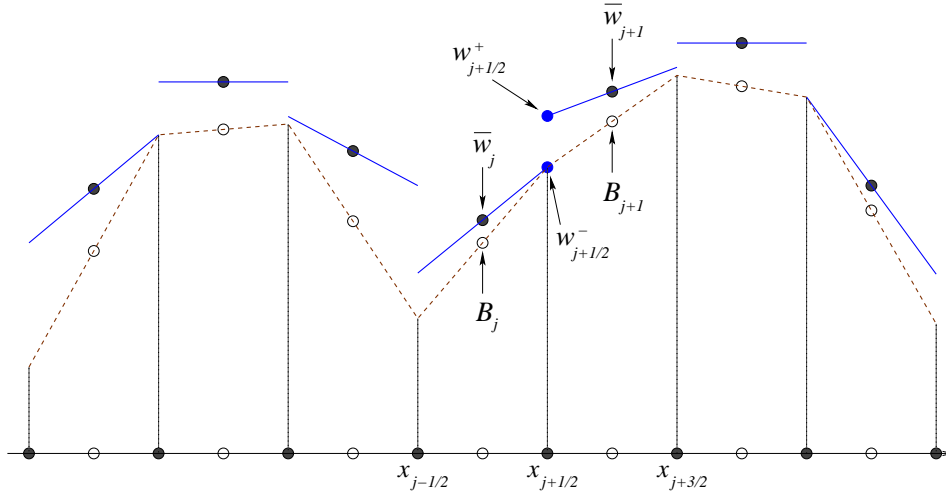


FIG. 2.2. Modified piecewise linear reconstruction for w .

However, they may be very small or even zero. This will not allow us to (accurately) compute the velocities $u_{j+\frac{1}{2}}^\pm$, required in the computation of the numerical flux and local speeds of propagation. To overcome this difficulty, we avoid the division by very small numbers by computing the velocities using the following formula (for simplicity we omit the \pm and $j \pm \frac{1}{2}$ indexes):

$$u = \frac{\sqrt{2}h(hu)}{\sqrt{h^4 + \max(h^4, \varepsilon)}}, \quad (2.17)$$

where ε is a small a-priori chosen positive number (in all our numerical experiments, $\varepsilon = (\Delta x)^4$).

In fact, there are several alternative ways to desingularize the quantity $(hu)/h$, and the simplest one is:

$$u = \frac{h(hu)}{h^2 + \varepsilon}, \quad (2.18)$$

However, we think that the suggested desingularization (2.17) has several advantages. First, it gives the exact value of u for $h \geq \varepsilon^{1/4}$. Second, for small (but not tiny) h , it minimizes the effect of the desingularization on the resulting value of u . In particular, formula (2.17) gives the corrected value

$$u^{\text{corr}} = \frac{(hu)}{h} \cdot \frac{\sqrt{2}}{\sqrt{1 + \frac{\varepsilon}{h^4}}}, \quad (2.19)$$

while the desingularization (2.18) results in:

$$u^{\text{corr}} = \frac{(hu)}{h} \cdot \frac{1}{1 + \frac{\varepsilon}{h^2}}. \quad (2.20)$$

Notice that when $h^4 \sim \varepsilon$, the ‘‘correction’’ factor in (2.19) is ~ 1 , while in the similar regime of h , namely when $h^2 \sim \varepsilon$, the ‘‘correction’’ factor in (2.20) is $\sim 1/2$. One may obviously try to minimize the effects of the desingularization by using the simpler formula (2.18) with tiny ε . Unfortunately, this may not work since, as it has been clearly indicated by our numerical experiments, the division of small values of $h(hu)$ by $h^2 + \varepsilon$ may lead to significant numerical errors, which in turn may lead to huge values of u and thus to dramatic slowdown of the computation.

As one can easily see, (2.17) reduces to $u = (hu)/h$ for large values of h , but when h is small, the entire algorithm remains consistent only if we recompute the discharge (hu) using

$$(hu) := h \cdot u, \quad (2.21)$$

where u is computed by (2.17). Notice that if (hu) is not replaced by $h \cdot u$, then the proof of Theorem 2.1 fails and, moreover, the scheme may produce negative values of h (this was confirmed by our numerical results). To the best of our knowledge, this fact has never been mentioned in the literature before.

Finally, equipped with the values of $h_{j+\frac{1}{2}}^\pm$ and $u_{j+\frac{1}{2}}^\pm$, we compute the one-sided local speeds of propagation:

$$a_{j+\frac{1}{2}}^+ = \max \left\{ u_{j+\frac{1}{2}}^+ + \sqrt{gh_{j+\frac{1}{2}}^+}, u_{j+\frac{1}{2}}^- + \sqrt{gh_{j+\frac{1}{2}}^-}, 0 \right\}, \quad (2.22)$$

$$a_{j+\frac{1}{2}}^- = \min \left\{ u_{j+\frac{1}{2}}^+ - \sqrt{gh_{j+\frac{1}{2}}^+}, u_{j+\frac{1}{2}}^- - \sqrt{gh_{j+\frac{1}{2}}^-}, 0 \right\}. \quad (2.23)$$

2.3. Positivity preserving property of the one-dimensional scheme.

In this section, we prove that the new version of the central-upwind scheme is not only well-balanced, but also preserves the positivity of the fluid depth h . We show this in the case when the system of ODEs (2.2) is discretized in time, using the forward Euler method or a higher-order SSP ODE solver. We consider two time levels, $t = t^n$ and $t = t^{n+1} := t^n + \Delta t$, and specify the time level, using the corresponding superscript. For example, the cell averages of the water surface at time $t = t^n$ will be denoted by \bar{w}_j^n . A similar notation is also used in §3.3, where the positivity preserving property of the new 2-D central-upwind scheme is proved.

The main result of this section is the following theorem.

THEOREM 2.1. *Consider the system (2.1) and the central-upwind semi-discrete scheme (2.2), (2.5), (2.8), (2.10)–(2.23). Assume that the system of ODEs (2.2) is solved by the forward Euler method and that for all j , $\bar{w}_j^n - B_j \geq 0$. Then, for all j , $\bar{w}_j^{n+1} - B_j \geq 0$, provided that $\Delta t \leq \frac{\Delta x}{2a}$, where $a := \max_j \left\{ \max \{ a_{j+\frac{1}{2}}^+, -a_{j+\frac{1}{2}}^- \} \right\}$.*

Proof. The first component in equation (2.2), together with the forward Euler temporal discretization can be written as:

$$\bar{w}_j^{n+1} = \bar{w}_j^n - \lambda \left(\mathbf{H}_{j+\frac{1}{2}}^{(1)} - \mathbf{H}_{j-\frac{1}{2}}^{(1)} \right), \quad \lambda := \frac{\Delta t}{\Delta x}, \quad (2.24)$$

where the numerical fluxes are evaluated at time level $t = t^n$. Using (2.11) and (2.14), we obtain:

$$\begin{aligned} \mathbf{H}_{j+\frac{1}{2}}^{(1)} &= \frac{a_{j+\frac{1}{2}}^+ (hu)_{j+\frac{1}{2}}^- - a_{j+\frac{1}{2}}^- (hu)_{j+\frac{1}{2}}^+}{a_{j+\frac{1}{2}}^+ - a_{j+\frac{1}{2}}^-} + \frac{a_{j+\frac{1}{2}}^+ a_{j+\frac{1}{2}}^-}{a_{j+\frac{1}{2}}^+ - a_{j+\frac{1}{2}}^-} \left[w_{j+\frac{1}{2}}^+ - w_{j+\frac{1}{2}}^- \right] \\ &= \frac{a_{j+\frac{1}{2}}^+ (hu)_{j+\frac{1}{2}}^- - a_{j+\frac{1}{2}}^- (hu)_{j+\frac{1}{2}}^+}{a_{j+\frac{1}{2}}^+ - a_{j+\frac{1}{2}}^-} + \frac{a_{j+\frac{1}{2}}^+ a_{j+\frac{1}{2}}^-}{a_{j+\frac{1}{2}}^+ - a_{j+\frac{1}{2}}^-} \left[h_{j+\frac{1}{2}}^+ - h_{j+\frac{1}{2}}^- \right]. \end{aligned}$$

Notice that (2.9) and (2.14) yield:

$$\bar{w}_j^n - B_j = \frac{w_{j-\frac{1}{2}}^+ + w_{j+\frac{1}{2}}^-}{2} - \frac{B_{j-\frac{1}{2}} + B_{j+\frac{1}{2}}}{2} = \frac{1}{2} h_{j-\frac{1}{2}}^+ + \frac{1}{2} h_{j+\frac{1}{2}}^-, \quad (2.25)$$

and thus subtracting B_j from both sides of (2.24) and using (2.25), we arrive at:

$$\begin{aligned} \bar{h}_j^{n+1} &= \left[\frac{1}{2} + \lambda a_{j-\frac{1}{2}}^- \left(\frac{a_{j-\frac{1}{2}}^+ - u_{j-\frac{1}{2}}^+}{a_{j-\frac{1}{2}}^+ - a_{j-\frac{1}{2}}^-} \right) \right] h_{j-\frac{1}{2}}^+ + \left[\frac{1}{2} - \lambda a_{j+\frac{1}{2}}^+ \left(\frac{u_{j+\frac{1}{2}}^- - a_{j+\frac{1}{2}}^-}{a_{j+\frac{1}{2}}^+ - a_{j+\frac{1}{2}}^-} \right) \right] h_{j+\frac{1}{2}}^- \\ &\quad - \lambda a_{j+\frac{1}{2}}^- \left(\frac{a_{j+\frac{1}{2}}^+ - u_{j+\frac{1}{2}}^+}{a_{j+\frac{1}{2}}^+ - a_{j+\frac{1}{2}}^-} \right) h_{j+\frac{1}{2}}^+ + \lambda a_{j-\frac{1}{2}}^+ \left(\frac{u_{j-\frac{1}{2}}^- - a_{j-\frac{1}{2}}^-}{a_{j-\frac{1}{2}}^+ - a_{j-\frac{1}{2}}^-} \right) h_{j-\frac{1}{2}}^-, \quad (2.26) \end{aligned}$$

where, as mentioned in (2.21), we use the fact that $(hu)_{j\pm\frac{1}{2}}^\pm = u_{j\pm\frac{1}{2}}^\pm \cdot h_{j\pm\frac{1}{2}}^\pm$.

