SIAM J. SCI. COMPUT.
Vol. 42, No. 4, pp. A2206–A2229

MOVING-WATER EQUILIBRIA PRESERVING PARTIAL RELAXATION SCHEME FOR THE SAINT-VENANT SYSTEM*

XIN LIU†, XI CHEN‡, SHI JIN§, ALEXANDER KURGANOV¶, TONG WU‖, AND HUI YU#

Abstract. We develop a new moving-water equilibria preserving numerical scheme for the Saint-Venant system. The new scheme is designed in two major steps. First, the geometric source term is incorporated into the discharge flux, which results in a hyperbolic system with a global flux. Second, the discharge equation is relaxed so that the nonlinearity is moved into the stiff right-hand side of the added auxiliary equation. The main advantages of the new scheme are that (i) no special treatment of the geometric source term is required, and (ii) no nonlinear (cubic) equations should be solved to obtain the point values of the water depth out of the reconstructed equilibrium variables, as it must be done in the existing alternative methods. We also develop a hybrid numerical flux, which helps to handle various flow regimes in a stable manner. Several numerical experiments are performed to verify that the proposed scheme is capable of exactly preserving general moving-water steady states and accurately capturing their small perturbations.

Key words. Saint-Venant system of shallow water equations, partial relaxation scheme, well-balanced method, steady-state solutions (equilibria), moving-water and still-water equilibria

AMS subject classifications. 76M12, 65M08, 35L65, 86-08, 86A05

DOI. 10.1137/19M1258098

1. Introduction.

The Saint-Venant system of shallow water equations proposed in [16] has been widely used to predict fluid flows in estuaries, oceans, coastal regions, lakes, rivers, and channels. It generally describes a depth-averaged thin layer free-surface flow of constant density under hydrostatic assumption over a rigid bottom. In the one-dimensional (1-D) case, the Saint-Venant system over a frictional bottom topography reads as

$$\begin{cases} h_t + q_x = 0, \\ q_t + \left(hu^2 + \dfrac{g}{2}h^2\right)_x = -ghB_x - g\dfrac{n^2}{h^{7/3}}|q|q, \end{cases} \tag{1.1}$$

where $h(x,t)$ is the water depth, $u(x,t)$ is the depth-averaged velocity, $q(x,t) := h(x,t)u(x,t)$ is the flow discharge, g is the constant gravitational acceleration, $B(x)$

*Submitted to the journal's Methods and Algorithms for Scientific Computing section April 24, 2019; accepted for publication (in revised form) March 31, 2020; published electronically July 20, 2020.

https://doi.org/10.1137/19M1258098

Funding: The work of the third author was partially supported by the NSFC under grants 31571071 and 11871297. The work of the fourth author was partially supported by the NSFC under grant 11771201 and by the Guangdong Provincial Key Laboratory of Computational Science and Material Design under grant 2019B030301001.

†Department of Mathematics, Southern University of Science and Technology, Shenzhen, 518055, China, and Numerical Environmental Prediction Section, Canadian Meteorological Centre, Environment and Climate Change Canada, Dorval, QC, H9P 1J3, Canada (xliu111@uottawa.ca).

‡Department of Mathematics, Harbin Institute of Technology, Harbin, 150001, China, and Department of Mathematics, Southern University of Science and Technology, Shenzhen, 518055, China (chenx2018@mail.sustech.edu.cn).

§School of Mathematical Sciences, Institute of Natural Sciences, MOE-LSC, Shanghai Jiao Tong University, Shanghai, 200240, China (shijin-m@sjtu.edu.cn).

¶Department of Mathematics and SUSTech International Center for Mathematics, Southern University of Science and Technology, Shenzhen, 518055, China (alexander@sustech.edu.cn).

‖Mathematics Department, Tulane University, New Orleans, LA 70118 (twu2@tulane.edu).

#Yau Mathematical Science Center, Tsinghua University, Beijing, 100084, China (huiyu@tsinghua.edu.cn).

is the bottom topography, and n is the Manning friction coefficient of the bottom.

The system (1.1) is a hyperbolic system of balance laws, which admits both smooth and nonsmooth solutions. A special class of the solutions of (1.1) are steady-state solutions at which $h_t \equiv 0$ and $q_t \equiv 0$ due to the fact that the flux gradient in the second equation in (1.1) is exactly balanced by the geometric and friction source terms. Two major classes of steady states are generally considered in the water flow computations. The first class consists of the "lake at rest" (still-water) steady states with the zero flow velocity

$$q \equiv 0, \quad h + B \equiv \text{Const}. \tag{1.2}$$

The second class of steady states consists of general moving-water steady states at which $q \not\equiv 0$. If the bottom friction is neglected, the general smooth moving-water equilibria are given by

$$q \equiv \text{Const}, \quad E := \frac{u^2}{2} + g(h + B) \equiv \text{Const}. \tag{1.3}$$

For some particular (moving-water) ($q \not\equiv 0$) steady states, we refer the reader to [13] and references therein. Notice that the still-water steady state (1.2) is a special case of the general moving-water steady state (1.3) with $u \equiv 0$.

Steady states are of great practical importance since many physically relevant solutions of (1.1) are, in fact, small perturbations of steady states. Numerically capturing such solutions is a challenging task since straightforward application of shock-capturing methods may lead to spurious oscillations. Such nonphysical oscillations may strongly disturb the simulation of the physical waves unless a very fine mesh is used, which, in turn, causes high, often unaffordable computational cost. Therefore, it is necessary to design well-balanced numerical methods which guarantee that the discretized numerical flux gradient is exactly balanced by the approximated source terms at both the steady states (1.2) and (1.3), so that the steady-state solutions can be exactly preserved independently of the grid size.

For still-water equilibria preserving numerical methods, the well-balanced property hinges on a special approximation of the bed-slope source term $-ghB_x$. We refer the readers to the nonexhaustive list of references on still-water equilibria preserving numerical methods [1, 2, 3, 18, 20, 21, 25, 26, 28, 29, 33]. However, these methods fail to accurately capture flows, which are at or near the moving-water equilibria.

Preserving moving-water equilibria is substantially more complicated. The main difficulty is related to the fact that well-balanced approximations of the geometric sources now need to include terms that are small for smooth solutions but may become artificially large at discontinuities. Among the papers in which moving-water steady-state-preserving methods were developed, the schemes in [5, 6, 9, 10, 13, 24, 25, 32, 34, 35] require nontrivial root-finding based on the energy balance, and some methods require substantial effort on local reconstruction [4, 5, 6, 19] or complex reconstruction of the geometric source terms [7, 32, 40]. The motivation of the current study is to avoid aforementioned complexities in designing a moving-water steady-state-preserving method.

In order to avoid complex discretization of geometric source term, we follow the idea from [9, 11, 14] and incorporate the source terms of the discharge equation into its flux and rewrite (1.1) in the following equivalent form:

$$\begin{cases} h_t + q_x = 0, \\ q_t + K_x = 0, \end{cases} \tag{1.4}$$

where

$$K := hu^2 + \frac{g}{2}h^2 + R, \tag{1.5}$$

so that K is a global equilibrium variable with

$$R(x,t) := g\int^x \left[h(\xi,t)B_x(\xi) + \frac{n^2}{h^{7/3}(\xi,t)}|q(\xi,t)|q(\xi,t)\right]\mathrm{d}\xi. \tag{1.6}$$

Accordingly, the general (moving-water) steady states can be expressed in terms of q and K as

$$q \equiv \text{Const}, \quad K \equiv \text{Const}.$$

We note that an idea of incorporating the source terms into the fluxes has also recently been used in the derivation of well-balanced schemes for both 1-D and two-dimensional (2-D) Euler equations with gravitation [11].

In a moving-water equilibria preserving central-upwind scheme, recently presented in [9], the equilibrium variables K and q are reconstructed to obtain the point values of K and q needed to evaluate the numerical fluxes at cell interfaces. In addition, one needs to compute the point values of h there. This is done by solving cubic equations at each cell interface at every time step. This is not only computationally costly but also quite challenging near the sonic points; also see [10, 24, 32, 41]. Moreover, the system (1.4)–(1.6) is a hyperbolic system with a global flux, which makes the development of an upwind scheme based on the solution of (generalized) Riemann problems difficult or even impossible. In [17], it was proposed to build a relaxation model for both equations in (1.4) so that one can easily develop an upwind scheme for the linear part of the relaxation system. However, the numerical methods proposed in [17] are not well-balanced.

In order to avoid the aforementioned numerical difficulties, we develop a novel numerical scheme using a partial relaxation technique for the second equation in (1.4). The proposed numerical scheme is well-balanced without any need of a special treatment for the source terms, so that it is capable of exactly preserving both the still-water (1.2) and moving-water (1.3) steady states. Our scheme is based on the system (1.4)–(1.6) with a global flux, while solving cubic equations for computing the point values of h is avoided. In order to deal with different flow regimes, we propose to use the upwind method for the subcritical flow and the central upwind method for the trans- and supercritical flows at each cell interface. A piecewise linear reconstruction with a generalized minmod limiter and a second-order steady state and sign preserving semi-implicit Runge–Kutta ODE solver from [12] are used in the developed numerical method to achieve the second-order accuracy in both space and time.

The paper is organized as follows. A new partial relaxation approximation of the Saint-Venant system is proposed in section 2. The numerical scheme based on the hybrid numerical fluxes is derived in section 3. The new numerical scheme is tested on a variety of numerical examples in section 4. Some concluding remarks can be found in section 5.

2. Partial relaxation approximation of (1.4)–(1.6). In this section, we present a relaxation approximation of the Saint-Venant system (1.4)–(1.6).

2.1. A brief overview of relaxation schemes. We begin with a review of relaxation schemes, which were originally proposed in [22] for systems of conservation laws. These schemes are developed by constructing a linear hyperbolic system with a

stiff lower-order term that approximates the original system with a small dissipative correction. For the 1-D conservation law of the form

$$\boldsymbol{U}_t + \boldsymbol{F}(\boldsymbol{U})_x = \boldsymbol{0}, \tag{2.1}$$

one can introduce the relaxation approximation

$$\begin{aligned} &\boldsymbol{U}_t + \boldsymbol{P}_x = \boldsymbol{0}, \\ &\boldsymbol{P}_t + a^2 \boldsymbol{U}_x = -\frac{1}{\varepsilon}(\boldsymbol{P} - \boldsymbol{F}(\boldsymbol{U})), \end{aligned} \tag{2.2}$$

which is equivalent to the original system (2.1) in the small relaxation limit when the relaxation parameter $\varepsilon \to 0^+$. In (2.2), $\boldsymbol{P}$ is an introduced auxiliary variable and a is a positive constant satisfying the subcharacteristic condition $a \geq \max_{\boldsymbol{U}} \rho(\partial \boldsymbol{F}(\boldsymbol{U})/\partial \boldsymbol{U})$, where ρ denotes the spectral radius of the matrix.