Next, we argue as in [12, Theorem 4.1] and show that \bar{h}_j^{n+1} is a linear combination of the nonnegative values $h_{j\pm\frac{1}{2}}^\pm$ with nonnegative coefficients. To this end, we note that our special reconstruction procedure in §2.2 guarantees that all $h_{j\pm\frac{1}{2}}^\pm \geq 0$, if the cell averages of the fluid depth at the previous time level $t = t^n$ are nonnegative for all j , namely if $\bar{h}_j^n := \bar{w}_j^n - B_j \geq 0$. Also, it follows from (2.22)–(2.23) that $a_{j+\frac{1}{2}}^+ \geq 0$, $a_{j+\frac{1}{2}}^- \leq 0$, $a_{j+\frac{1}{2}}^+ - u_{j+\frac{1}{2}}^+ \geq 0$, and $u_{j+\frac{1}{2}}^- - a_{j+\frac{1}{2}}^- \geq 0$ for all j , and hence the last two terms in (2.26) are nonnegative. By the same argument, $0 \leq \frac{a_{j-\frac{1}{2}}^+ - u_{j-\frac{1}{2}}^+}{a_{j-\frac{1}{2}}^+ - a_{j-\frac{1}{2}}^-} \leq 1$ and $0 \leq \frac{u_{j+\frac{1}{2}}^- - a_{j+\frac{1}{2}}^-}{a_{j+\frac{1}{2}}^+ - a_{j+\frac{1}{2}}^-} \leq 1$ for all j , and thus the first two terms in (2.26) will be also nonnegative, provided we have the CFL restriction $\lambda a \leq 1/2$, with $a := \max \left\{ \max \{ a_{j+\frac{1}{2}}^+, -a_{j+\frac{1}{2}}^- \} \right\}$. Therefore, the cell average of the fluid depth computed at the next time-level, $\bar{h}_j^{n+1} \geq 0$, for all j , and the proof is completed. \square

REMARK 2.3. *Theorem 2.1 is still valid if one uses a higher-order SSP ODE solver (either the Runge-Kutta or the multistep one), because such solvers can be written as a convex combination of several forward Euler steps, see [7].*

3. Two-dimensional scheme

In this section, we describe our new second-order semi-discrete central-upwind scheme for the 2-D Saint-Venant system:

$$\begin{cases} h_t + (hu)_x + (hv)_y = 0, \\ (hu)_t + \left(hu^2 + \frac{1}{2}gh^2 \right)_x + (huv)_y = -ghB_x, \\ (hv)_t + (huv)_x + \left(hv^2 + \frac{1}{2}gh^2 \right)_y = -ghB_y, \end{cases} \quad (3.1)$$

where the function $B(x, y)$ represents the bottom elevation, h is the fluid depth above the bottom, $(u, v)^T$ is the velocity vector, and g is the gravitational constant. As in the 1-D case, we denote the water surface by $w := h + B$ and rewrite the system (3.1) in terms of the new unknown vector $\mathbf{U} := (w, hu, hv)^T$:

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U}, B)_x + \mathbf{G}(\mathbf{U}, B)_y = \mathbf{S}(\mathbf{U}, B), \quad (3.2)$$

where the fluxes and the source terms are:

$$\mathbf{F}(\mathbf{U}, B) = \left(hu, \frac{(hu)^2}{w-B} + \frac{1}{2}g(w-B)^2, \frac{(hu)(hv)}{w-B} \right)^T, \quad (3.3)$$

$$\mathbf{G}(\mathbf{U}, B) = \left(hv, \frac{(hu)(hv)}{w-B}, \frac{(hv)^2}{w-B} + \frac{1}{2}g(w-B)^2 \right)^T, \quad (3.4)$$

$$\mathbf{S}(\mathbf{U}, B) = \left(0, -(w-B)B_x, -(w-B)B_y \right)^T. \quad (3.5)$$

We denote by $C_{j,k}$ the computational cells $C_{j,k} := [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}] \times [y_{k-\frac{1}{2}}, y_{k+\frac{1}{2}}]$, where $x_\alpha = \alpha\Delta x$ and $y_\beta = \beta\Delta y$. As in the 1-D case, a central-upwind semi-discretization of (3.2)–(3.5) results in the system of ODEs (see [12, 13] for details),

$$\frac{d}{dt} \bar{\mathbf{U}}_{j,k}(t) = -\frac{\mathbf{H}_{j+\frac{1}{2},k}^x(t) - \mathbf{H}_{j-\frac{1}{2},k}^x(t)}{\Delta x} - \frac{\mathbf{H}_{j,k+\frac{1}{2}}^y(t) - \mathbf{H}_{j,k-\frac{1}{2}}^y(t)}{\Delta y} + \bar{\mathbf{S}}_{j,k}(t), \quad (3.6)$$

for the time evolution of the approximations $\bar{\mathbf{U}}_{j,k}(t)$ of the cell averages:

$$\bar{\mathbf{U}}_{j,k}(t) \approx \frac{1}{\Delta x \Delta y} \int \int_{C_{j,k}} \mathbf{U}(x, y, t) \, dx dy.$$

The system (3.6) should be solved by a stable ODE solver of an appropriate order. In our numerical experiments, we have used the third-order SSP-RK ODE solver.

The numerical fluxes \mathbf{H}^x and \mathbf{H}^y in (3.6) are given by (see [12, 13] for details):

$$\begin{aligned} \mathbf{H}_{j+\frac{1}{2},k}^x &= \frac{a_{j+\frac{1}{2},k}^+ \mathbf{F}(\mathbf{U}_{j,k}^E, B(x_{j+\frac{1}{2}}, y_k)) - a_{j+\frac{1}{2},k}^- \mathbf{F}(\mathbf{U}_{j+1,k}^W, B(x_{j+\frac{1}{2}}, y_k))}{a_{j+\frac{1}{2},k}^+ - a_{j+\frac{1}{2},k}^-} \\ &\quad + \frac{a_{j+\frac{1}{2},k}^+ a_{j+\frac{1}{2},k}^-}{a_{j+\frac{1}{2},k}^+ - a_{j+\frac{1}{2},k}^-} [\mathbf{U}_{j+1,k}^W - \mathbf{U}_{j,k}^E], \\ \mathbf{H}_{j,k+\frac{1}{2}}^y &= \frac{b_{j,k+\frac{1}{2}}^+ \mathbf{G}(\mathbf{U}_{j,k}^N, B(x_j, y_{k+\frac{1}{2}})) - b_{j,k+\frac{1}{2}}^- \mathbf{G}(\mathbf{U}_{j,k+1}^S, B(x_j, y_{k+\frac{1}{2}}))}{b_{j,k+\frac{1}{2}}^+ - b_{j,k+\frac{1}{2}}^-} \\ &\quad + \frac{b_{j,k+\frac{1}{2}}^+ b_{j,k+\frac{1}{2}}^-}{b_{j,k+\frac{1}{2}}^+ - b_{j,k+\frac{1}{2}}^-} [\mathbf{U}_{j,k+1}^S - \mathbf{U}_{j,k}^N]. \end{aligned} \quad (3.7)$$

Here, $\mathbf{U}_{j,k}^{E,W,N,S}$ are the point values of the piecewise linear reconstruction $\tilde{\mathbf{U}} \equiv (\tilde{w}, \tilde{hu}, \tilde{hv})$ for \mathbf{U} ,

$$\tilde{\mathbf{U}}(x, y) := \bar{\mathbf{U}}_{j,k} + (\mathbf{U}_x)_{j,k}(x - x_j) + (\mathbf{U}_y)_{j,k}(y - y_k), \quad (x, y) \in C_{j,k}, \quad (3.8)$$

at $(x_{j+\frac{1}{2}}, y_k)$, $(x_{j-\frac{1}{2}}, y_k)$, $(x_j, y_{k+\frac{1}{2}})$, and $(x_j, y_{k-\frac{1}{2}})$, respectively. Namely, we have:

$$\begin{aligned} \mathbf{U}_{j,k}^E &:= \tilde{\mathbf{U}}(x_{j+\frac{1}{2}} - 0, y_k) = \bar{\mathbf{U}}_{j,k} + \frac{\Delta x}{2} (\mathbf{U}_x)_{j,k}, \\ \mathbf{U}_{j,k}^W &:= \tilde{\mathbf{U}}(x_{j-\frac{1}{2}} + 0, y_k) = \bar{\mathbf{U}}_{j,k} - \frac{\Delta x}{2} (\mathbf{U}_x)_{j,k}, \\ \mathbf{U}_{j,k}^N &:= \tilde{\mathbf{U}}(x_j, y_{k+\frac{1}{2}} - 0) = \bar{\mathbf{U}}_{j,k} + \frac{\Delta y}{2} (\mathbf{U}_y)_{j,k}, \\ \mathbf{U}_{j,k}^S &:= \tilde{\mathbf{U}}(x_j, y_{k-\frac{1}{2}} + 0) = \bar{\mathbf{U}}_{j,k} - \frac{\Delta y}{2} (\mathbf{U}_y)_{j,k}. \end{aligned} \tag{3.9}$$

The numerical derivatives $(\mathbf{U}_x)_{j,k}$ and $(\mathbf{U}_y)_{j,k}$ are (at least) first-order component-wise approximations of $\mathbf{U}_x(x_j, y_k, t)$ and $\mathbf{U}_y(x_j, y_k, t)$, respectively, and are computed using a nonlinear limiter. The one-sided local speeds in the x - and y -directions, $a_{j+\frac{1}{2},k}^\pm$ and $b_{j,k+\frac{1}{2}}^\pm$, are obtained from the largest and the smallest eigenvalues of the Jacobians $\frac{\partial \mathbf{F}}{\partial \mathbf{U}}$ and $\frac{\partial \mathbf{G}}{\partial \mathbf{U}}$, respectively. As before, we suppress the dependence of $\bar{\mathbf{U}}_{j,k}$, $\mathbf{U}_{j,k}^{E,W,N,S}$, $(\mathbf{U}_x)_{j,k}$, $(\mathbf{U}_y)_{j,k}$, $a_{j+\frac{1}{2},k}^\pm$, and $b_{j,k+\frac{1}{2}}^\pm$ on t to simplify the notation.

To guarantee that the scheme (3.6) is well-balanced, an appropriate discretization of the cell average of the source term,

$$\bar{\mathbf{S}}_{j,k}(t) \approx \frac{1}{\Delta x \Delta y} \int \int_{C_{j,k}} \mathbf{S}(\mathbf{U}(x,y,t), B(x,y)) \, dx dy,$$

is needed. Similarly to the 1-D case, the following quadrature formula for the nonzero components of $\bar{\mathbf{S}}_{j,k}(t)$ was proposed in [12]:

$$\begin{aligned} \bar{\mathbf{S}}_{j,k}^{(2)}(t) &\approx -g \frac{B(x_{j+\frac{1}{2}}, y_k) - B(x_{j-\frac{1}{2}}, y_k)}{\Delta x} \cdot \frac{(w_{j,k}^E - B(x_{j+\frac{1}{2}}, y_k)) + (w_{j,k}^W - B(x_{j-\frac{1}{2}}, y_k))}{2}, \\ \bar{\mathbf{S}}_{j,k}^{(3)}(t) &\approx -g \frac{B(x_j, y_{k+\frac{1}{2}}) - B(x_j, y_{k-\frac{1}{2}})}{\Delta y} \cdot \frac{(w_{j,k}^N - B(x_j, y_{k+\frac{1}{2}})) + (w_{j,k}^S - B(x_j, y_{k-\frac{1}{2}}))}{2}. \end{aligned} \tag{3.10}$$

The semi-discrete central-upwind scheme (3.6)–(3.10) was derived in [12]. It is well-balanced, but is not positivity preserving and can be applied only to problems, in which the bottom topography function B is continuous.

Next, we introduce a modification of this scheme, which guarantees that the resulting method is both well-balanced and positivity preserving and can be applied to discontinuous bottom functions B as well. The modification steps are as in the 1-D case, namely, we propose a continuous piecewise bilinear approximation to B , correction of the reconstruction \tilde{w} , and recalculation of the values of the reconstructions $\tilde{h}\tilde{u}$ and $\tilde{h}\tilde{v}$ for small values of the reconstruction \tilde{h} . The details are described in Sections 3.1–3.3.