In [17], a slightly modified version of the relaxation scheme (2.2), applied to the Saint-Venant system (1.4)–(1.6), resulted in the following relaxation approximation:

$$\begin{aligned} &\boldsymbol{U}_t + \boldsymbol{P}_x = \boldsymbol{0}, \\ &\boldsymbol{P}_t + A^2 \boldsymbol{U}_x = -\frac{1}{\varepsilon}\left(\boldsymbol{P} - \boldsymbol{F}(\boldsymbol{U})\right), \end{aligned} \tag{2.3}$$

where

$$\boldsymbol{U} = \begin{pmatrix} h \\ q \end{pmatrix}, \quad \boldsymbol{P} = \begin{pmatrix} w \\ v \end{pmatrix}, \quad A = \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix}, \quad \boldsymbol{F}(\boldsymbol{U}) = \begin{pmatrix} q \\ K \end{pmatrix}, \tag{2.4}$$

in which the relaxation source term rapidly drives $\boldsymbol{P} \to \boldsymbol{F}(\boldsymbol{U})$ in the limit as $\varepsilon \to 0^+$. One can thus build an upwind semidiscrete approximation of the relaxation system (2.3), (2.4). Its first-order form reads as

$$\frac{d}{\mathrm{d}t}\overline{\boldsymbol{U}}_j + \frac{\overline{\boldsymbol{P}}_{j+1} - \overline{\boldsymbol{P}}_{j-1}}{2\Delta x} = \frac{\Delta x}{2} \cdot A \, \frac{\overline{\boldsymbol{U}}_{j+1} - 2\overline{\boldsymbol{U}}_j + \overline{\boldsymbol{U}}_{j-1}}{(\Delta x)^2}, \tag{2.5}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\overline{\boldsymbol{P}}_j + A^2 \, \frac{\overline{\boldsymbol{U}}_{j+1} - \overline{\boldsymbol{U}}_{j-1}}{2\Delta x} = -\frac{1}{\varepsilon}\left(\overline{\boldsymbol{P}}_j - \boldsymbol{F}(\overline{\boldsymbol{U}}_j)\right) + \frac{\Delta x}{2} \cdot \frac{\overline{\boldsymbol{P}}_{j+1} - 2\overline{\boldsymbol{P}}_j + \overline{\boldsymbol{P}}_{j-1}}{(\Delta x)^2}, \tag{2.6}$$

where $\overline{(\cdot)}$ denotes the cell averages, which will be introduced in the beginning of section 3.

Unfortunately, the numerical scheme (2.5), (2.6) as well as its higher-order modifications introduced in [17] are not well-balanced since the dissipative terms on the right-hand side (RHS) of (2.5) and (2.6) do not vanish at steady states and thus introduce spurious waves.

2.2. Partial relaxation system. In order to construct a well-balanced relaxation scheme for the system (1.4)–(1.6), we introduce the following *partial relaxation system*, in which we only relax the second equation in (1.4):

$$h_t + q_x = 0, \tag{2.7}$$

$$q_t + v_x = 0, \tag{2.8}$$

$$v_t + a^2 q_x = -\frac{1}{\varepsilon}(v - K), \tag{2.9}$$

in which v is an auxiliary variable. In the numerical results reported in section 4, we have used $\varepsilon = 10^{-6}$. Such a relaxation is similar to Suliciu's relaxation [36], which also found other applications; see, e.g., [15].

In the small relaxation limit ($\varepsilon \to 0^+$), equations (2.8) and (2.9) can be approximated to leading order by

$$v = K, \quad q_t + K_x = 0, \tag{2.10}$$

and one can find that the local equilibrium (2.10) is the momentum equation in (1.4).

It is well known that the relaxation approximation is dissipative and thus well-posed provided a is sufficiently large, that is, it satisfies a subcharacteristic condition which is derived below. In order to establish a proper subcharacteristic condition for the partial relaxation system (2.7)–(2.9), we consider a simple case with $B \equiv \text{Const}$ and $n = 0$ (that is, the case in which the RHS of (1.1) vanishes) in our analysis. In this case,

$$K = hu^2 + \frac{g}{2}h^2 = \frac{q^2}{h} + \frac{g}{2}h^2, \tag{2.11}$$

and thus

$$K_x = 2uq_x - (u^2 - gh)h_x. \tag{2.12}$$

Next, using (2.10)–(2.12), we obtain

$$K_t = 2uq_t - (u^2 - gh)h_t = -2uK_x - (u^2 - gh)h_t = 2u(u^2 - gh)h_x - (3u^2 + gh)q_x.$$

Using the Chapman–Enskog expansion [8], one can obtain the following first-order approximation of the proposed partial relaxation system (2.7)–(2.9):

$$v = K - \varepsilon\left[2u(u^2 - gh)h_x + (a^2 - 3u^2 - gh)q_x\right] + \mathcal{O}(\varepsilon^2),$$

which can then be substituted into (2.8) to eliminate the auxiliary variable v. This results in

$$\begin{pmatrix} h \\ q \end{pmatrix}_t + \begin{pmatrix} q \\ K \end{pmatrix}_x = \varepsilon\left[\begin{pmatrix} 0 & 0 \\ 2u(u^2 - gh) & a^2 - 3u^2 - gh \end{pmatrix}\begin{pmatrix} h \\ q \end{pmatrix}_x\right]_x + \begin{pmatrix} 0 \\ \mathcal{O}(\varepsilon^2) \end{pmatrix}, \tag{2.13}$$

which introduces a higher-order dissipative perturbation of the system (1.4)–(1.6). As one can easily see, the system (2.13) is dissipative provided the following subcharacteristic condition is satisfied:

$$a > \sqrt{3u^2 + gh}. \tag{2.14}$$

3. Numerical scheme. In this section, we develop a semidiscrete numerical scheme for the partially relaxed system (2.7)–(2.9), which can be rewritten in the following vector form:

$$\boldsymbol{U}_t + \boldsymbol{F}(\boldsymbol{U})_x = \boldsymbol{S}(\boldsymbol{U}), \tag{3.1}$$

where

$$\boldsymbol{U} = \begin{pmatrix} h \\ q \\ v \end{pmatrix}, \quad \boldsymbol{F}(\boldsymbol{U}) = \begin{pmatrix} q \\ v \\ a^2 q \end{pmatrix}, \quad \boldsymbol{S}(\boldsymbol{U}) = \begin{pmatrix} 0 \\ 0 \\ -\dfrac{1}{\varepsilon}(v - K) \end{pmatrix}. \tag{3.2}$$

The computational domain is discretized by using the finite-volume cells $I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$ of size Δx centered at $x_j = \big(x_{j-\frac{1}{2}} + x_{j+\frac{1}{2}}\big)/2$ with $j = j_\ell, \dots, j_r$. We assume that the cell averages of $\boldsymbol{U}$,

$$\overline{\boldsymbol{U}}_j(t) \approx \frac{1}{\Delta x}\int_{I_j} \boldsymbol{U}(x,t)\,\mathrm{d}x, \tag{3.3}$$

are available at a certain time level t and define the cell averages of $\boldsymbol{S}$ by

$$\overline{\boldsymbol{S}}_j(t) \approx \frac{1}{\Delta x}\int_{I_j} \boldsymbol{S}(\boldsymbol{U}(x,t))\,\mathrm{d}x.$$

We note that in the semidiscrete framework, all of the indexed quantities that depend on $\boldsymbol{U}$ automatically depend on t, but we will omit this time dependence in the rest of the paper for the sake of brevity.

3.1. Piecewise linear reconstruction and point values. Adopting the second-order finite-volume approach, we use the cell averages (3.3) to approximate the computed solution using a second-order piecewise linear reconstruction

$$\widetilde{\boldsymbol{U}}(x) := \overline{\boldsymbol{U}}_j + (\boldsymbol{U}_x)_j(x - x_j), \quad x \in I_j, \tag{3.4}$$

where the slopes $\boldsymbol{U}_x$ are to be computed in a nonoscillatory manner using a nonlinear limiter. In the numerical results reported in section 4, we have used the generalized minmod limiter (see, e.g., [30, 31, 37, 38]):

$$(\boldsymbol{U}_x)_j = \mathrm{minmod}\left(\theta\frac{\overline{\boldsymbol{U}}_{j+1} - \overline{\boldsymbol{U}}_j}{\Delta x}, \frac{\overline{\boldsymbol{U}}_{j+1} - \overline{\boldsymbol{U}}_{j-1}}{2\Delta x}, \theta\frac{\overline{\boldsymbol{U}}_j - \overline{\boldsymbol{U}}_{j-1}}{\Delta x}\right),$$

with the minmod function

$$\mathrm{minmod}(z_1, z_2, \dots) := \begin{cases} \min(z_1, z_2, \dots) & \text{if } z_i > 0 \ \ \forall i, \\ \max(z_1, z_2, \dots) & \text{if } z_i < 0 \ \ \forall i, \\ 0 & \text{otherwise.} \end{cases}$$

The parameter $\theta \in [1, 2]$ controls the amount of numerical dissipation: the larger the θ, the smaller the numerical dissipation.

Using the piecewise linear reconstruction (3.4), we obtain the left- and right-sided point values of $\boldsymbol{U}$ by

$$\boldsymbol{U}^-_{j+\frac{1}{2}} = \overline{\boldsymbol{U}}_j + \frac{\Delta x}{2}(\boldsymbol{U}_x)_j, \quad \boldsymbol{U}^+_{j+\frac{1}{2}} = \overline{\boldsymbol{U}}_{j+1} - \frac{\Delta x}{2}(\boldsymbol{U}_x)_{j+1}.$$

These values can be then used to compute the corresponding values of $u = q/h$. In order to be able to simulate the flows near wet-dry fronts, we use the desingularization procedure to compute the point value of velocity (see, e.g., [25, 28]), namely

$$u^\pm_{j+\frac{1}{2}} = \frac{\sqrt{2}h^\pm_{j+\frac{1}{2}}q^\pm_{j+\frac{1}{2}}}{\sqrt{\big(h^\pm_{j+\frac{1}{2}}\big)^4 + \max\big\{\big(h^\pm_{j+\frac{1}{2}}\big)^4, \tau\big\}}},$$

where $\tau > 0$ is the prescribed positive parameter which is used to avoid division by very small numbers. In the numerical experiments reported in section 4, we have used $\tau = 10^{-10}$. For consistency, we then recompute the corresponding values of q by

$$q^\pm_{j+\frac{1}{2}} = h^\pm_{j+\frac{1}{2}} \cdot u^\pm_{j+\frac{1}{2}}.$$

Equipped with the reconstructed point values of h and u, we use (2.14) and the eigenvalues of the Jacobian of the original Saint-Venant system (1.1) to set

$$a=\max_j\Big\{\max\Big(|u^+_{j+\frac12}|+\sqrt{gh^+_{j+\frac12}},\,|u^-_{j+\frac12}|+\sqrt{gh^-_{j+\frac12}},\\ \sqrt{3(u^+_{j+\frac12})^2+gh^+_{j+\frac12}},\,\sqrt{3(u^-_{j+\frac12})^2+gh^-_{j+\frac12}}\Big)\Big\}. \tag{3.5}$$

Notice that a is to be recomputed at every time step.