3.1. Piecewise bilinear approximation of the bottom. We start by replacing the bottom topography function B with its continuous piecewise bilinear

approximation \tilde{B} , which at each cell $C_{j,k}$ is given by the bilinear form:

$$\begin{aligned} \tilde{B}(x, y) = & B_{j-\frac{1}{2}, k-\frac{1}{2}} + \left(B_{j+\frac{1}{2}, k-\frac{1}{2}} - B_{j-\frac{1}{2}, k-\frac{1}{2}} \right) \cdot \frac{x - x_{j-\frac{1}{2}}}{\Delta x} \\ & + \left(B_{j-\frac{1}{2}, k+\frac{1}{2}} - B_{j-\frac{1}{2}, k-\frac{1}{2}} \right) \cdot \frac{y - y_{k-\frac{1}{2}}}{\Delta y} \\ & + \left(B_{j+\frac{1}{2}, k+\frac{1}{2}} - B_{j+\frac{1}{2}, k-\frac{1}{2}} - B_{j-\frac{1}{2}, k+\frac{1}{2}} + B_{j-\frac{1}{2}, k-\frac{1}{2}} \right) \frac{(x - x_{j-\frac{1}{2}})(y - y_{k-\frac{1}{2}})}{\Delta x \Delta y}, \quad (x, y) \in C_{j,k}. \end{aligned} \quad (3.11)$$

Here, $B_{j\pm\frac{1}{2}, k\pm\frac{1}{2}}$ are the values of \tilde{B} at the corners of the cell $C_{j,k}$, computed according to the following formula:

$$\begin{aligned} B_{j\pm\frac{1}{2}, k\pm\frac{1}{2}} := & \frac{1}{2} \left(\max_{\xi^2 + \eta^2 = 1} \lim_{h, \ell \rightarrow 0} B(x_{j\pm\frac{1}{2}} + h\xi, y_{k\pm\frac{1}{2}} + \ell\eta) \right. \\ & \left. + \min_{\xi^2 + \eta^2 = 1} \lim_{h, \ell \rightarrow 0} B(x_{j\pm\frac{1}{2}} + h\xi, y_{k\pm\frac{1}{2}} + \ell\eta) \right), \end{aligned}$$

which reduces to

$$B_{j\pm\frac{1}{2}, k\pm\frac{1}{2}} = B(x_{j\pm\frac{1}{2}}, y_{k\pm\frac{1}{2}}),$$

if the function B is continuous at $(x_{j\pm\frac{1}{2}}, y_{k\pm\frac{1}{2}})$.

Note that the restriction of the interpolant \tilde{B} along each of the lines $x = x_j$ or $y = y_k$ is a continuous piecewise linear function, and, as in the 1-D case (see (2.9)), the cell average of \tilde{B} over the cell $C_{j,k}$ is equal to its value at the center of the cell and is also equal to the average of the values of \tilde{B} at the midpoints of the edges of $C_{j,k}$, namely we have:

$$\begin{aligned} B_{j,k} := \tilde{B}(x_j, y_k) &= \frac{1}{\Delta x \Delta y} \int \int_{C_{j,k}} \tilde{B}(x, y) \, dx dy \\ &= \frac{1}{4} \left(B_{j+\frac{1}{2}, k} + B_{j-\frac{1}{2}, k} + B_{j, k+\frac{1}{2}} + B_{j, k-\frac{1}{2}} \right), \end{aligned} \quad (3.12)$$

where

$$B_{j+\frac{1}{2}, k} := \tilde{B}(x_{j+\frac{1}{2}}, y_k) = \frac{1}{2} \left(B_{j+\frac{1}{2}, k+\frac{1}{2}} + B_{j+\frac{1}{2}, k-\frac{1}{2}} \right), \quad (3.13)$$

and

$$B_{j, k+\frac{1}{2}} := \tilde{B}(x_j, y_{k+\frac{1}{2}}) = \frac{1}{2} \left(B_{j+\frac{1}{2}, k+\frac{1}{2}} + B_{j-\frac{1}{2}, k+\frac{1}{2}} \right). \quad (3.14)$$

Formulae (3.12)–(3.14) are crucial for the proof of the positivity preserving property of our 2-D well-balanced central-upwind scheme (see Theorem 3.1). One of the reasons the well-balanced central-upwind scheme from [12] fails to preserve the nonnegativity of the fluid depth h is that (3.12) is not true when the original (non-reconstructed) values of B are used.

Finally, using the notation introduced in this section and taking into account that \widetilde{w} and \widetilde{B} are piecewise linear functions, we obtain the following relations between their point values:

$$\overline{w}_{j,k} = \frac{w_{j,k}^E + w_{j,k}^W}{2} = \frac{w_{j,k}^S + w_{j,k}^N}{2}, \quad B_{j,k} = \frac{B_{j+\frac{1}{2},k} + B_{j-\frac{1}{2},k}}{2} = \frac{B_{j,k+\frac{1}{2}} + B_{j,k-\frac{1}{2}}}{2}. \quad (3.15)$$

Then, applying the source quadratures (3.10) to \widetilde{B} and using formulae (3.15), we arrive at:

$$\overline{\mathbf{S}}_{j,k}^{(2)}(t) \approx -g(\overline{w}_{j,k} - B_{j,k}) \frac{B_{j+\frac{1}{2},k} - B_{j-\frac{1}{2},k}}{\Delta x}, \quad \overline{\mathbf{S}}_{j,k}^{(3)}(t) \approx -g(\overline{w}_{j,k} - B_{j,k}) \frac{B_{j,k+\frac{1}{2}} - B_{j,k-\frac{1}{2}}}{\Delta y}. \quad (3.16)$$

Notice that here, $B_{j+\frac{1}{2},k} \neq B(x_{j+\frac{1}{2}}, y_k)$ and $B_{j,k+\frac{1}{2}} \neq B(x_j, y_{k+\frac{1}{2}})$ even when B is continuous, and that for our new scheme, the numerical fluxes (3.7) become:

$$\begin{aligned} \mathbf{H}_{j+\frac{1}{2},k}^x &= \frac{a_{j+\frac{1}{2},k}^+ \mathbf{F}(\mathbf{U}_{j,k}^E, B_{j+\frac{1}{2},k}) - a_{j+\frac{1}{2},k}^- \mathbf{F}(\mathbf{U}_{j+1,k}^W, B_{j+\frac{1}{2},k})}{a_{j+\frac{1}{2},k}^+ - a_{j+\frac{1}{2},k}^-} \\ &\quad + \frac{a_{j+\frac{1}{2},k}^+ a_{j+\frac{1}{2},k}^-}{a_{j+\frac{1}{2},k}^+ - a_{j+\frac{1}{2},k}^-} [\mathbf{U}_{j+1,k}^W - \mathbf{U}_{j,k}^E], \\ \mathbf{H}_{j,k+\frac{1}{2}}^y &= \frac{b_{j,k+\frac{1}{2}}^+ \mathbf{G}(\mathbf{U}_{j,k}^N, B_{j,k+\frac{1}{2}}) - b_{j,k+\frac{1}{2}}^- \mathbf{G}(\mathbf{U}_{j,k+1}^S, B_{j,k+\frac{1}{2}})}{b_{j,k+\frac{1}{2}}^+ - b_{j,k+\frac{1}{2}}^-} \\ &\quad + \frac{b_{j,k+\frac{1}{2}}^+ b_{j,k+\frac{1}{2}}^-}{b_{j,k+\frac{1}{2}}^+ - b_{j,k+\frac{1}{2}}^-} [\mathbf{U}_{j,k+1}^S - \mathbf{U}_{j,k}^N], \end{aligned} \quad (3.17)$$

where $B_{j+\frac{1}{2},k}$ and $B_{j,k+\frac{1}{2}}$ are given by (3.13) and (3.14).

REMARK 3.1. *As in the 1-D case, we note that replacing B with \widetilde{B} does not affect the (formal) order of the central-upwind scheme since the piecewise bilinear interpolant (3.11) is second order accurate for smooth B .*

3.2. Positivity preserving reconstruction for w . In this section, we extend the positivity preserving reconstruction, introduced in §2.2, to two space dimensions. As in the 1-D case, we begin with computing the numerical derivatives, $(\mathbf{U}_x)_{j,k} \equiv ((\widetilde{w}_x)_{j,k}, ((\widetilde{hu})_x)_{j,k}, ((\widetilde{hv})_x)_{j,k})$ and $(\mathbf{U}_y)_{j,k} \equiv ((\widetilde{w}_y)_{j,k}, ((\widetilde{hu})_y)_{j,k}, ((\widetilde{hv})_y)_{j,k})$ with the help of a nonlinear limiter. In our numerical experiments, we have used a one-parameter family of the generalized minmod limiters with $\theta \in [1, 2]$:

$$\begin{aligned} (\mathbf{U}_x)_{j,k} &= \text{minmod} \left(\theta \frac{\overline{\mathbf{U}}_{j,k} - \overline{\mathbf{U}}_{j-1,k}}{\Delta x}, \frac{\overline{\mathbf{U}}_{j+1,k} - \overline{\mathbf{U}}_{j-1,k}}{2\Delta x}, \theta \frac{\overline{\mathbf{U}}_{j+1,k} - \overline{\mathbf{U}}_{j,k}}{\Delta x} \right), \\ (\mathbf{U}_y)_{j,k} &= \text{minmod} \left(\theta \frac{\overline{\mathbf{U}}_{j,k} - \overline{\mathbf{U}}_{j,k-1}}{\Delta y}, \frac{\overline{\mathbf{U}}_{j,k+1} - \overline{\mathbf{U}}_{j,k-1}}{2\Delta y}, \theta \frac{\overline{\mathbf{U}}_{j,k+1} - \overline{\mathbf{U}}_{j,k}}{\Delta y} \right). \end{aligned} \quad (3.18)$$

Similarly to the 1-D case, the piecewise linear reconstruction (3.8)–(3.9), (3.18) (even the first-order piecewise constant one) cannot guarantee the positivity of the point values $h_{j,k}^{E,W,N,S}$, obtained from the corresponding reconstructed values $w_{j,k}^{E,W,N,S}$ using

$$\begin{aligned} h_{j,k}^E &:= w_{j,k}^E - B_{j+\frac{1}{2},k}, & h_{j,k}^W &:= w_{j,k}^W - B_{j-\frac{1}{2},k}, \\ h_{j,k}^N &:= w_{j,k}^N - B_{j,k+\frac{1}{2}}, & h_{j,k}^S &:= w_{j,k}^S - B_{j,k-\frac{1}{2}}. \end{aligned} \quad (3.19)$$