3.2. Source term evaluation. The third component of the source term $\boldsymbol{S}_j$ is computed using the midpoint quadrature:

$$\overline{S}^{(3)}_j=-\frac{1}{\varepsilon}\left(\overline{v}_j-K_j\right),$$

where the values of the global variable K at the cell centers x_j are computed by

$$K_j=\frac{\sqrt{2}\,\overline{h}_j\overline{q}_j^{\,2}}{\sqrt{\overline{h}_j^{\,4}+\max\left\{\overline{h}_j^{\,4},\tau\right\}}}+\frac{g}{2}\overline{h}_j^{\,2}+R_j,\quad j=j_\ell,\ldots,j_r. \tag{3.6}$$

Here, R_j can be computed using (1.6):

$$R_j=R(x_j,t)=g\int^{x_j}\left[h(\xi,t)B_x(\xi)+\frac{n^2}{h^{7/3}(\xi,t)}|q(\xi,t)|q(\xi,t)\right]\mathrm{d}\xi,\quad j=j_\ell,\ldots,j_r. \tag{3.7}$$

We notice that formula (3.7) can be rewritten in the following recursive way:

$$R_j=R_{j-1}+g\int_{x_{j-1}}^{x_j}\left[h(\xi,t)B_x(\xi)+\frac{n^2}{h^{7/3}(\xi,t)}|q(\xi,t)|q(\xi,t)\right]\mathrm{d}\xi,\quad j=j_\ell+1,\ldots,j_r, \tag{3.8}$$

and then the integral in (3.8) is discretized using a second-order quadrature to obtain

$$R_j=R_{j-1}+\frac{g}{2}(\overline{h}_{j-1}+\overline{h}_j)(B_j-B_{j-1})+\frac{gn^2}{2}\Bigg[\left(\frac{2\,\overline{h}_{j-1}}{\overline{h}_{j-1}^{\,2}+\max\left\{\overline{h}_{j-1}^{\,2},\tau\right\}}\right)^{7/3}|\overline{q}_{j-1}|\overline{q}_{j-1}\\ +\left(\frac{2\,\overline{h}_j}{\overline{h}_j^{\,2}+\max\left\{\overline{h}_j^{\,2},\tau\right\}}\right)^{7/3}|\overline{q}_j|\overline{q}_j\Bigg]\Delta x,\quad j=j_\ell+1,\ldots,j_r\,. \tag{3.9}$$

It should be observed that the recursive formula (3.9) requires a starting value R_{j_ℓ}, which is obtained by setting $R_{j_\ell}:=(R_{j_\ell-\frac12}+R_{j_\ell+\frac12})/2$, where $R_{j_\ell-\frac12}:=0$ and $R_{j_\ell+\frac12}$ is computed using a slightly different quadrature:

$$R_{j_\ell+\frac12}=g\overline{h}_{j_\ell}(B_{j_\ell+\frac12}-B_{j_\ell-\frac12})+gn^2|\overline{q}_{j_\ell}|\overline{q}_{j_\ell}\left(\frac{2\overline{h}_{j_\ell}}{\overline{h}_{j_\ell}^{\,2}+\max\left\{\overline{h}_{j_\ell}^{\,2},\tau\right\}}\right)^{7/3}\Delta x.$$

3.3. Upwind numerical fluxes. A semidiscrete finite-volume method for (3.1), (3.2) can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\overline{\boldsymbol{U}}_j=-\frac{\boldsymbol{\mathcal{F}}_{j+\frac12}-\boldsymbol{\mathcal{F}}_{j-\frac12}}{\Delta x}+\overline{\boldsymbol{S}}_j, \tag{3.10}$$

where $\boldsymbol{\mathcal{F}}_{j+\frac{1}{2}}$ are the numerical fluxes at the cell interfaces.

We note that equation in (2.7) is, in fact, coupled with equations (2.8) and (2.9) through the source term only. We therefore first consider the subsystem

$$\begin{pmatrix} q \\ v \end{pmatrix}_t + \begin{pmatrix} 0 & 1 \\ a^2 & 0 \end{pmatrix} \begin{pmatrix} q \\ v \end{pmatrix}_x = \begin{pmatrix} 0 \\ -\dfrac{1}{\varepsilon}(v-K) \end{pmatrix}. \tag{3.11}$$

Since the left-hand side (LHS) of (3.11) is linear with constant coefficients, one can easily derive upwind numerical fluxes $\mathcal{F}^{(2)}_{j+\frac{1}{2}}$ and $\mathcal{F}^{(3)}_{j+\frac{1}{2}}$, which read as

$$\mathcal{F}^{(2)}_{j+\frac{1}{2}} = \frac{v^+_{j+\frac{1}{2}} + v^-_{j+\frac{1}{2}}}{2} - \frac{a}{2}\left(q^+_{j+\frac{1}{2}} - q^-_{j+\frac{1}{2}}\right), \tag{3.12}$$

$$\mathcal{F}^{(3)}_{j+\frac{1}{2}} = a^2\,\frac{q^+_{j+\frac{1}{2}} + q^-_{j+\frac{1}{2}}}{2} - \frac{a}{2}\left(v^+_{j+\frac{1}{2}} - v^-_{j+\frac{1}{2}}\right). \tag{3.13}$$

We then use the relation $F^{(3)} = a^2 F^{(1)}$ to obtain the remaining component of the numerical flux:

$$\mathcal{F}^{(1)}_{j+\frac{1}{2}} = \frac{q^+_{j+\frac{1}{2}} + q^-_{j+\frac{1}{2}}}{2} - \frac{1}{2a}\left(v^+_{j+\frac{1}{2}} - v^-_{j+\frac{1}{2}}\right). \tag{3.14}$$

3.4. Modified equation analysis. We now analyze the stability of the developed semidiscrete upwind scheme using a modified equation analysis. To this end, we take the zero-relaxation limit, in which $\bar{v}_j = K_j$, use the first-order piecewise constant reconstruction obtained by setting $\boldsymbol{U}_j \equiv 0$ in (3.4), and substitute the numerical fluxes $\boldsymbol{\mathcal{F}}_{j+\frac{1}{2}}$ and $\boldsymbol{\mathcal{F}}_{j-\frac{1}{2}}$ into (3.10) to obtain

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t}\bar{h}_j + \frac{\bar{q}_{j+1} - \bar{q}_{j-1}}{2\Delta x} = \frac{\Delta x}{2a}\cdot\frac{K_{j+1} - 2K_j + K_{j-1}}{(\Delta x)^2}, \\ &\frac{\mathrm{d}}{\mathrm{d}t}\bar{q}_j + \frac{K_{j+1} - K_{j-1}}{2\Delta x} = \frac{a\Delta x}{2}\cdot\frac{\bar{q}_{j+1} - 2\bar{q}_j + \bar{q}_{j-1}}{(\Delta x)^2}. \end{aligned}$$

Thus, the modified equations for the first-order semidiscrete upwind scheme are

$$h_t + q_x = \frac{\Delta x}{2a}\,K_{xx}, \tag{3.15}$$

$$q_t + K_x = \frac{a\Delta x}{2}\,q_{xx}. \tag{3.16}$$

For simplicity, we consider the case with $B \equiv \mathrm{Const}$ and $n = 0$, in which K is given by (2.11). We then differentiate (2.12) with respect to x to obtain

$$K_{xx} = 2uq_{xx} + (gh - u^2)h_{xx} + \frac{2}{h}(q_x)^2 + \left[\frac{2u}{h} + g\right](h_x)^2 - \frac{4u}{h}q_x h_x,$$

which we substitute into (3.15) and rewrite the modified equations in the following vector form:

$$\begin{aligned} \begin{pmatrix} h \\ q \end{pmatrix}_t + \begin{pmatrix} q \\ v \end{pmatrix}_x = {} & \frac{\Delta x}{2a}\begin{pmatrix} gh - u^2 & 2u \\ 0 & a^2 \end{pmatrix}\begin{pmatrix} h \\ q \end{pmatrix}_{xx} \\ & + \begin{pmatrix} \dfrac{2}{h}(q_x)^2 + \left[\dfrac{2u}{h} + g\right](h_x)^2 - \dfrac{4u}{h}q_x h_x \\ 0 \end{pmatrix}. \end{aligned} \tag{3.17}$$

As one can see, the system (3.17) is dissipative as long as

$$|u| < \sqrt{gh}, \tag{3.18}$$

which is true for subcritical flows only. This indicates that the upwind numerical flux approximations (3.12)–(3.14) may lead to a stable approximate solution only under the subcritical flow condition (3.18).

Remark 3.1. We have conducted a series of numerical experiments (not reported in this paper), which clearly indicate that the developed partial relaxation scheme based on the upwind fluxes (3.12)–(3.14) is indeed unstable in the supercritical flow regime; see, e.g., Example 1, Case (a).

3.5. Hybrid upwind/central-upwind numerical fluxes. According to the analysis in section 3.4, at those cell interfaces $x = x_{j+\frac12}$, where either

$$\left|u^+_{j+\frac12}\right| > \sqrt{gh^+_{j+\frac12}} \quad \text{or} \quad \left|u^-_{j+\frac12}\right| > \sqrt{gh^-_{j+\frac12}},$$

we replace the upwind numerical fluxes (3.12)–(3.14) with the central-upwind ones, originally developed in [27] for general hyperbolic systems of conservation laws:

$$\boldsymbol{\mathcal{F}}_{j+\frac12} = \frac{a^+_{j+\frac12}\boldsymbol{F}(\boldsymbol{U}^-_{j+\frac12}) - a^-_{j+\frac12}\boldsymbol{F}(\boldsymbol{U}^+_{j+\frac12})}{a^+_{j+\frac12} - a^-_{j+\frac12}} + \frac{a^+_{j+\frac12}a^-_{j+\frac12}}{a^+_{j+\frac12} - a^-_{j+\frac12}}\left(\boldsymbol{U}^+_{j+\frac12} - \boldsymbol{U}^-_{j+\frac12}\right). \tag{3.19}$$

Here, $a^+_{j+\frac12}$ and $a^-_{j+\frac12}$ are one-sided local speeds of propagation. Their careful estimation allows one to control the amount of the numerical diffusion present in the central-upwind flux. If one wants to apply the central-upwind flux in all of the flow regimes (both sub- and supercritical), one has to set $a^+_{j+\frac12} = -a^-_{j+\frac12} \equiv a$ for all j, which would lead to a very diffusive and non-well-balanced discretization of the system (2.7)–(2.9). However, as we use the central-upwind flux in the supercritical regime only, we can safely estimate the one-sided local speeds in (3.19) by using the largest and smallest eigenvalues of the original system (1.1), namely by setting

$$\begin{aligned} a^+_{j+\frac12} &= \max\left\{u^+_{j+\frac12} + \sqrt{gh^+_{j+\frac12}}, u^-_{j+\frac12} + \sqrt{gh^-_{j+\frac12}}, 0\right\},\\ a^-_{j+\frac12} &= \min\left\{u^+_{j+\frac12} - \sqrt{gh^+_{j+\frac12}}, u^-_{j+\frac12} - \sqrt{gh^-_{j+\frac12}}, 0\right\}. \end{aligned}$$

Remark 3.2. We note that the presented flux hybridization does not guarantee the positivity of the computed water depth. In order to enforce the positivity, we modify the first component of the numerical fluxes, $\mathcal{F}^{(1)}_{j+\frac12}$, using the "draining" time-step strategy proposed in [3] (also see [2, 9, 10]).

3.6. Fully discrete semi-implicit scheme. The semidiscretization (3.10) results in a system of time-dependent ODEs, which should be solved using an efficient, stable, and sufficiently accurate ODE solver. Since the source term in (3.10) is stiff, explicit ODE solvers typically used for solving nonlinear hyperbolic systems may be extremely inefficient, as they suffer from severe stability time-step restrictions. We therefore implement the second-order steady state and sign preserving semi-implicit Runge–Kutta ODE solver SI-RK3 from [12, equation (3.2)]. In this section, we provide all of the required details, as application of the SI-RK3 method to the ODE

system (3.10) is not straightforward due to the presence of the nondamping source term K/ε.

Let us assume that at time level $t = t^m$ the computed cell averages, $\{\bar{h}_j^m\}$, $\{\bar{q}_j^m\}$, and $\{\bar{v}_j^m\}$, are available. We then evolve them to the next time level $t = t^{m+1} = t^m + \Delta t^m$ according to the following algorithm.

Step 1. Compute a by (3.5), and use it to evaluate

$$\Delta t^m = C\,\frac{\Delta x}{a},$$

where C is a CFL number. In the numerical results reported in section 4, we have used $C = 0.4$.

Step 2. Use $\{\bar{h}_j^m\}$, $\{\bar{q}_j^m\}$, and $\{\bar{v}_j^m\}$ to compute $\{\boldsymbol{\mathcal{F}}_{j+\frac{1}{2}}^m\}$ according to sections 3.1, 3.3, and 3.5.

Step 3. Compute

$$\begin{aligned}
\bar{h}_j^I &= \bar{h}_j^m - \frac{\Delta t^m}{\Delta x}\left(\mathcal{F}_{j+\frac{1}{2}}^{m,(1)} - \mathcal{F}_{j-\frac{1}{2}}^{m,(1)}\right),\\
\bar{q}_j^I &= \bar{q}_j^m - \frac{\Delta t^m}{\Delta x}\left(\mathcal{F}_{j+\frac{1}{2}}^{m,(2)} - \mathcal{F}_{j-\frac{1}{2}}^{m,(2)}\right).
\end{aligned} \tag{3.20}$$

Step 4. Use $\{\bar{h}_j^I\}$ and $\{\bar{q}_j^I\}$ to compute $\{K_j^I\}$ by (3.6), (3.9).

Step 5. Compute

$$\bar{v}_j^I = \frac{\varepsilon}{\varepsilon + \Delta t^m}\left[\bar{v}_j^m - \frac{\Delta t^m}{\Delta x}\left(\mathcal{F}_{j+\frac{1}{2}}^{m,(3)} - \mathcal{F}_{j-\frac{1}{2}}^{m,(3)}\right) + \frac{\Delta t^m}{\varepsilon}K_j^I\right].$$

Step 6. Use $\{\bar{h}_j^I\}$, $\{\bar{q}_j^I\}$, and $\{\bar{v}_j^I\}$ to compute $\{\boldsymbol{\mathcal{F}}_{j+\frac{1}{2}}^I\}$ according to sections 3.1, 3.3, and 3.5.

Step 7. Compute

$$\begin{aligned}
\bar{h}_j^{II} &= \frac{3}{4}\bar{h}_j^m + \frac{1}{4}\left[\bar{h}_j^I - \frac{\Delta t^m}{\Delta x}\left(\mathcal{F}_{j+\frac{1}{2}}^{I,(1)} - \mathcal{F}_{j-\frac{1}{2}}^{I,(1)}\right)\right],\\
\bar{q}_j^{II} &= \frac{3}{4}\bar{q}_j^m + \frac{1}{4}\left[\bar{q}_j^I - \frac{\Delta t^m}{\Delta x}\left(\mathcal{F}_{j+\frac{1}{2}}^{I,(2)} - \mathcal{F}_{j-\frac{1}{2}}^{I,(2)}\right)\right].
\end{aligned} \tag{3.21}$$

Step 8. Use $\{\bar{h}_j^{II}\}$ and $\{\bar{q}_j^{II}\}$ to compute $\{K_j^{II}\}$ by (3.6), (3.9).

Step 9. Compute

$$\bar{v}_j^{II} = \frac{3}{4}\bar{v}_j^m + \frac{\varepsilon}{4(\varepsilon + \Delta t^m)}\left[\bar{v}_j^I - \frac{\Delta t^m}{\Delta x}\left(\mathcal{F}_{j+\frac{1}{2}}^{I,(3)} - \mathcal{F}_{j-\frac{1}{2}}^{I,(3)}\right) + \frac{\Delta t^m}{\varepsilon}K_j^{II}\right].$$

Step 10. Use $\{\bar{h}_j^{II}\}$, $\{\bar{q}_j^{II}\}$, and $\{\bar{v}_j^{II}\}$ to compute $\{\boldsymbol{\mathcal{F}}_{j+\frac{1}{2}}^{II}\}$ according to sections 3.1, 3.3, and 3.5.

Step 11. Compute

$$\begin{aligned}
\bar{h}_j^{m+1} &= \frac{1}{3}\bar{h}_j^m + \frac{2}{3}\left[\bar{h}_j^{II} - \frac{\Delta t^m}{\Delta x}\left(\mathcal{F}_{j+\frac{1}{2}}^{II,(1)} - \mathcal{F}_{j-\frac{1}{2}}^{II,(1)}\right)\right],\\
\bar{q}_j^{m+1} &= \frac{1}{3}\bar{q}_j^m + \frac{2}{3}\left[\bar{q}_j^{II} - \frac{\Delta t^m}{\Delta x}\left(\mathcal{F}_{j+\frac{1}{2}}^{II,(2)} - \mathcal{F}_{j-\frac{1}{2}}^{II,(2)}\right)\right].
\end{aligned} \tag{3.22}$$

Step 12. Use $\{\bar{h}_j^{m+1}\}$ and $\{\bar{q}_j^{m+1}\}$ to compute $\{K_j^{m+1}\}$ by (3.6), (3.9).
Step 13. Compute

$$\bar{v}_j^{III} = \frac{1}{3}\bar{v}_j^m + \frac{2\varepsilon}{3(\varepsilon+\Delta t^m)}\left[\bar{v}_j^{II} - \frac{\Delta t^m}{\Delta x}\left(\mathcal{F}_{j+\frac{1}{2}}^{II,(3)} - \mathcal{F}_{j-\frac{1}{2}}^{II,(3)}\right) + \frac{\Delta t^m}{\varepsilon}K_j^{m+1}\right].$$

Step 14. Use $\{\bar{h}_j^{m+1}\}$, $\{\bar{q}_j^{m+1}\}$ and $\{\bar{v}_j^{III}\}$ to compute $\{\mathcal{F}_{j+\frac{1}{2}}^{III,(3)}\}$ according to sections 3.1, 3.3, and 3.5.
Step 15. Compute

$$\bar{v}_j^{m+1} = \frac{\varepsilon^2}{\varepsilon^2+(\Delta t^m)^2}\left[\bar{v}_j^{III} - \frac{(\Delta t^m)^2}{\varepsilon\Delta x}\left(\mathcal{F}_{j+\frac{1}{2}}^{III,(3)} - \mathcal{F}_{j-\frac{1}{2}}^{III,(3)}\right) + \left(\frac{\Delta t^m}{\varepsilon}\right)^2 K_j^{m+1}\right].$$

Remark 3.3. We would like to stress once again that the numerical fluxes $\mathcal{F}_{j+\frac{1}{2}}^{m,(1)}$, $\mathcal{F}_{j+\frac{1}{2}}^{I,(1)}$, and $\mathcal{F}_{j+\frac{1}{2}}^{II,(1)}$ in (3.20), (3.21), and (3.22), respectively, are to be modified using the "draining" time-step strategy from [2, 3, 9, 10] in order to guarantee the positivity of the computed values of h.

Remark 3.4. At time $t = 0$, we set $\bar{v}_j^0 = K_j^0$, which are computed using (3.6), (3.9).

4. Numerical examples. In this section, we test the performance of the proposed well-balanced partial relaxation scheme on a number of numerical examples. In these tests, we consider the cases of both continuous and discontinuous bottom topographies. In all of the examples, we perform computations in the computational domain $[0, 25]$ split into N uniform cells with simple transmissive boundary conditions and use the following parameters: the gravitational acceleration $g = 9.812$ and the minmod parameter $\theta = 1.9$.

4.1. Frictionless Saint-Venant system ($n = 0$). We first consider a frictionless case, for which the Manning friction coefficient $n = 0$.

Example 1—convergence to steady states. In this numerical example, we study the convergence of the solutions computed using the proposed partial relaxation scheme towards steady flow over a hump. The bottom topography function is continuous and defined by

$$B(x) = \begin{cases} 0.2 - 0.05(x-10)^2 & \text{if } 8 \le x \le 12, \\ 0 & \text{otherwise.} \end{cases} \tag{4.1}$$

Depending on the initial and boundary conditions, the flow may be subcritical, supercritical, or transcritical with or without a steady shock. We consider the four sets of initial data and boundary conditions taken from [10]; also see [23, 39]. The data for Cases (a), (b), (c), and (d) are summarized in Table 4.1 in which − represents a zero-extrapolation boundary condition.

Remark 4.1. In Case (c), the downstream boundary condition ($h(25, t) \equiv 0.66$) is imposed only when the flow is subcritical.

In all of these four cases, the numerical solutions are obtained by the proposed partial relaxation schemes at $t = 300$ using $N = 200$ uniform cells. The results ($h+B$, q, K, and E) are shown in Figures 4.1–4.4. As one can see, the proposed numerical

TABLE 4.1
Example 1: *Initial data (ID) and boundary conditions (BC).*

Case	Description	ID	BC at $x = 0$	BC at $x = 25$
(a)	Supercritical flow	$h(x,0) = 2 - B(x)$	$h(0,t) \equiv 2$	–
		$q(x,0) \equiv 0$	$q(0,t) \equiv 24$	–
(b)	Subcritical flow	$h(x,0) = 2 - B(x)$	–	$h(25,t) \equiv 2$
		$q(x,0) \equiv 0$	$q(0,t) \equiv 4.42$	–
(c)	Transcritical flow	$h(x,0) = 0.66 - B(x)$	–	$h(25,t) \equiv 0.66$
	without a jump	$q(x,0) \equiv 0$	$q(0,t) \equiv 1.53$	–
(d)	Transcritical flow	$h(x,0) = 0.33 - B(x)$	–	$h(25,t) \equiv 0.33$
	with a jump	$q(x,0) \equiv 0$	$q(0,t) \equiv 0.18$	–

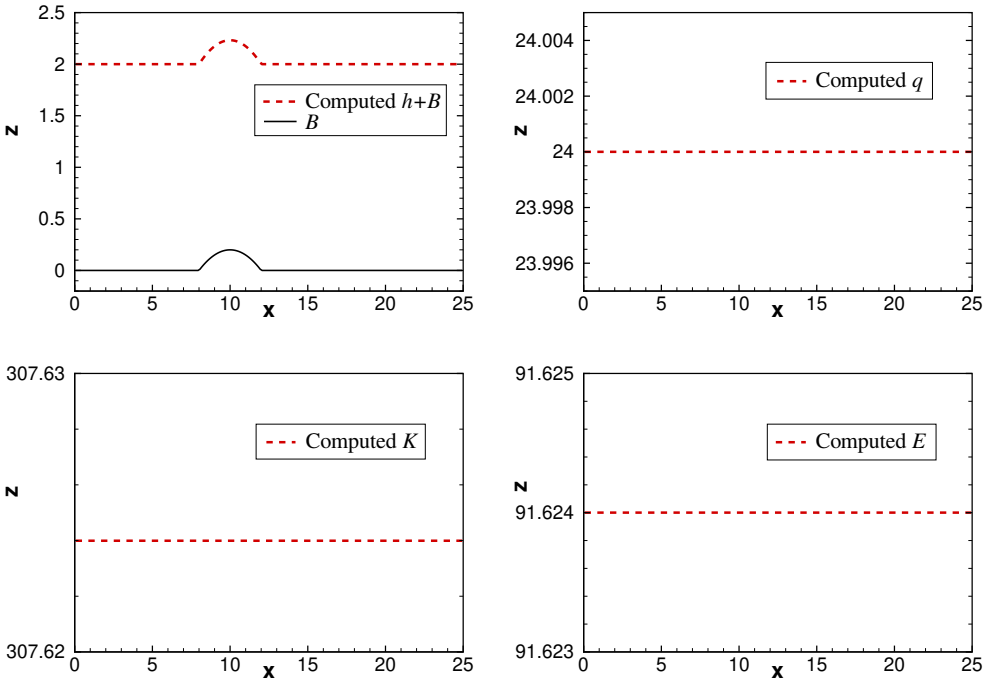

FIG. 4.1. *Example* 1, *Case* (a): *Computed* $h + B$, q, K *and* E.

scheme can capture the steady states of different flow regimes, and the computed solutions are comparable with those reported in [9, 10, 23, 39].