Therefore, we correct the basic reconstruction (3.8)–(3.9),(3.18) to enforce $h_{j,k}^{\text{E,W,N,S}} \geq 0$. The corrections are needed only for the \tilde{w} component of $\tilde{\mathbf{U}}$ and only in the following four cases:

$$\begin{aligned} \text{if } w_{j,k}^{\text{E}} < B_{j+\frac{1}{2},k}, \text{ then take } (w_x)_{j,k} &:= \frac{B_{j+\frac{1}{2},k} - \bar{w}_{j,k}}{\Delta x/2}, \\ \implies w_{j,k}^{\text{E}} = B_{j+\frac{1}{2},k}, \quad w_{j,k}^{\text{W}} = 2\bar{w}_{j,k} - B_{j+\frac{1}{2},k}; \end{aligned} \quad (3.20)$$

$$\begin{aligned} \text{if } w_{j,k}^{\text{W}} < B_{j-\frac{1}{2},k}, \text{ then take } (w_x)_{j,k} &:= \frac{\bar{w}_{j,k} - B_{j-\frac{1}{2},k}}{\Delta x/2}, \\ \implies w_{j,k}^{\text{E}} = 2\bar{w}_{j,k} - B_{j-\frac{1}{2},k}, \quad w_{j,k}^{\text{W}} = B_{j-\frac{1}{2},k}; \end{aligned} \quad (3.21)$$

$$\begin{aligned} \text{if } w_{j,k}^{\text{N}} < B_{j,k+\frac{1}{2}}, \text{ then take } (w_y)_{j,k} &:= \frac{B_{j,k+\frac{1}{2}} - \bar{w}_{j,k}}{\Delta y/2}, \\ \implies w_{j,k}^{\text{N}} = B_{j,k+\frac{1}{2}}, \quad w_{j,k}^{\text{S}} = 2\bar{w}_{j,k} - B_{j,k+\frac{1}{2}}; \end{aligned} \quad (3.22)$$

$$\begin{aligned} \text{if } w_{j,k}^{\text{S}} < B_{j,k-\frac{1}{2}}, \text{ then take } (w_y)_{j,k} &:= \frac{\bar{w}_{j,k} - B_{j,k-\frac{1}{2}}}{\Delta y/2}, \\ \implies w_{j,k}^{\text{N}} = 2\bar{w}_{j,k} - B_{j,k-\frac{1}{2}}, \quad w_{j,k}^{\text{S}} = B_{j,k-\frac{1}{2}}. \end{aligned} \quad (3.23)$$

The correction procedure (3.20)–(3.23) guarantees that the reconstruction of w is conservative and its restrictions on the lines $y = y_k$ and $x = x_j$ are above $\tilde{B}(x, y_k)$ and $\tilde{B}(x_j, y)$, respectively. Hence the point values of the water height, defined by (3.19), will be nonnegative. Notice that unlike the 1-D case, this does not guarantee the non-negativity of $\tilde{w} - \tilde{B}$ in the entire cell. However, this is not a problem since our goal is to preserve positivity of the cell averages of h and its point-values used in the scheme ($h_{j,k}^{\text{E}}, h_{j,k}^{\text{W}}, h_{j,k}^{\text{S}},$ and $h_{j,k}^{\text{N}}$).

As in the 1-D case, the obtained values of h may be very small (or even zero). Therefore, the corresponding velocities should be calculated in a way similar to (2.17) (we omit the E, W, S, N, j , k , indexes):

$$u = \frac{\sqrt{2}h(hu)}{\sqrt{h^4 + \max(h^4, \varepsilon)}}, \quad v = \frac{\sqrt{2}h(hv)}{\sqrt{h^4 + \max(h^4, \varepsilon)}}, \quad (3.24)$$

where ε is a prescribed tolerance (we have taken $\varepsilon = \max\{(\Delta x)^4, (\Delta y)^4\}$ in all our computations). After evaluating h , u , and v , we recompute the x - and y -discharges accordingly, that is, we set:

$$(hu) := h \cdot u, \quad (hv) := h \cdot v, \quad (3.25)$$

at the points where the fluxes are to be calculated.

Finally, the local one-sided speeds of propagation are computed by:

$$a_{j+\frac{1}{2},k}^+ = \max\left\{u_{j,k}^{\text{E}} + \sqrt{gh_{j,k}^{\text{E}}}, u_{j+1,k}^{\text{W}} + \sqrt{gh_{j+1,k}^{\text{W}}}, 0\right\}, \quad (3.26)$$

$$a_{j+\frac{1}{2},k}^- = \min\left\{u_{j,k}^{\text{E}} - \sqrt{gh_{j,k}^{\text{E}}}, u_{j+1,k}^{\text{W}} - \sqrt{gh_{j+1,k}^{\text{W}}}, 0\right\}, \quad (3.27)$$

$$b_{j,k+\frac{1}{2}}^+ = \max\left\{v_{j,k}^{\text{N}} + \sqrt{gh_{j,k}^{\text{N}}}, v_{j,k+1}^{\text{S}} + \sqrt{gh_{j,k+1}^{\text{S}}}, 0\right\}, \quad (3.28)$$

$$b_{j,k+\frac{1}{2}}^- = \min\left\{v_{j,k}^{\text{N}} - \sqrt{gh_{j,k}^{\text{N}}}, v_{j,k+1}^{\text{S}} - \sqrt{gh_{j,k+1}^{\text{S}}}, 0\right\}. \quad (3.29)$$

3.3. Positivity preserving property of the two-dimensional scheme.

In this section, we prove the positivity preserving property of our new well-balanced 2-D central-upwind scheme in the case when the system of ODEs (3.6) is discretized in time, using the forward Euler method or a higher-order SSP ODE solver. The notation we use here is similar to the one used in the 1-D case. The following theorem holds.

THEOREM 3.1. *Consider the system (3.2)–(3.5) and the central-upwind semi-discrete scheme (3.6),(3.9),(3.13)–(3.14), (3.16)–(3.29). Assume that the system of ODEs (3.6) is solved with the forward Euler method and that for all (j,k) , $\bar{w}_{j,k}^n - B_{j,k} \geq 0$. Then, for all (j,k) , $\bar{w}_{j,k}^{n+1} - B_{j,k} \geq 0$, provided that $\Delta t \leq \min \left\{ \frac{\Delta x}{4a}, \frac{\Delta y}{4b} \right\}$, where a and b are given by $a := \max_{j,k} \left\{ \max \{ a_{j+\frac{1}{2},k}^+, -a_{j+\frac{1}{2},k}^- \} \right\}$, $b := \max_{j,k} \left\{ \max \{ b_{j,k+\frac{1}{2}}^+, -b_{j,k+\frac{1}{2}}^- \} \right\}$.*

Proof. The proof of this theorem is a straightforward extension of the proof of Theorem 2.1. We write the first component in equation (3.6) together with the forward Euler temporal discretization as:

$$\bar{w}_{j,k}^{n+1} = \bar{w}_{j,k}^n - \lambda \left((\mathbf{H}^x)^{(1)}_{j+\frac{1}{2},k} - (\mathbf{H}^x)^{(1)}_{j-\frac{1}{2},k} \right) - \mu \left((\mathbf{H}^y)^{(1)}_{j,k+\frac{1}{2}} - (\mathbf{H}^y)^{(1)}_{j,k-\frac{1}{2}} \right), \quad (3.30)$$

where $\lambda := \Delta t / \Delta x$, $\mu := \Delta t / \Delta y$, and the numerical fluxes are evaluated at time level $t = t^n$. Using (3.17) and (3.19), we obtain:

$$\begin{aligned} (\mathbf{H}^x)^{(1)}_{j+\frac{1}{2},k} &= \frac{a_{j+\frac{1}{2},k}^+ (hu)_{j,k}^E - a_{j+\frac{1}{2},k}^- (hu)_{j+1,k}^W}{a_{j+\frac{1}{2},k}^+ - a_{j+\frac{1}{2},k}^-} + \frac{a_{j+\frac{1}{2},k}^+ a_{j+\frac{1}{2},k}^-}{a_{j+\frac{1}{2},k}^+ - a_{j+\frac{1}{2},k}^-} \left[w_{j+1,k}^W - w_{j,k}^E \right] \\ &= \frac{a_{j+\frac{1}{2},k}^+ (hu)_{j,k}^E - a_{j+\frac{1}{2},k}^- (hu)_{j+1,k}^W}{a_{j+\frac{1}{2},k}^+ - a_{j+\frac{1}{2},k}^-} + \frac{a_{j+\frac{1}{2},k}^+ a_{j+\frac{1}{2},k}^-}{a_{j+\frac{1}{2},k}^+ - a_{j+\frac{1}{2},k}^-} \left[h_{j+1,k}^W - h_{j,k}^E \right] \end{aligned}$$

and

$$\begin{aligned} (\mathbf{H}^y)^{(1)}_{j,k+\frac{1}{2}} &= \frac{b_{j,k+\frac{1}{2}}^+ (hv)_{j,k}^N - b_{j,k+\frac{1}{2}}^- (hv)_{j,k+1}^S}{b_{j,k+\frac{1}{2}}^+ - b_{j,k+\frac{1}{2}}^-} + \frac{b_{j,k+\frac{1}{2}}^+ b_{j,k+\frac{1}{2}}^-}{b_{j,k+\frac{1}{2}}^+ - b_{j,k+\frac{1}{2}}^-} \left[w_{j,k+1}^S - w_{j,k}^N \right] \\ &= \frac{b_{j,k+\frac{1}{2}}^+ (hv)_{j,k}^N - b_{j,k+\frac{1}{2}}^- (hv)_{j,k+1}^S}{b_{j,k+\frac{1}{2}}^+ - b_{j,k+\frac{1}{2}}^-} + \frac{b_{j,k+\frac{1}{2}}^+ b_{j,k+\frac{1}{2}}^-}{b_{j,k+\frac{1}{2}}^+ - b_{j,k+\frac{1}{2}}^-} \left[h_{j,k+1}^S - h_{j,k}^N \right]. \end{aligned}$$

It follows from (3.15) and (3.19) that

$$\begin{aligned} \bar{w}_{j,k}^n - B_{j,k} &= \frac{1}{4} (w_{j,k}^E + w_{j,k}^W + w_{j,k}^S + w_{j,k}^N) - \frac{1}{4} \left(B_{j+\frac{1}{2},k} + B_{j-\frac{1}{2},k} + B_{j,k+\frac{1}{2}} + B_{j,k-\frac{1}{2}} \right) \\ &= \frac{1}{4} (h_{j,k}^E + h_{j,k}^W + h_{j,k}^S + h_{j,k}^N), \end{aligned} \quad (3.31)$$

and thus, subtracting $B_{j,k}$ from both sides of (3.30) and applying (3.31), we arrive at:

$$\begin{aligned} \bar{h}_{j,k}^{n+1} = & \left[\frac{1}{4} + \lambda a_{j-\frac{1}{2},k}^- \left(\frac{a_{j-\frac{1}{2},k}^+ - u_{j,k}^W}{a_{j-\frac{1}{2},k}^+ - a_{j-\frac{1}{2},k}^-} \right) \right] h_{j,k}^W + \left[\frac{1}{4} - \lambda a_{j+\frac{1}{2},k}^+ \left(\frac{u_{j,k}^E - a_{j+\frac{1}{2},k}^-}{a_{j+\frac{1}{2},k}^+ - a_{j+\frac{1}{2},k}^-} \right) \right] h_{j,k}^E \\ & - \lambda a_{j+\frac{1}{2},k}^- \left(\frac{a_{j+\frac{1}{2},k}^+ - u_{j+1,k}^W}{a_{j+\frac{1}{2},k}^+ - a_{j+\frac{1}{2},k}^-} \right) h_{j+1,k}^W + \lambda a_{j-\frac{1}{2},k}^+ \left(\frac{u_{j-1,k}^E - a_{j-\frac{1}{2},k}^-}{a_{j-\frac{1}{2},k}^+ - a_{j-\frac{1}{2},k}^-} \right) h_{j-1,k}^E \\ & + \left[\frac{1}{4} + \mu b_{j,k-\frac{1}{2}}^- \left(\frac{b_{j,k-\frac{1}{2}}^+ - v_{j,k}^S}{b_{j,k-\frac{1}{2}}^+ - b_{j,k-\frac{1}{2}}^-} \right) \right] h_{j,k}^S + \left[\frac{1}{4} - \mu b_{j,k+\frac{1}{2}}^+ \left(\frac{v_{j,k}^N - b_{j,k+\frac{1}{2}}^-}{b_{j,k+\frac{1}{2}}^+ - b_{j,k+\frac{1}{2}}^-} \right) \right] h_{j,k}^N \\ & - \mu b_{j,k+\frac{1}{2}}^- \left(\frac{b_{j,k+\frac{1}{2}}^+ - v_{j,k+1}^S}{b_{j,k+\frac{1}{2}}^+ - b_{j,k+\frac{1}{2}}^-} \right) h_{j,k+1}^S + \mu b_{j,k-\frac{1}{2}}^+ \left(\frac{v_{j,k-1}^N - b_{j,k-\frac{1}{2}}^-}{b_{j,k-\frac{1}{2}}^+ - b_{j,k-\frac{1}{2}}^-} \right) h_{j,k-1}^N, \quad (3.32) \end{aligned}$$

where, as mentioned in (3.25), we have used the fact that $(hu) = h \cdot u$ and $(hv) = h \cdot v$. Next, we argue as in Theorem 2.1 to show that all eight terms on the right-hand (RHS) side of (3.32) are nonnegative. We use the definitions of the local speeds (3.26)–(3.29) and impose the CFL restriction $\lambda a \leq 1/4$, $\mu b \leq 1/4$. This completes the proof. \square

REMARK 3.2. *As in the 1-D case, Theorem 3.1 is still valid if one uses a higher-order SSP ODE solver (either the Runge-Kutta or the multistep one), because such solvers can be written as a convex combination of several forward Euler steps.*

4. Numerical experiments

We test our well-balanced positivity preserving central-upwind scheme on several 1-D and 2-D problems, in which (almost) dry stationary steady states and/or their small perturbations are present. In all examples, the proposed scheme highly accurately resolves stationary steady states and, at the same time, preserves the positivity of the fluid depth h (as proved in Theorems 2.1 and 3.1). We begin with an accuracy test, performed in Example 1. Then, in Example 2, we compare the new central-upwind scheme with the original one from [12]. Notice, however, that such comparison is not performed in other examples, since the scheme from [12] is not applicable in the case of discontinuous bottom topography function B .

In all the reported numerical experiments, the reconstruction parameter θ in (2.12) and (3.18) is selected to be $\theta = 1.3$ except for Examples 3 and 4, where $\theta = 1$ has been used. It is a well known fact that the smaller the θ , the more dissipative the scheme is and the less oscillations we have. However, in cases of systems, oscillations may appear even when $\theta = 1$.

Our experience shows that in most problems the value $\theta = 1.3$ seems to be optimal. In Examples 3 and 4, we solve problems with discontinuous bottom topography, in which case the geometric source term may become a product of a discontinuous function and a distribution. In the presence of such a nonconservative product, we would like to reduce possible oscillations, and this is the reason why we use $\theta = 1$ in these two examples.

In all figures, we plot the corresponding piecewise linear approximations of the bottom function B — not the actual B .

4.1. One-dimensional examples.

Example 1 — experimental order of accuracy. The goal of this numerical example, taken from [27], is to experimentally verify the order of accuracy of the

proposed central-upwind scheme. The scheme is applied to the system (1.1) with gravitational constant $g=9.812$ subject to the following initial data and bottom topography:

$$h(x,0)=5+e^{\cos(2\pi x)}, \quad hu(x,0)=\sin(\cos(2\pi x)), \quad B(x)=\sin^2(\pi x),$$

and 1-periodic boundary conditions.

Though the exact solution of this initial-boundary value problem is not available, it seems to be a generic problem for accuracy tests since most known solutions (in closed form) have special properties making the leading term in truncation errors in many schemes vanish (see [26, 27]). We compute the solution up to time $t=0.1$ and use the solution computed with 51200 cells over a period as a reference solution. The L^1 -errors (over one period) for both $w=h+B$ and hu are shown in Table 4.1, where one can clearly observe the experimental second order of accuracy.

TABLE 4.1. L^1 -errors and numerical orders of accuracy in Example 1.

Number of grid cells	w		hu	
	L^1 -error	Order	L^1 -error	Order
100	8.89e-03	–	6.63e-02	–
200	3.35e-03	1.41	2.64e-02	1.33
400	1.11e-03	1.59	8.82e-03	1.58
800	3.35e-04	1.73	2.70e-03	1.71
1600	9.30e-05	1.85	7.59e-04	1.83
3200	2.39e-05	1.96	1.96e-04	1.95

Example 2 — small perturbation of a stationary steady-state solution.

Here, we compare the performances of the new well-balanced positivity preserving central-upwind scheme (NEW scheme) and its fore-runner from [12] (OLD scheme). We solve a problem where a small perturbation of a steady-state solution propagates over an almost dry area. Our main goal is to demonstrate that the new techniques (both piecewise linear approximation of the bottom and a special piecewise linear reconstruction for the water surface) are essential for designing a robust central-upwind scheme.

We numerically solve the system (1.1) with gravitational constant $g=1$ and the following bottom topography function:

$$B(x) = \begin{cases} 10(x-0.3), & 0.3 \leq x \leq 0.4, \\ 1 - 0.0025 \sin^2(25\pi(x-0.4)), & 0.4 \leq x \leq 0.6, \\ -10(x-0.7), & 0.6 \leq x \leq 0.7, \\ 0, & \text{otherwise.} \end{cases} \quad (4.1)$$

The initial data,

$$(w(x,0), u(x,0)) = \begin{cases} (1.001, 0), & 0.1 < x < 0.2, \\ (1, 0), & \text{otherwise,} \end{cases} \quad (4.2)$$

is shown in Figure 4.1 and corresponds to a small perturbation (at $[0.1, 0.2]$ with amplitude $\epsilon=0.001$) of the stationary steady-state solution $w \equiv 1, u \equiv 0$.

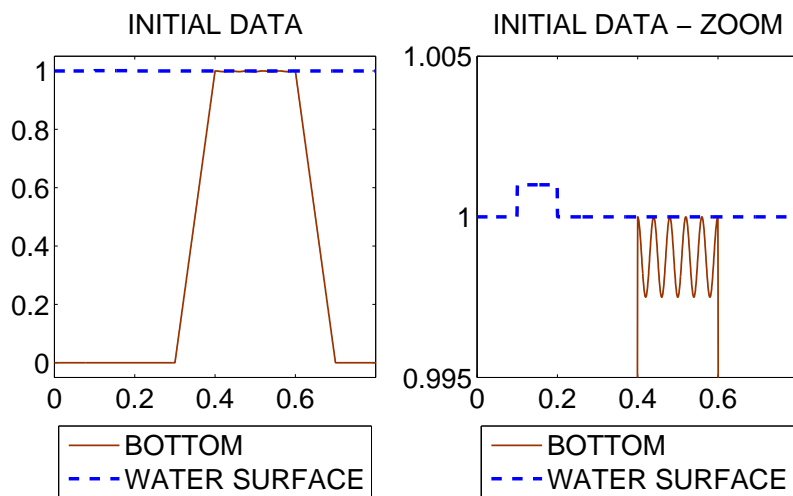


FIG. 4.1. Initial setting (water level w and bottom function B) for the IVP (1.1),(4.1)–(4.2).

With time, the perturbation, initially located at $[0.1, 0.2]$, splits into two pulses moving to opposite directions. The one moving to the right interacts with the nonflat part of the bottom and gets partially reflected. This is illustrated in Figure 4.2, where we plot the solution (water surface w) of the initial value problem IVP (1.1),(4.1)–(4.2), computed at time $t = 1$ by the OLD and NEW schemes using different uniform spatial grids with $\Delta x = 1/200$ and $\Delta x = 1/800$. As one can clearly see, the solution obtained by the OLD scheme is either very oscillatory (on the coarse grid) or not well resolved (on a finer grid). The latter fact can be better observed in Figure 4.3 (left), where we zoom at the right-moving wave front. When the mesh is further refined (see Figure 4.3 (right)), both schemes, as expected, provide similar results. Note, however, that such a fine mesh (with $\Delta x = 1/3200$) is clearly impractical.

REMARK 4.1. *In the examples that follow, we cannot compare the OLD scheme with the NEW scheme, since in these examples the bottom topography is discontinuous, and thus the OLD scheme is not applicable. However, the results obtained in Examples 3 and 4, obtained by the NEW scheme, are comparable with the ones obtained with the alternative methods tested in [1].*

Example 3 — dry state and discontinuous bottom. Our next example is taken from [1], where the bottom topography B is the step function:

$$B(x) = \begin{cases} 2, & x \leq 0.5, \\ 0.1, & x > 0.5, \end{cases} \quad (4.3)$$

the gravitational constant is $g = 2$ and the initial data are:

$$(w(x, 0), u(x, 0)) = \begin{cases} (2.222, -1), & x \leq 0.5, \\ (0.8246, -1.6359), & x > 0.5. \end{cases} \quad (4.4)$$

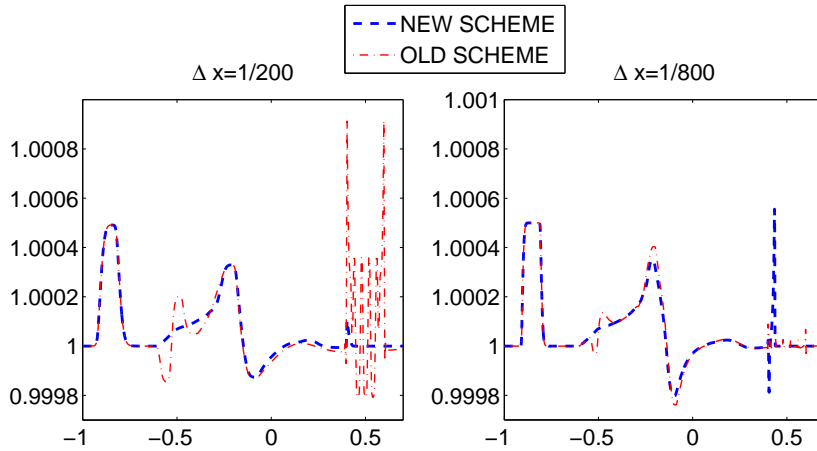


FIG. 4.2. Solutions (water level w) of the IVP (1.1), (4.1)–(4.2) computed by the NEW and OLD schemes. Zoom at the water surface.