We then study the dependence of the computed solutions on the relaxation parameter ε in Cases (a) and (b), in which the exact smooth solutions are available. We measure the L^∞-errors for h, q, K, and E for $\varepsilon = 10^{-16}$, 10^{-12}, 10^{-8}, and 10^{-4} and report the obtained results in Tables 4.2 and 4.3. As one can clearly see, the designed partial relaxation scheme is not sensitive to the choice of ε.

In Table 4.4, we have compared the proposed partial relaxation scheme with several previously reported moving-water equilibrium preserving methods with respect to the L^∞- and L^1-errors in q for the subcritical flows, that is, Case (b) in Table 4.1 with $N = 200$ uniform cells. As one can see, our new scheme is more accurate than those previously reported second-order moving-water equilibrium preserving methods

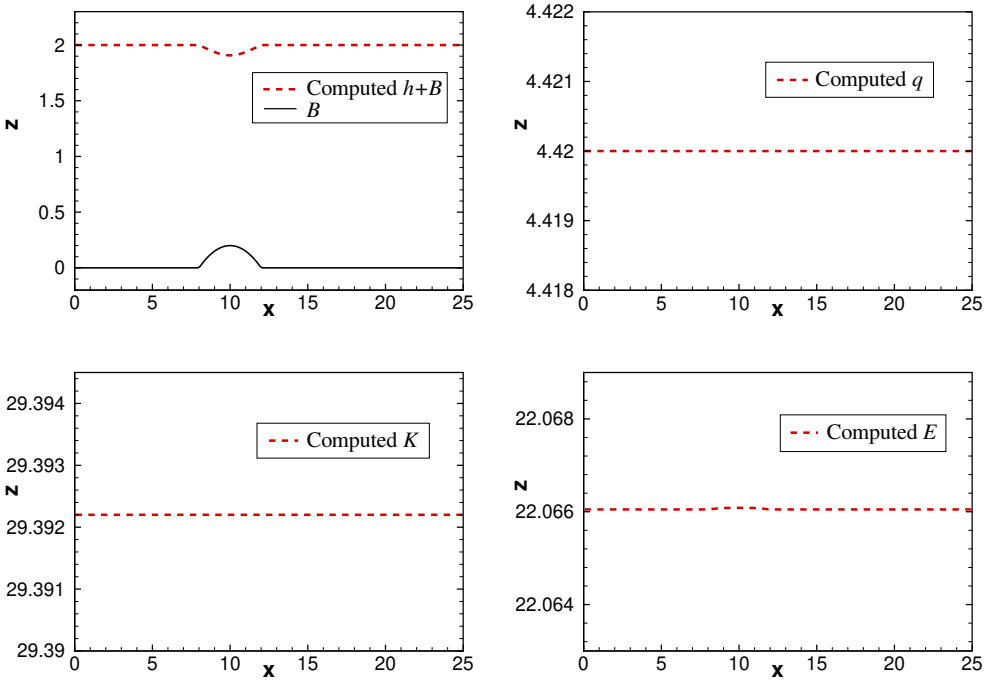

FIG. 4.2. *Example* 1, *Case* (b): *Computed* $h + B$, q, K, *and* E.

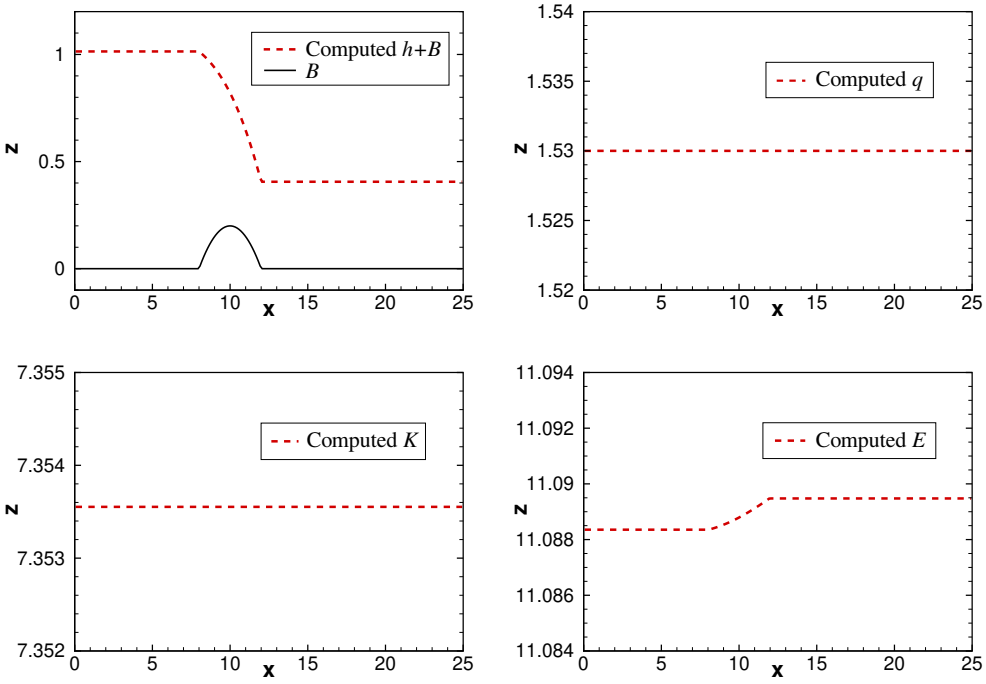

FIG. 4.3. *Example* 1, *Case* (c): *Computed* $h + B$, q, K, *and* E.

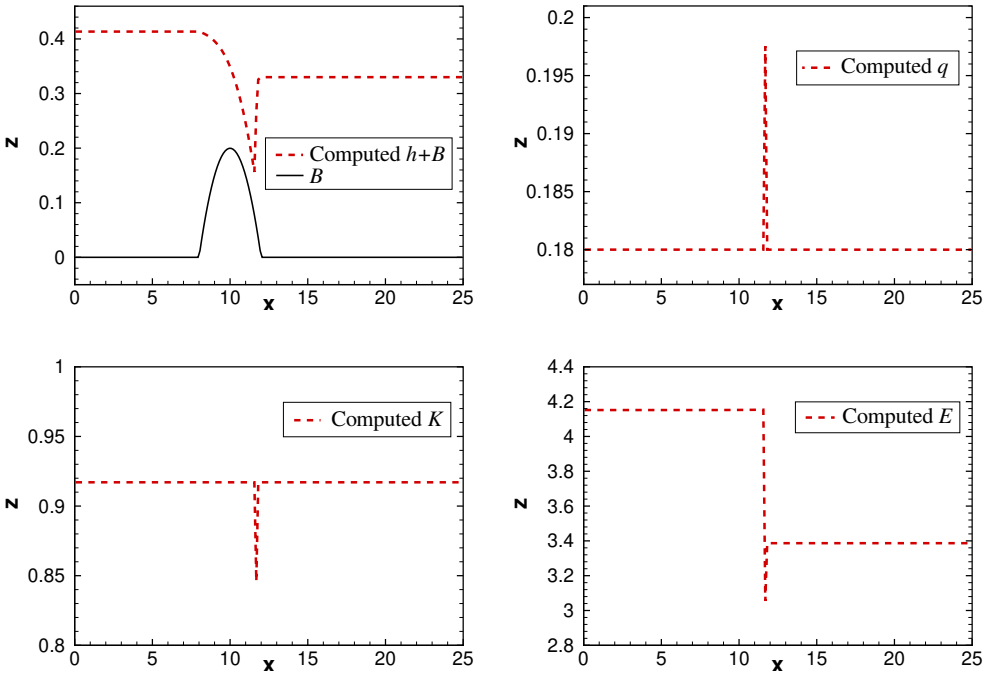

FIG. 4.4. *Example* 1, *Case* (d): *Computed* $h + B$, q, K, *and* E.

TABLE 4.2
Example 1: L^∞*-errors in Case* (a) *for different values of* ε.

ε	$\|h - h_{exact}\|_{L^\infty}$	$\|q - q_{exact}\|_{L^\infty}$	$\|K - K_{exact}\|_{L^\infty}$	$\|E - E_{exact}\|_{L^\infty}$
10^{-16}	3.26e-05	2.30e-12	4.00e-11	1.03e-06
10^{-12}	3.26e-05	2.20e-12	3.91e-11	1.03e-06
10^{-8}	3.26e-05	2.20e-12	4.00e-11	1.03e-06
10^{-4}	3.26e-05	2.10e-12	3.60e-11	1.03e-06

TABLE 4.3
Example 1: L^∞*-errors in Case* (b) *for different values of* ε.

ε	$\|h - h_{exact}\|_{L^\infty}$	$\|q - q_{exact}\|_{L^\infty}$	$\|K - K_{exact}\|_{L^\infty}$	$\|E - E_{exact}\|_{L^\infty}$
10^{-16}	3.31e-04	6.40e-13	5.40e-12	3.35e-05
10^{-12}	3.31e-04	6.60e-13	5.30e-12	3.35e-05
10^{-8}	3.31e-04	6.20e-13	5.20e-12	3.35e-05
10^{-4}	3.31e-04	4.50e-13	4.40e-12	3.35e-05

at the moving-water equilibrium. Several third- and fourth-order existing schemes yield, as expected, better accuracy, however, at the cost of more expensive spatial reconstruction.

Example 2—small perturbations of moving-water equilibria. This example is designed to test the capability of the proposed numerical scheme of capturing small perturbations of the moving-water equilibria studied in Example 1 over the continuous bottom topography function (4.1). The initial conditions in different flow

TABLE 4.4
Example 1, *Case* (b): *Comparisons of the* L^∞- *and* L^1-*errors in* q *for several moving-water equilibrium preserving schemes.*

Schemes	Formal order	L^∞-error in q	L^1-error in q
Partial relaxation scheme	2	6.6e-13	2.2e-13
George [19]	2	3.0e-6	-
Cheng and Kurganov [10]	2	-	5.3e-5
Xing [40]	3	4.2e-13	2.6e-14
Castro Díaz et al. [6]	3	-	1.0e-13
Noelle et al. [32]	4	1.77e-15	8.84e-17

regimes of this test are taken from the steady state solutions of Cases (a)–(c) in Example 1 with adding a small positive number 0.05 to the water depth in the interval $x \in [5.75, 6.25]$. Theoretically, this disturbance should split into two waves propagating with the flows. We run the tests with 200 uniform cells. We note that the same perturbations of the steady-state solutions were considered in [42].

We compute the numerical solutions using the proposed numerical scheme until the final times $t = 1$ in Case (a) and $t = 1.5$ in Cases (b) and (c) over a coarse and fine meshes with $N = 100$ and 1000 uniform cells, respectively. The obtained results are plotted in Figures 4.5–4.7. As one can observe, no spurious oscillations are generated and the propagating perturbations are well captured over both the coarse and the fine meshes.

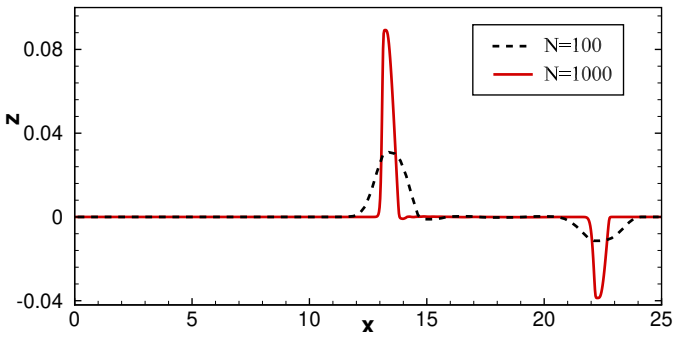

FIG. 4.5. *Example* 2, *Case* (a): *The difference between* h *and the background moving steady-state water depth computed using* 100 *and* 1000 *uniform cells.*

4.2. Saint-Venant system with friction ($n \neq 0$). We now test the proposed partial relaxation scheme in the presence of the Manning friction term, and we set $n = 0.05$. Examples 4–8 below are taken from [9].

Example 3—accuracy test. In this example, we modify the test taken from [28, 32] by adding the bottom friction. We verify the convergence rate of the proposed partial relaxation schemes in the case of a smooth solution. The following initial data and bottom topography,

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

are defined over the computational domain $[0, 1]$, and the periodic boundary conditions are imposed. Since the exact solution is not explicitly known, we use a numerical

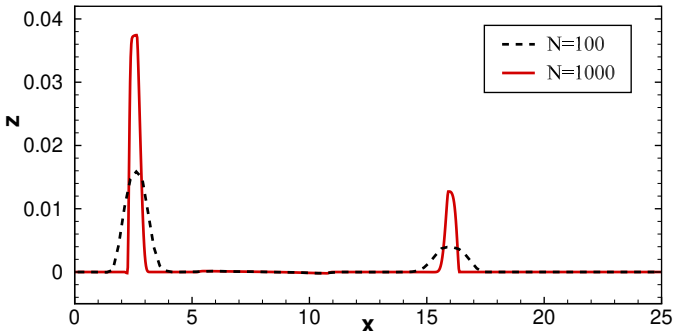

FIG. 4.6. *Example* 2, *Case* (b): *The difference between* h *and the background moving steady-state water depth computed using* 100 *and* 1000 *uniform cells.*

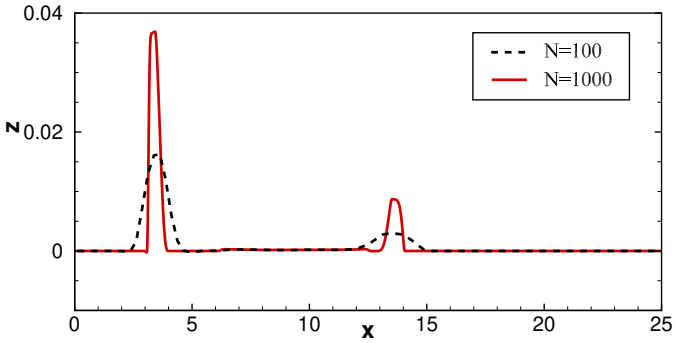

FIG. 4.7. *Example* 2, *Case* (c): *The difference between* h *and the background moving steady-state water depth computed using* 100 *and* 1000 *uniform cells.*

solution computed on a very fine mesh with $N = 8000$ uniform grid cells as a reference solution. We run the simulations until the final time $t = 0.1$ when the solution is still smooth (shocks are developed later in time). In this test, we set $a = 10$.

We measure the L^1-errors for both h and q over different grid sizes. The results are reported on Table 4.5, where one can observe the second order of accuracy as expected.

TABLE 4.5
Example 3: L^1*-errors and experimental convergence rates.*

Number of	h		q	
grid cells	L^1-error	rate	L^1-error	rate
200	1.05e-03	–	9.27e-03	–
400	2.66e-04	1.99	2.19e-03	2.08
800	6.47e-05	2.04	5.27e-04	2.06
1600	1.59e-05	2.02	1.26e-04	2.06
3200	4.27e-06	1.90	2.75e-05	2.20

Example 4—convergence to steady states (continuous bottom topography). This is a modification of Example 1 with the only difference that we now take into account the Manning friction term. We still consider the same super-, sub-,

and transcritical cases with the initial and boundary conditions as in Cases (a), (b), and (d) in Table 4.1, respectively.

In all of these three cases, the numerical solutions are obtained by the proposed partial relaxation scheme at time $t = 300$ using 200 uniform cells. The obtained results ($h + B$, q, and K) are shown in Figures 4.8–4.10. As one can see, the proposed numerical scheme can capture the steady states of different flow regimes in the presence of bed friction.

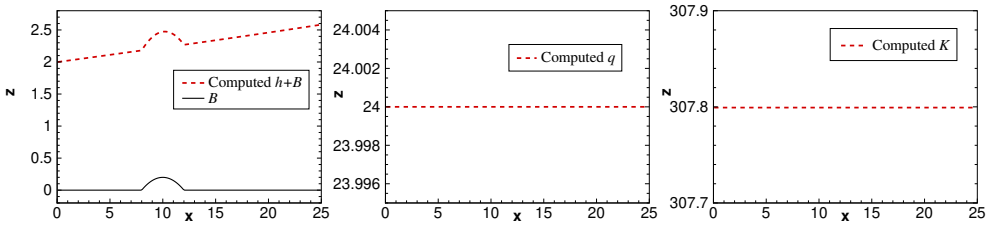

FIG. 4.8. *Example* 4, *Case* (a): *Computed* $h + B$, q, *and* K.

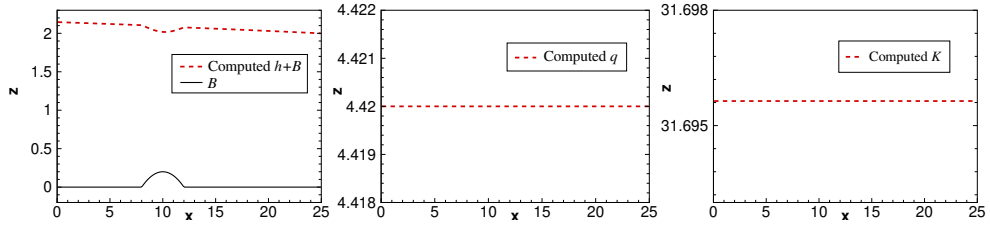

FIG. 4.9. *Example* 4, *Case* (b): *Computed* $h + B$, q, *and* K.

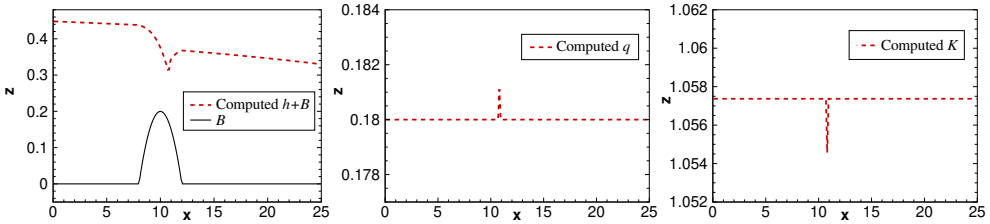

FIG. 4.10. *Example* 4, *Case* (d): *Computed* $h + B$, q, *and* K.

Example 5—small perturbations of moving-water equilibria (continuous bottom topography). In this example, we test the ability of the proposed partial relaxation scheme to capture the propagation of small perturbations of the moving-water equilibria. We consider the same bottom topography definition (4.1), and the initial data are obtained by adding a small positive number 0.001 to the water depth in the interval $x \in [4.5, 5.5]$ to the steady-state solutions obtained in Cases (a) and (b) in Example 4.

We compute the numerical solutions until the final times $t = 1$ in Case (a) and $t = 1.5$ in Case (b) using coarse and fine meshes with 100 and 1000 uniform cells, respectively. The obtained results are shown in Figures 4.11 and 4.12, where one can observe that no spurious oscillations are generated.

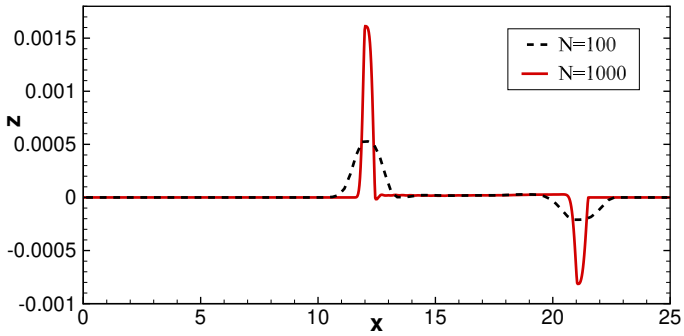

FIG. 4.11. *Example* 5, *Case* (a): *The difference between* h *and the background moving steady-state water depth computed using* 100 *and* 1000 *uniform cells.*

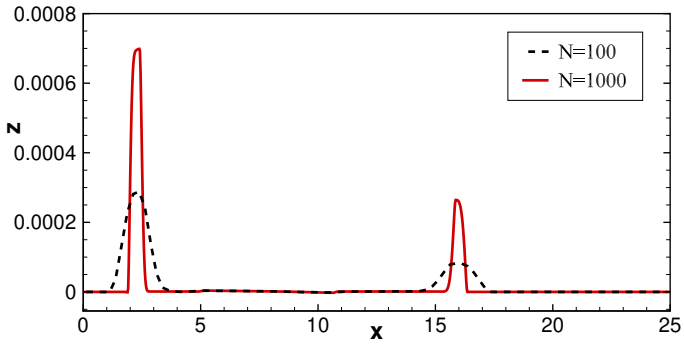

FIG. 4.12. *Example* 5, *Case* (b): *The difference between* h *and the background moving steady-state water depth computed using* 100 *and* 1000 *uniform cells.*

Example 6—convergence to steady states (discontinuous bottom topography). Next, we numerically study the convergence towards a steady flow over a discontinuous bump. The bottom topography is given by

$$B(x) = \begin{cases} 0.2 & \text{if } \ 8 \le x \le 12, \\ 0 & \text{otherwise.} \end{cases} \tag{4.2}$$

We consider the same super-, sub-, and transcritical cases with the initial and boundary conditions as in Cases (a), (b), and (d) in Table 4.1, respectively. In all of these three cases, the numerical solutions are obtained by the proposed partial relaxation scheme at time $t = 300$ using 200 uniform cells.