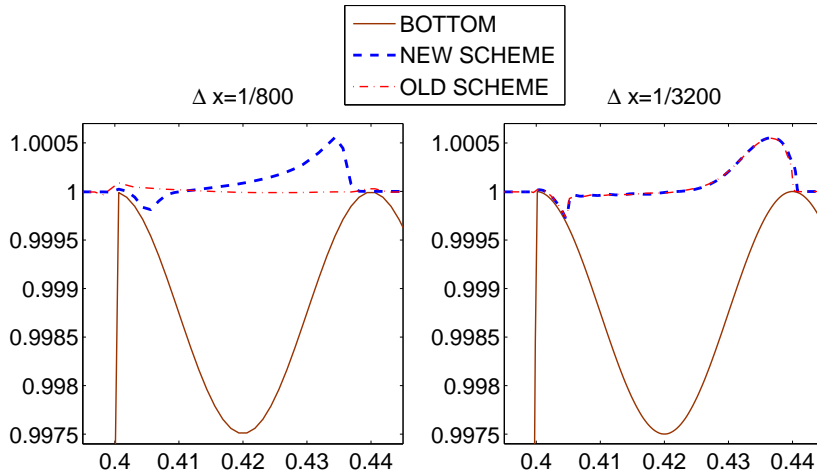
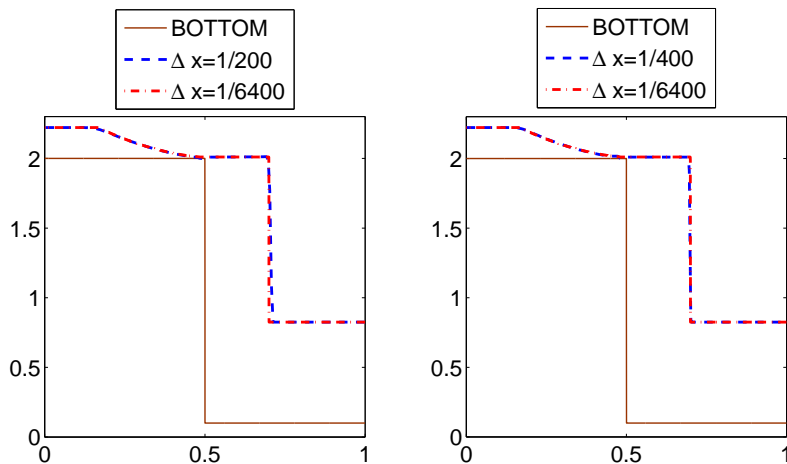
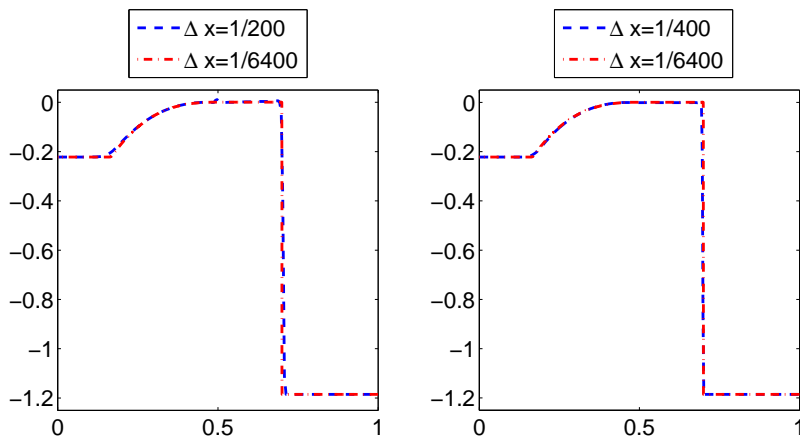


FIG. 4.3. Solutions (water level w) of the IVP (1.1), (4.1)–(4.2) computed by the NEW and OLD schemes. Zoom at the right-going wave front.

The exact solution consists of a rarefaction, followed by a contact discontinuity and a shock.

A numerical solution, computed at time $t=0.2$ using the new well-balanced positivity preserving central-upwind scheme, is presented in Figures 4.4 and 4.5. We show the solutions, computed on a reasonably coarse uniform spatial grids with $\Delta x=1/200$ and $\Delta x=1/400$ and a reference solution, obtained on a very fine mesh with $\Delta x=1/6400$. One can clearly observe a remarkably high resolution achieved by the new scheme.

FIG. 4.4. Solutions (water surface w) of the IVP (1.1),(4.3)–(4.4).FIG. 4.5. Solutions (discharge hu) of the IVP (1.1),(4.3)–(4.4).

Example 4 — composite wave and discontinuous bottom. This test problem is also taken from [1]. As in Example 3, the bottom topography B is the step function:

$$B(x) = \begin{cases} 1.5, & x \leq 0.5, \\ 1.1, & x > 0.5, \end{cases} \quad (4.5)$$

the gravitational constant is $g=2$ and the Riemann initial data are:

$$(w(x,0), u(x,0)) = \begin{cases} (5, 1), & x \leq 0.5, \\ (1.6, -2), & x > 0.5. \end{cases} \quad (4.6)$$

The exact solution of the IVP (1.1),(4.5)–(4.6) contains a composite triple wave, which consists of a sonic rarefaction, attached to a contact, followed by a zero speed shock and another contact. A numerical solution, computed at time $t=0.15$ using the well-

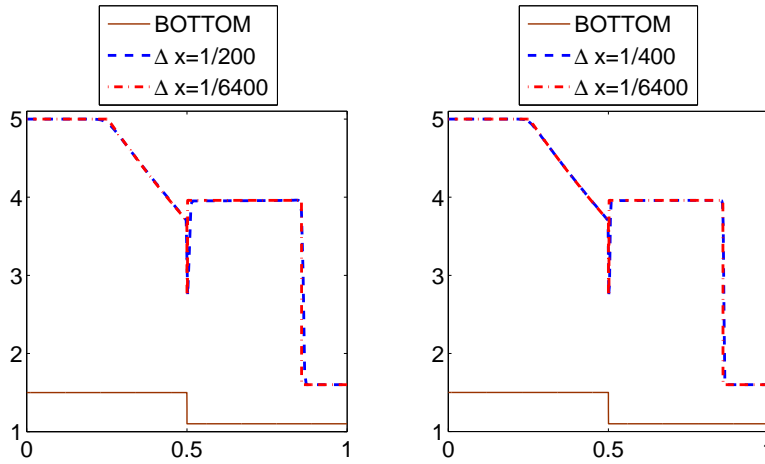


FIG. 4.6. Solutions (water level w) of the IVP (1.1),(4.5)–(4.6).

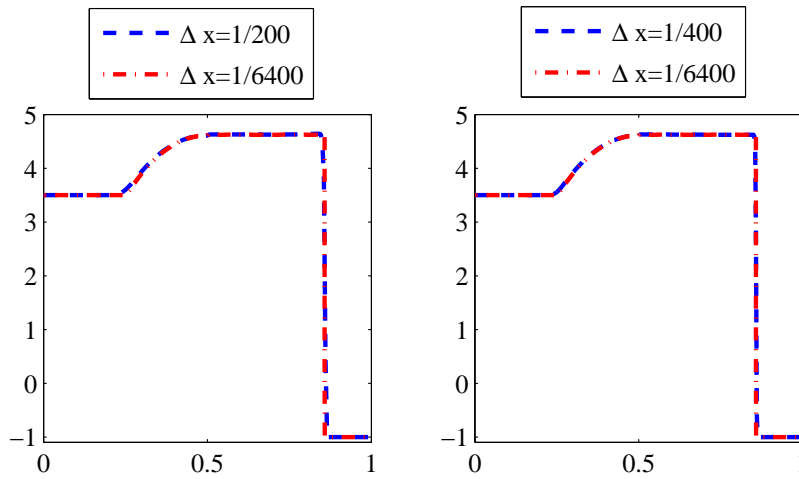
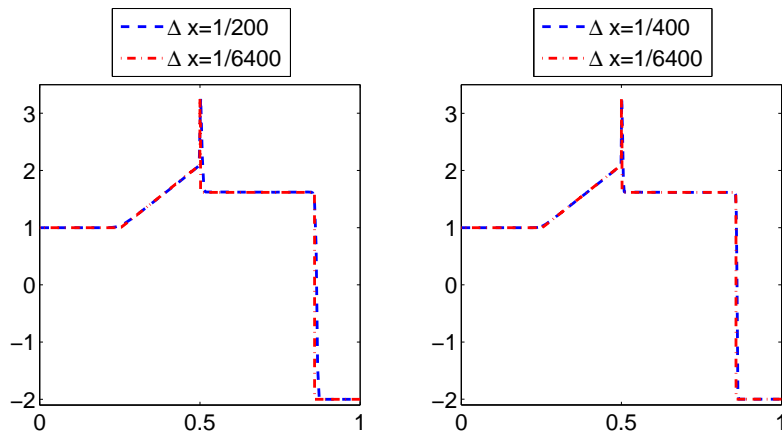


FIG. 4.7. Solutions (discharge hu) of the IVP (1.1),(4.5)–(4.6).

balanced positivity preserving central-upwind scheme, is presented in Figures 4.6–4.8. We show the solutions, computed on a reasonably coarse uniform spatial grids with $\Delta x=1/200$ and $\Delta x=1/400$ and a reference solution, obtained on a very fine mesh with $\Delta x=1/6400$. One can observe a very high overall resolution achieved by our

FIG. 4.8. Solutions (velocity u) of the IVP (1.1),(4.5)–(4.6).

scheme. However, there are significant under- and overshoots in the neighborhood of the bottom discontinuity, especially prominent in the computation of the water surface w (see Figure 4.6) and the velocity u (see Figure 4.8). Note that (according to [1], where several numerical methods were tested) such under- and overshoots are present in all numerical solutions, computed using the schemes from [2, 3, 21].

REMARK 4.2. *An overshoot was originally present in the computed discharge hu as well, but it has been removed by a post-processing technique developed in [9] to eliminate momentum spikes appearing in the computation of slowly moving shocks. Unfortunately, this approach cannot be implemented for the other conservative variable w , and therefore, the obtained w and $u = hu/(w - B)$ are not oscillation-free.*

Example 5 — Saint-Venant system with friction and discontinuous bottom. It is a well-known fact that more realistic shallow water models based on the Saint-Venant system (1.1) should include additional friction and/or viscosity terms. In [5], such models were derived from the Navier-Stokes equations for incompressible flows with a free moving boundary. Presence of friction and viscosity terms guarantees uniqueness of the steady state solution, especially in the case when the fluid propagates into a certain region and gradually occupies parts of initially dry areas, as, for example, in Figures 4.9 and 4.10.

We consider the simplest model, in which only a friction term, $-\kappa(h)u$, is added to the RHS of the second equation in (1.1):

$$\begin{cases} h_t + (hu)_x = 0, \\ (hu)_t + \left(hu^2 + \frac{1}{2}gh^2\right)_x = -ghB' - \kappa(h)u. \end{cases} \quad (4.7)$$

We numerically solve the system (4.7) assuming that the gravitational constant has value $g = 1$, the friction coefficient is $\kappa(h) = 0.001(1 + 10h)^{-1}$, and the bottom topog-

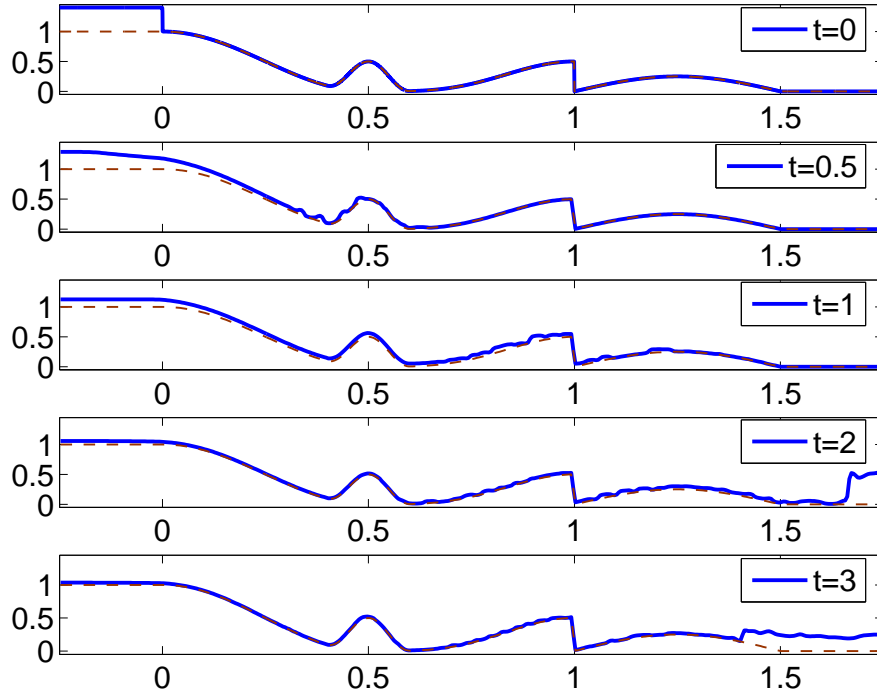


FIG. 4.9. Solution of the IBVP (4.7)–(4.9) computed by the well-balanced positivity preserving central-upwind scheme: water level w (solid line) and the interpolated bottom B (dashed line).

raphy function has a discontinuity at $x = 1$ and is given by:

$$B(x) = \begin{cases} 1, & x < 0, \\ \cos^2(\pi x), & 0 \leq x \leq 0.4, \\ \cos^2(\pi x) + 0.25(\cos(10\pi(x - 0.5)) + 1), & 0.4 \leq x \leq 0.5, \\ 0.5\cos^4(\pi x) + 0.25(\cos(10\pi(x - 0.5)) + 1), & 0.5 \leq x \leq 0.6, \\ 0.5\cos^4(\pi x), & 0.6 \leq x < 1, \\ 0.25\sin(2\pi(x - 1)), & 1 < x \leq 1.5, \\ 0, & x > 1.5. \end{cases} \quad (4.8)$$

We take the following initial data:

$$(w(x,0), u(x,0)) = \begin{cases} (1.4, 0), & x < 0, \\ (B(x), 0), & x > 0, \end{cases} \quad (4.9)$$

and implement solid wall boundary conditions. These data correspond to the situation when the second of the three dams, initially located at $x = -0.25$ (the left boundary of the computational domain), 0, and 1.75 (the right boundary of the computational domain), breaks down at time $t = 0$, the water propagates into the initially dry area $[0, 1.75]$, and a stationary steady state is achieved after a certain period of time.

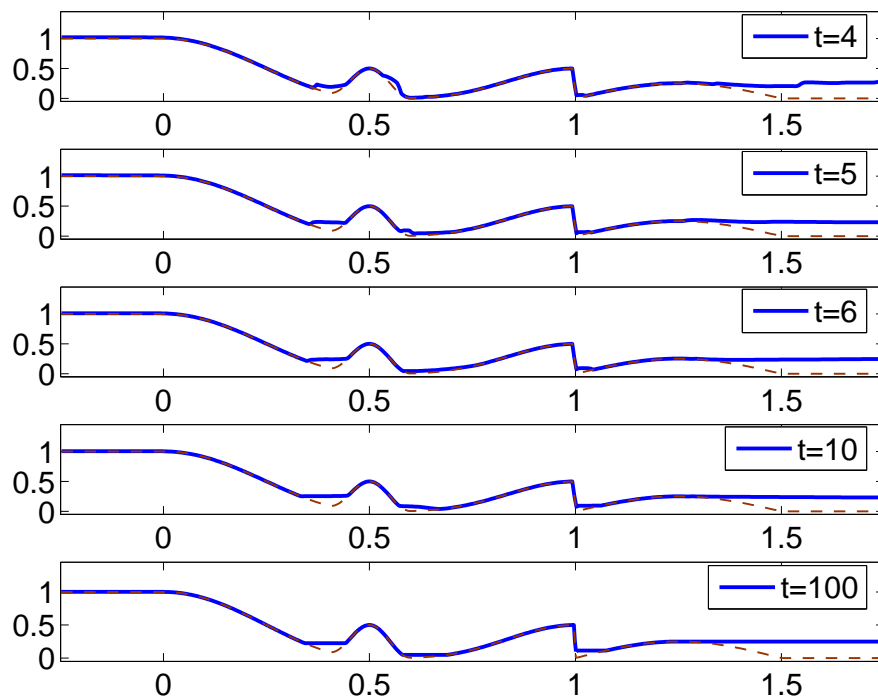


FIG. 4.10. The same as in Figure 4.9, but at larger times.

We apply the 1-D well-balanced positivity preserving central-upwind scheme (2.2),(2.5),(2.8), (2.10)–(2.23) to this initial-boundary value problem (IBVP). The friction term in (4.7) is discretized in a straightforward manner. Notice that this affects neither the well-balanced (since $u \equiv 0$ at stationary steady states) nor the positivity preserving (since the first equation has not been changed) properties of our scheme.

The solution of the IBVP (4.7)–(4.9) at times $t=0, 0.5, 1, 2, 3, 4, 5, 6$, and $t=10$ is computed using a uniform grid with $\Delta x = 1/200$, while the solution at a very large time $t=100$, when the stationary steady-state is practically achieved, is obtained with $\Delta x = 1/400$. These solutions are shown in Figures 4.9 and 4.10, where one can clearly see the dynamics of the fluid flow as it moves from the region $[-0.25, 0]$ into the initially dry area $[0, 1.75]$ and gradually settles down into a stationary steady state. Notice that this state includes dry areas and therefore its computation requires a method that is both well balanced and positivity preserving on the entire computational domain.

4.2. Two-dimensional example. Here, we use the new 2-D well-balanced positivity preserving central-upwind scheme to numerically solve the 2-D system (3.1) in a domain $[-1, 1] \times [-1, 1]$. The gravitational constant is $g=2$, the bottom topogra-

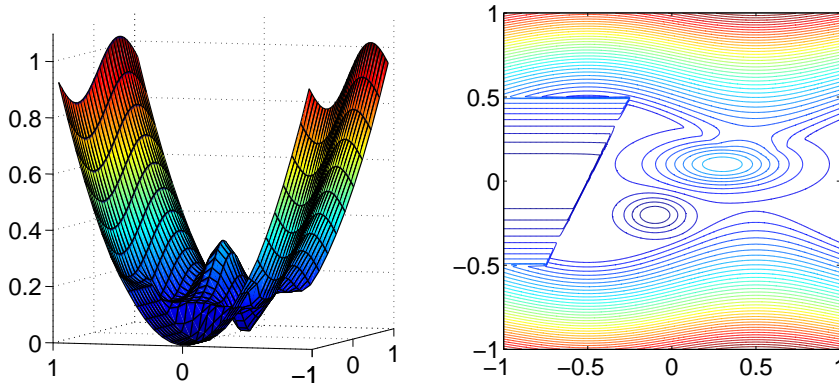


FIG. 4.11. River valley topography for the IBVP (3.1),(4.10)–(4.11), 3-D view (left) and contour plot (right).

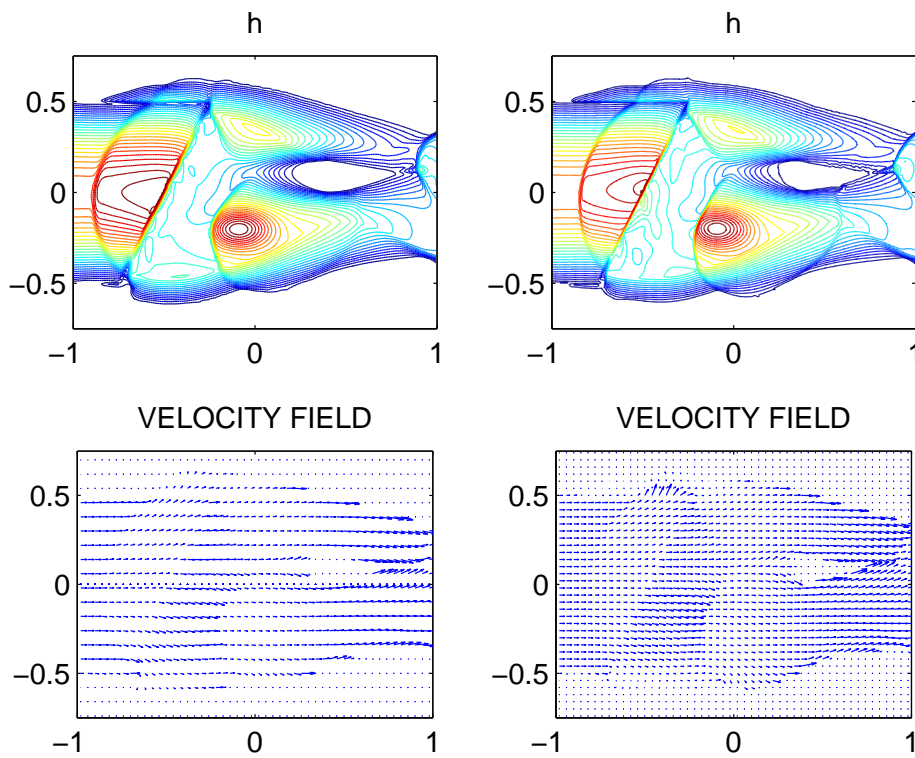


FIG. 4.12. Solution of the IBVP (3.1),(4.10)–(4.11): $t=1$; $\Delta x=\Delta y=1/100$ (left) and $\Delta x=\Delta y=1/200$ (right).

phy function is given by:

$$\begin{aligned}
 B(x,y) &= \frac{7}{32}e^{-8(x-0.3)^2-60(y-0.1)^2} - \frac{1}{8}e^{-30(x+0.1)^2-90(y+0.2)^2} \\
 &+ \begin{cases} y^2, & |y| \leq \frac{1}{2}, x \leq \frac{y-1}{2}, \\ y^2 + \frac{1}{10}\sin(\pi x), & |y| > \frac{1}{2}, x \leq \frac{y-1}{2}, \\ \max\{\frac{1}{8}, y^2 + \frac{1}{10}\sin(\pi x)\}, & x > \frac{y-1}{2}, \end{cases} \quad (4.10)
 \end{aligned}$$