The obtained results ($h + B$, q, and K) are shown in Figures 4.13–4.15. As one case see, in Cases (a) and (b), the quality of the obtained results is practically not affected by the presence of the discontinuity in B, while in the transcritical case (Case (d)), somewhat larger (compared to Example 4) jumps in q and K at the flow regime transition location near $x = 12$ can be observed. Yet, one can conclude that the proposed numerical scheme is capable of capturing the steady states of different flow regimes over the discontinuous bottom bed.

Example 7—small perturbations of moving-water equilibria (discontinuous bottom topography). In this example, we investigate the ability of the

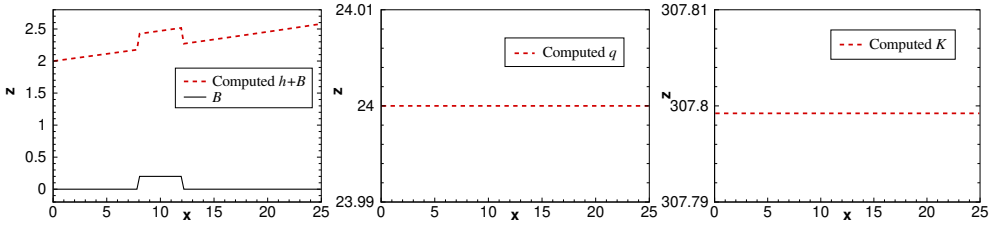

FIG. 4.13. *Example* 6, *Case* (a): *Computed* $h + B$, q, *and* K.

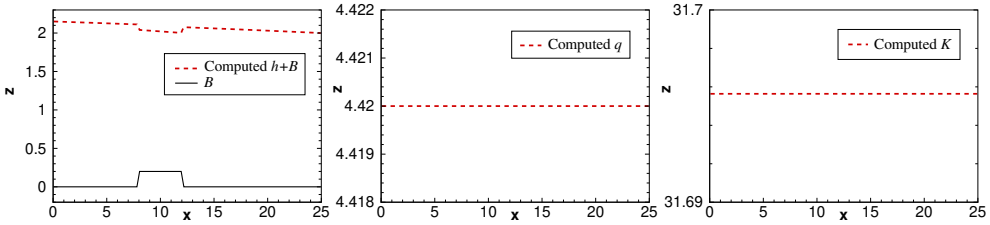

FIG. 4.14. *Example* 6, *Case* (b): *Computed* $h + B$, q, *and* K.

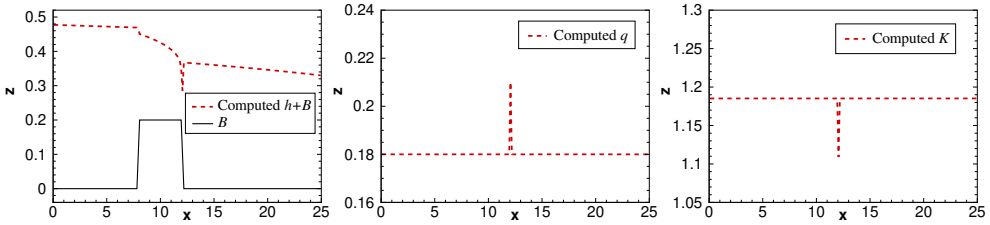

FIG. 4.15. *Example* 6, *Case* (d): *Computed* $h + B$, q, *and* K.

proposed partial relaxation scheme to capture small perturbations of the moving-water equilibria over the discontinuous bottom topography given by (4.2). The initial data are obtained by adding a small positive number 0.001 to the water depth in the interval $x \in [5.75, 6.25]$ to the steady-state solutions obtained in Cases (a) and (b) in Example 6. In this test, we compute the solutions until the final times $t = 1$ in Case (a) and $t = 1.5$ in Case (b) using either 100 or 1000 uniform grid cells. The obtained results, reported in Figures 4.16 and 4.17, demonstrate that in the case of discontinuous B, the proposed numerical scheme is able to capture small perturbations of the moving-water steady state in a nonoscillatory manner.

Example 8—Riemann problem. In this example, we test the performance of the proposed partial relaxation scheme on the test problem with a dry bed by numerically solving the initial value problem with the Riemann initial data

$$h(x,0) = \begin{cases} 2 & \text{if } x < 5, \\ 0 & \text{otherwise}, \end{cases} \qquad q(x,0) = \begin{cases} 24 & \text{if } x < 5, \\ 0 & \text{otherwise} \end{cases}$$

and the same bottom topography and boundary conditions as in Example 1, Case (a); see Table 4.1. We compute the solution using 100 uniform grid cells and plot its snapshots at times $t = 0.1$, 0.5, 1, 2, and 5 in Figure 4.18. One can observe that the obtained results are stable, the water flow runs through the bump, and by $t = 5$ it

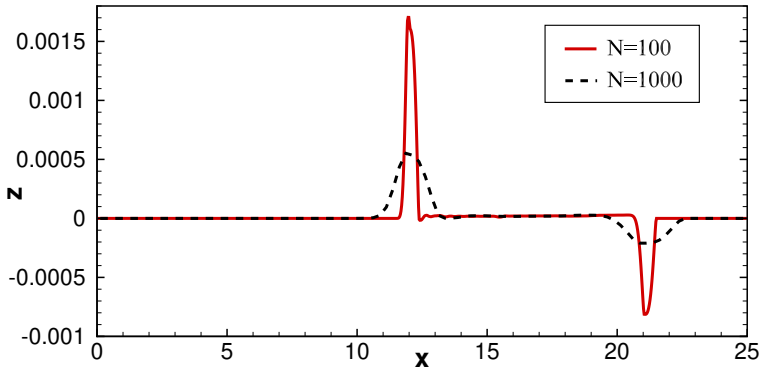

FIG. 4.16. *Example* 7, *Case* (a): *the difference between* h *and the background moving steady-state water depth computed using* 100 *and* 1000 *uniform cells.*

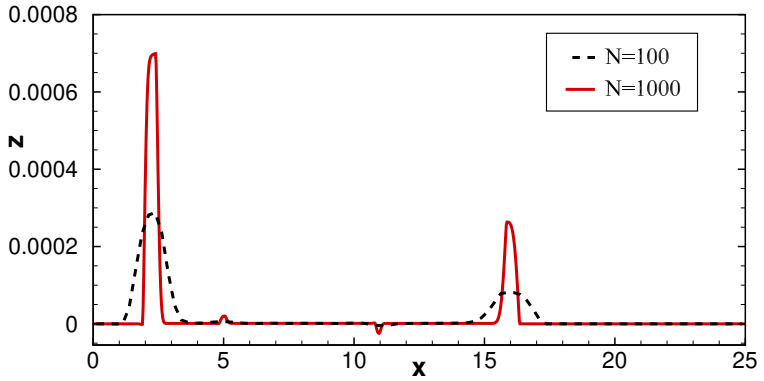

FIG. 4.17. *Example* 7, *Case* (b): *the difference between* h *and the background moving steady-state water depth computed using* 100 *and* 1000 *uniform cells.*

reaches the same steady state as in Example 5, Case (a) as expected.

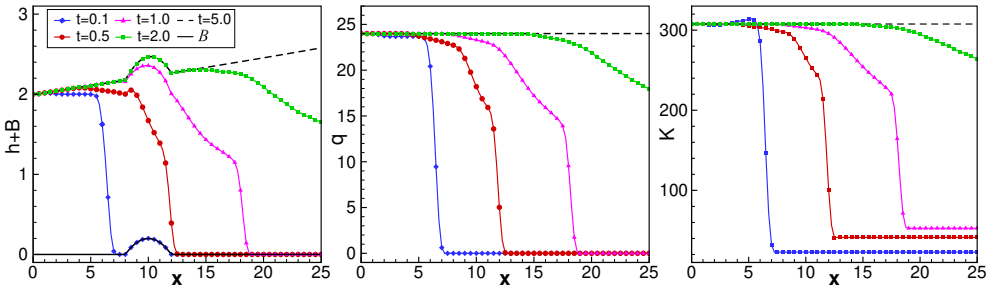

FIG. 4.18. *Example* 8: *Computed* $h + B$ *(left),* q *(middle), and* K *(right).*

4.3. Comparison with an alternative well-balanced method. We now compare the performance of the proposed partial relaxation scheme (referred to below as "new scheme") with an alternative moving-water equilibrium preserving method—the central-upwind scheme recently introduced in [9] (hereafter referred to as "old

scheme"). The scheme from [9] is based on a global flux formulation (1.4)–(1.6), piecewise linear reconstruction of the equilibrium variables q and K, and recovering of the point values $h^{\pm}_{j+\frac{1}{2}}$ by exactly solving the cubic equations

$$\frac{(q^{\pm}_{j+\frac{1}{2}})^2}{h}+\frac{g}{2}h^2+R_{j+\frac{1}{2}}-K^{\pm}_{j+\frac{1}{2}}=0 \tag{4.3}$$

for h; see [9] for details.

Example 9—steady state of smooth transcritical flow. In the final example, we consider a frictionless ($n=0$) case and take the following bottom topography over a $[0,100]$ computational domain,

$$B(x)=\begin{cases}0 & \text{if } \ x<8 \ \text{ or } \ x>92,\\ 0.2-0.05(x-10)^2 & \text{if } \ 8\le x<10,\\ 0.15+0.05e^{-36\sin^{36}(\pi(x-2)/8)} & \text{if } \ 27<x<35 \ \text{ or } \ 65<x<73,\\ 0.2-0.05(x-90)^2 & \text{if } \ 90<x\le 92,\\ 0.2 & \text{otherwise,}\end{cases}$$

and the same initial and boundary conditions as in the transcritical case (Case (c)) studied in Example 1, namely

$$\begin{aligned}&h(x,0)=0.66-B(x),\quad q(x,0)\equiv 0,\\ &q(0,t)=1.53,\quad h(25,t)=0.66.\end{aligned}$$

The solution of this initial boundary value problem is expected to converge to a smooth transition steady flow with a large transcritical zone [10, 90].

We first investigate how the proposed new scheme performs over a coarse mesh (with $N=100$ uniform cells) against the old scheme. To this end, we use the convergent numerical results over a very fine mesh (with $N=1000$ uniform cells) as reference solutions, and run the computations until a very large final time $t=5000$, by which both the new and the old schemes converge to their corresponding steady states. In Figure 4.19, where we plot the obtained numerical solutions, one can observe that even though both schemes are moving-water equilibria preserving, the new scheme yields more accurate steady-state results over the coarse grid than the old scheme. The reason why the old scheme performs relatively poorly in this test problem is that when solving the cubic equation (4.3) in the transcritical zone, the nonlinear solver faces an unavoidable instability at some cell interfaces there. In these points, the old scheme switches to a non-well-balanced reconstruction of the water surface $h+B$ (see [9] for details), which makes the old scheme not truly well-balanced. We note that this switch did not cause any problems in the benchmarks considered in [9] since the transcritical zone there contained one point only, while here the transcritical zone is large.

We then take the obtained numerical steady states and add a small perturbation $-0.2e^{-(x-50)^2}$ to its water depth field. We run the code for an additional 20 time units and compare the solutions of both the new and the old schemes over the coarse grid with the corresponding reference solution. The obtained results, plotted in Figure 4.20, clearly demonstrate that the proposed new scheme yields more accurate results than the old scheme.

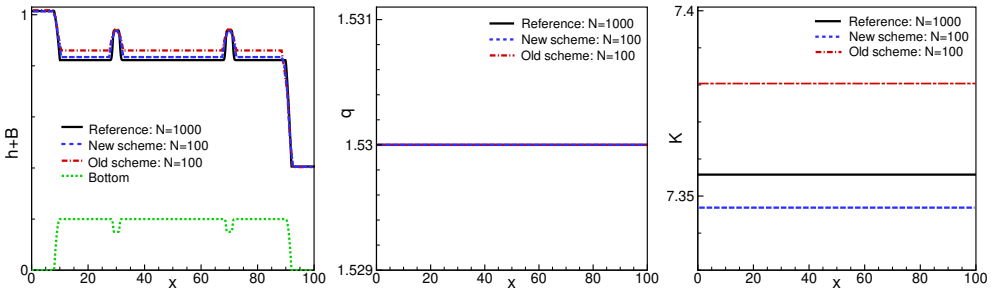

FIG. 4.19. *Example* 9: *Computed* $h + B$ *(left),* q *(middle), and* K *(right) at numerical steady states.*

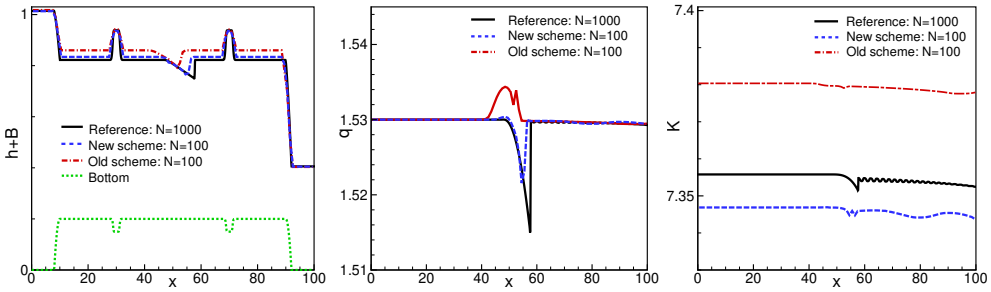

FIG. 4.20. *Example* 9: *Computed* $h + B$ *(left),* q *(middle), and* K *(right) at* $t = 20$ *after adding a perturbation to the steady state shown in Figure* 4.19.

5. Conclusion. In this paper, we have developed a novel partial relaxation scheme for the 1-D Saint-Venant system of shallow water equations. The scheme is well-balanced by construction as the source term is first incorporated into the global flux and then shifted into the nonlinear part of the relaxation approximation. This approach is advantageous compared to the existing alternatives in two aspects. First, no special treatment of the geometric source term is required. This is important since the well-balanced discretization proposed in [32] is, in fact, inconsistent for discontinuous solutions. Second, no nonlinear (cubic) equations should be solved to obtain the point values of the water depth out of the reconstructed equilibrium variables. This is important since solving such cubic equations may be impossible (typically in critical and transcritical cases), which forces one to use non-well-balanced reconstruction, giving up the well-balanced property (at least locally); see, e.g., [9, 10].

The partial relaxation scheme was successfully tested on a large number of benchmarks in order to demonstrate the high accuracy and robustness of the proposed method. In our future work, we plan to extend the developed scheme to several more complicated shallow water models, including the 2-D ones, and to other hyperbolic systems of balance laws, such as compressible Euler equations with gravitation. The 2-D extension will be carried out through the flux globalization approach used in [11].

Acknowledgment. We would like to thank Dr. Theodoros Katsaounis for fruitful discussions and helpful comments.

REFERENCES

[1] E. AUDUSSE, F. BOUCHUT, M.-O. BRISTEAU, R. KLEIN, AND B. PERTHAME, *A fast and stable*

well-balanced scheme with hydrostatic reconstruction for shallow water flows, SIAM J. Sci. Comput., 25 (2004), pp. 2050–2065, https://doi.org/10.1137/S1064827503431090.

[2] A. Bollermann, G. Chen, A. Kurganov, and S. Noelle, *A well-balanced reconstruction of wet/dry fronts for the shallow water equations*, J. Sci. Comput., 56 (2013), pp. 267–290.

[3] A. Bollermann, S. Noelle, and M. Lukáčová-Medviďová, *Finite volume evolution Galerkin methods for the shallow water equations with dry beds*, Commun. Comput. Phys., 10 (2011), pp. 371–404.

[4] F. Bouchut and T. Morales de Luna, *A subsonic-well-balanced reconstruction scheme for shallow water flows*, SIAM J. Numer. Anal., 48 (2010), pp. 1733–1758, https://doi.org/10.1137/090758416.

[5] 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 (2007), pp. 2055–2113.

[6] 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 (2013), pp. 242–264.

[7] L. Cea and M. E. Vázquez-Cendón, *Unstructured finite volume discretisation of bed friction and convective flux in solute transport models linked to the shallow water equations*, J. Comput. Phys., 231 (2012), pp. 3317–3339.

[8] S. Chapman and T. G. Cowling, *The Mathematical Theory of Non-uniform Gases. An Account of the Kinetic Theory of Viscosity, Thermal Conduction and Diffusion in Gases*, 3rd ed., Cambridge University Press, London, 1970.

[9] Y. Cheng, A. Chertock, M. Herty, A. Kurganov, and T. Wu, *A new approach for designing moving-water equilibria preserving schemes for the shallow water equations*, J. Sci. Comput., 80 (2019), pp. 538–554.

[10] Y. Cheng and A. Kurganov, *Moving-water equilibria preserving central-upwind schemes for the shallow water equations*, Commun. Math. Sci., 14 (2016), pp. 1643–1663.

[11] A. Chertock, S. Cui, A. Kurganov, Ş. N. Özcan, and E. Tadmor, *Well-balanced schemes for the Euler equations with gravitation: Conservative formulation using global fluxes*, J. Comput. Phys., 358 (2018), pp. 36–52.

[12] A. Chertock, S. Cui, A. Kurganov, and W. Tong, *Steady state and sign preserving semi-implicit Runge–Kutta methods for ODEs with stiff damping term*, SIAM J. Numer. Anal., 53 (2015), pp. 2008–2029, https://doi.org/10.1137/151005798.

[13] 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 (2015), pp. 355–383.

[14] A. Chertock, M. Herty, and Ş. N. Özcan, *Well-balanced central-upwind schemes for* 2×2 *systems of balance laws*, in Theory, Numerics and Applications of Hyperbolic Problems I, Springer Proc. Math. Stat. 236, Springer, Cham, 2018, pp. 345–361.

[15] F. Coquel and B. Perthame, *Relaxation of energy and approximate Riemann solvers for general pressure laws in fluid dynamics*, SIAM J. Numer. Anal., 35 (1998), pp. 2223–2249, https://doi.org/10.1137/S0036142997318528.

[16] A. J. C. de Saint-Venant, *Thèorie du mouvement non-permanent des eaux, avec application aux crues des rivière at à l'introduction des marèes dans leur lit*, C. R. Acad. Sci. Paris, 73 (1871), pp. 147–154, 237–240.

[17] A. I. Delis and T. Katsaounis, *Relaxation schemes for the shallow water equations*, Internat. J. Numer. Methods Fluids, 41 (2003), pp. 695–719.

[18] J. M. Gallardo, C. Parés, and M. Castro, *On a well-balanced high-order finite volume scheme for shallow water equations with topography and dry areas*, J. Comput. Phys., 227 (2007), pp. 574–601.

[19] D. L. George, *Augmented Riemann solvers for the shallow water equations over variable topography with steady states and inundation*, J. Comput. Phys., 227 (2008), pp. 3089–3113.

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

[21] S. Jin and X. Wen, *Two interface-type numerical methods for computing hyperbolic systems with geometrical source terms having concentrations*, SIAM J. Sci. Comput., 26 (2005), pp. 2079–2101, https://doi.org/10.1137/040605825.

[22] S. Jin and Z. P. Xin, *The relaxation schemes for systems of conservation laws in arbitrary space dimensions*, Comm. Pure Appl. Math., 48 (1995), pp. 235–276.

[23] A. A. Khan and W. Lai, *Modeling Shallow Water Flows Using the Discontinuous Galerkin Method*, CRC Press, Taylor & Francis, New York, 2014.

[24] C. Klingenberg, A. Kurganov, Y. Liu, and M. Zenk, *Moving-Water Equilibria Preserving*

HLL-Type Schemes for the Shallow Water Equations, preprint.

[25] A. Kurganov, *Finite-volume schemes for shallow-water equations*, Acta Numer., 27 (2018), pp. 289–351.

[26] A. Kurganov and D. Levy, *Central-upwind schemes for the Saint-Venant system*, M2AN Math. Model. Numer. Anal., 36 (2002), pp. 397–425.

[27] A. Kurganov, S. Noelle, and G. Petrova, *Semidiscrete central-upwind schemes for hyperbolic conservation laws and Hamilton–Jacobi equations*, SIAM J. Sci. Comput., 23 (2001), pp. 707–740, https://doi.org/10.1137/S1064827500373413.

[28] A. Kurganov and G. Petrova, *A second-order well-balanced positivity preserving central-upwind scheme for the Saint-Venant system*, Commun. Math. Sci., 5 (2007), pp. 133–160.

[29] R. J. LeVeque, *Balancing source terms and flux gradients in high-resolution Godunov methods: The quasi-steady wave-propagation algorithm*, J. Comput. Phys., 146 (1998), pp. 346–365.

[30] 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 (2003), pp. 1157–1174, https://doi.org/10.1137/S1064827501392880.

[31] H. Nessyahu and E. Tadmor, *Nonoscillatory central differencing for hyperbolic conservation laws*, J. Comput. Phys., 87 (1990), pp. 408–463.

[32] 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 (2007), pp. 29–58.

[33] M. Ricchiuto and A. Bollermann, *Stabilized residual distribution for shallow water simulations*, J. Comput. Phys., 228 (2009), pp. 1071–1115.

[34] G. Russo, *Central schemes for conservation laws with application to shallow water equations*, in Trends and Applications of Mathematics to Mechanics, Springer, Milan, 2005, pp. 225–246.

[35] G. Russo and A. Khe, *High order well-balanced schemes based on numerical reconstruction of the equilibrium variables*, in Proceedings of the 15th Conference on Waves and Stability in Continuous Media (WASCOM 2009), World Scientific, Hackensack, NJ, 2010, pp. 230–241.

[36] I. Suliciu, *On the thermodynamics of rate-type fluids and phase transitions.* I. *Rate-type fluids*, Internat. J. Engrg. Sci., 36 (1998), pp. 921–947.

[37] P. K. Sweby, *High resolution schemes using flux limiters for hyperbolic conservation laws*, SIAM J. Numer. Anal., 21 (1984), pp. 995–1011, https://doi.org/10.1137/0721062.

[38] B. van Leer, *Towards the ultimate conservative difference scheme.* V. *A second-order sequel to Godunov's method*, J. Comput. Phys., 32 (1979), pp. 101–136.

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

[40] Y. Xing, *Exactly well-balanced discontinuous Galerkin methods for the shallow water equations with moving water equilibrium*, J. Comput. Phys., 257 (2014), pp. 536–553.

[41] Y. Xing and C.-W. Shu, *A survey of high order schemes for the shallow water equations*, J. Math. Study, 47 (2014), pp. 221–249.

[42] 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 (2011), pp. 339–349.