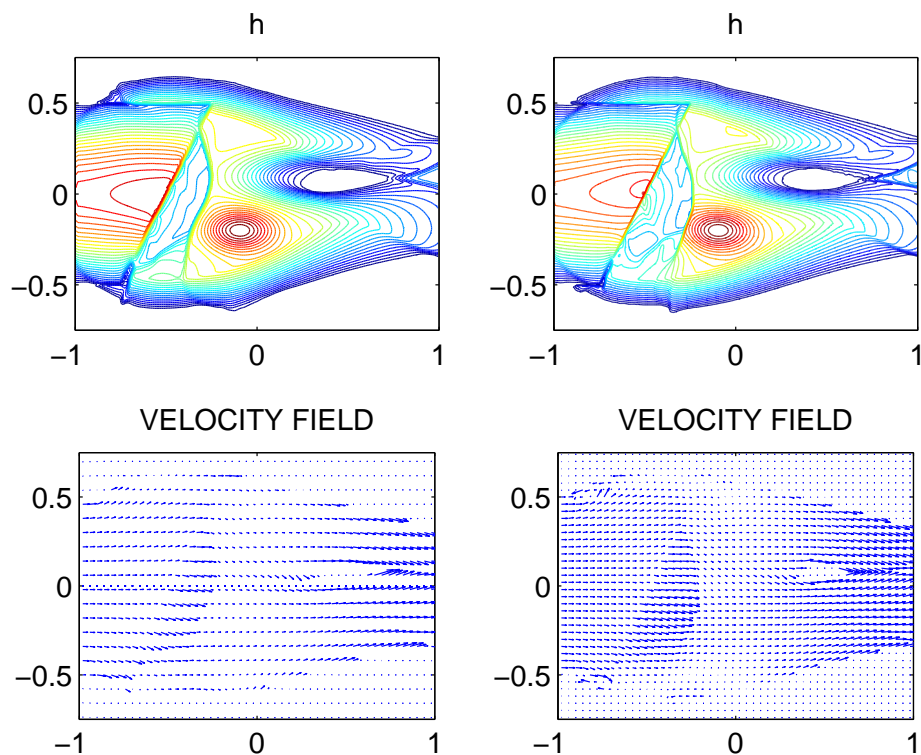


FIG. 4.13. Solution of the IBVP (3.1),(4.10)–(4.11): $t=2$; $\Delta x = \Delta y = 1/100$ (left) and $\Delta x = \Delta y = 1/200$ (right).

and the initial data are:

$$w(x, y, 0) = \max \left\{ \frac{1}{4}, B(x, y) \right\}, \quad v(x, y, 0) \equiv 0, \quad u(x, y, 0) = \begin{cases} \frac{1}{2}, & |y| \leq \frac{1}{2}, \\ 0, & \text{otherwise.} \end{cases} \quad (4.11)$$

We implement inflow boundary conditions at the left boundary and outflow boundary conditions at the right boundary.

This example describes a fluid flow in a mountain river valley, which, together with the surrounding mountains, is shown in Figure 4.11. The valley has a discontinuity across the line $y = 2x + 1$, a Gaussian-shaped hole, centered at $(-0.1, -0.2)$, and a Gaussian-shaped island, centered at $(0.3, 0.1)$. Initially, the water surface is flat and the river flows from left to the right with a constant velocity, as stated in (4.11).

Due to the nontrivial geometry of the bottom topography in this IBVP, the dynamics of the flow is quite complicated. The numerical solutions, computed on uniform grids with $\Delta x = \Delta y = 1/100$ and $\Delta x = \Delta y = 1/200$ at times $t = 1, 2$, and 4 , are shown in Figures 4.12–4.14, respectively. One can observe that the solutions computed using the two different grids are in good agreement.

Notice that the IBVP (3.1),(4.10)–(4.11) mimics a realistic situation one may face when simulating river flows in mountains. The solution has dry areas (the island and the surrounding mountain areas) and thus the positivity preserving property

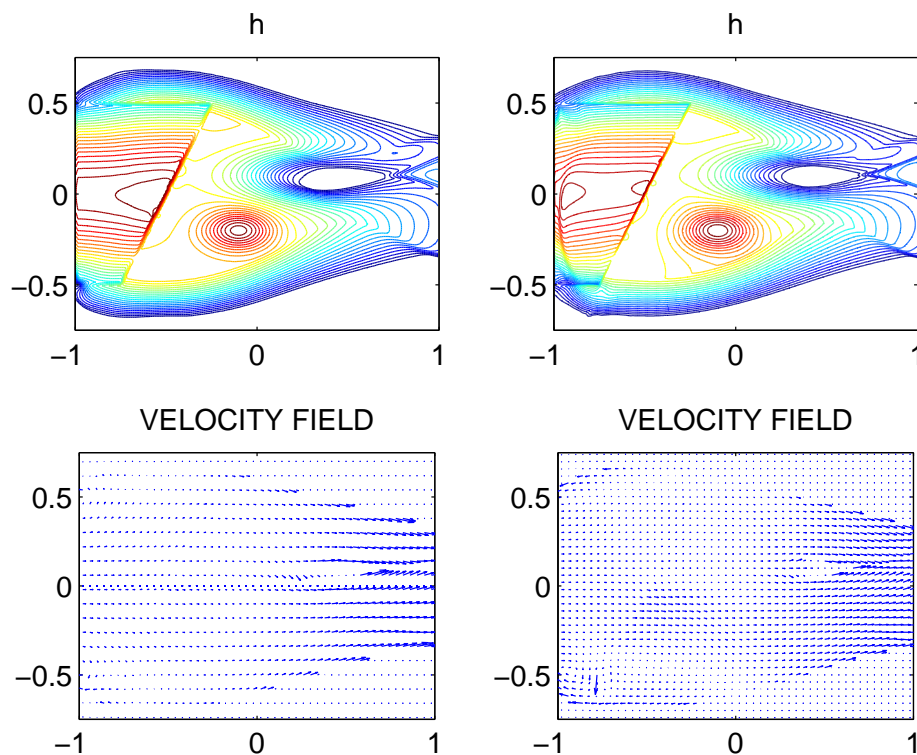


FIG. 4.14. Solution of the IBVP (3.1),(4.10)–(4.11): $t=4$; $\Delta x = \Delta y = 1/100$ (left) and $\Delta x = \Delta y = 1/200$ (right).

of our scheme is essential in the computation of the flow. At the same time, the nature of the bottom topography requires a numerical method that is capable to treat discontinuous bottom functions. Our new scheme is an excellent candidate that satisfies both aforementioned properties.

Acknowledgement. The research of A. Kurganov was supported in part by the NSF Grants # DMS-0310585 and DMS-0610430. The research of G. Petrova was supported in part by the NSF Grant # DMS-0505501.

REFERENCES

- [1] N. Andrianov, *Testing numerical schemes for the shallow water equations*, preprint, 2004, available at <http://www-ian.math.uni-magdeburg.de/home/andriano/CONSTRUCT/testing.ps.gz>.
- [2] 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, 2050-2065, 2004.
- [3] F. Bouchut, *Nonlinear Stability of Finite Volume Methods for Hyperbolic Conservation Laws and Well-balanced Schemes for Sources*, Frontiers in Math Series, Birkhäuser Verlag, Basel, 2004.
- [4] T. Gallouët, J.-M. Hérard and N. Seguin, *Some approximate Godunov schemes to compute shallow-water equations with topography*, Comput. & Fluids, 32, 479-513, 2003.

- [5] J.-F. Gerbeau and B. Perthame, *Derivation of viscous Saint-Venant system for laminar shallow water; numerical validation*, Discrete Contin. Dyn. Syst. Ser. B, 1, 89-102, 2001.
- [6] E. Godlewski and P.-A. Raviart, *Numerical Approximation of Hyperbolic Systems of Conservation Laws*, Springer-Verlag, New York, 1996.
- [7] S. Gottlieb, C.-W. Shu and E. Tadmor, *High order time discretization methods with the strong stability property*, SIAM Review, 43, 89-112, 2001.
- [8] S. Jin, *A steady-state capturing method for hyperbolic system with geometrical source terms*, M2AN Math. Model. Numer. Anal., 35, 631-645, 2001.
- [9] S. Jin and J.-G. Liu, *The effects of numerical viscosities. I. Slowly moving shocks*, J. Comput. Phys., 126, 373-389, 1996.
- [10] S. Jin and X. Wen, *Two interface-type numerical methods for computing hyperbolic systems with geometrical source terms having concentrations*, SIAM J. Sci. Comput., 26, 2079-2101, 2005.
- [11] D. Kröner, *Numerical Schemes for Conservation Laws*, Wiley, Chichester, 1997.
- [12] A. Kurganov and D. Levy, *Central-upwind schemes for the Saint-Venant system*, M2AN Math. Model. Numer. Anal., 36, 397-425, 2002.
- [13] A. Kurganov, S. Noelle and G. Petrova, *Semi-discrete central-upwind schemes for hyperbolic conservation laws and Hamilton-Jacobi equations*, SIAM J. Sci. Comput., 23, 707-740, 2001.
- [14] A. Kurganov and E. Tadmor, *New high-resolution central schemes for nonlinear conservation laws and convection-diffusion equations*, J. Comput. Phys., 160, 214-282, 2000.
- [15] B. van Leer, *Towards the ultimate conservative difference scheme, V. A second order sequel to Godunov's method*, J. Comput. Phys., 32, 101-136, 1979.
- [16] R. J. LeVeque, *Balancing source terms and flux gradients in high-resolution Godunov methods: the quasi-steady wave-propagation algorithm*, J. Comput. Phys., 146, 346-365, 1998.
- [17] R. LeVeque, *Finite Volume Methods for Hyperbolic Problems*, Cambridge Texts in Applied Mathematics, Cambridge University Press, 2002.
- [18] K.-A. Lie and S. Noelle, *On the artificial compression method for second-order nonoscillatory central difference schemes for systems of conservation laws*, SIAM J. Sci. Comput., 24, 1157-1174, 2003.
- [19] H. Nessyahu and E. Tadmor, *Non-oscillatory central differencing for hyperbolic conservation laws*, J. Comput. Phys., 87, 408-463, 1990.
- [20] S. Noelle, N. Pankratz, G. Puppo and J. Natvig, *Well-balanced finite volume schemes of arbitrary order of accuracy for shallow water flows*, J. Comput. Phys., 213, 474-499, 2006.
- [21] B. Perthame and C. Simeoni, *A kinetic scheme for the Saint-Venant system with a source term*, Calcolo, 38, 201-231, 2001.
- [22] G. Russo, *Central schemes for balance laws*, Hyperbolic problems: theory, numerics, applications, Magdeburg, I, II, 821-829, 2000. Internat. Ser. Numer. Math., 140, 141, Birkhäuser, Basel, 2001.
- [23] G. Russo, *Central schemes for conservation laws with application to shallow water equations*, Trends and applications of mathematics to mechanics: STAMM 2002, S. Rionero and G. Romano (eds.), Springer-Verlag Italia SRL, 225-246, 2005.
- [24] A. J. C. de Saint-Venant, *Théorie du mouvement non-permanent des eaux, avec application aux crues des rivières et à l'introduction des vagues dans leur lit*, C. R. Acad. Sci. Paris, 73, 147-154, 1871.
- [25] P. K. Sweby, *High resolution schemes using flux limiters for hyperbolic conservation laws*, SIAM J. Numer. Anal., 21, 995-1011, 1984.
- [26] S. Vukovic and L. Sopta, *ENO and WENO schemes with the exact conservation property for one-dimensional shallow water equations*, J. Comput. Phys., 179, 593-621, 2002.
- [27] Y. Xing and C.-W. Shu, *High order finite difference WENO schemes with the exact conservation property for the shallow water equations*, J. Comput. Phys., 208, 206-227, 2005.
- [28] Y. Xing and C.-W. Shu, *A new approach of high order well-balanced finite volume WENO schemes and discontinuous Galerkin methods for a class of hyperbolic systems with source terms*, Commun. Comput. Phys., 1, 100-134, 2006.