Journal of Computational Physics 257 (2014) 536–553

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

Exactly well-balanced discontinuous Galerkin methods for the shallow water equations with moving water equilibrium ☆

Yulong Xing [a,b,*]

[a] Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, United States
[b] Department of Mathematics, University of Tennessee, Knoxville, TN 37996, United States

ARTICLE INFO

Article history:
Received 15 November 2012
Received in revised form 3 October 2013
Accepted 6 October 2013
Available online 14 October 2013

Keywords:
Shallow water equations
Discontinuous Galerkin method
Moving water equilibrium
High order accuracy
Well-balanced
Positivity-preserving methods

ABSTRACT

Hyperbolic conservation laws with source terms often admit steady state solutions where the fluxes and source terms balance each other. To capture this balance and near-equilibrium solutions, well-balanced methods have been introduced and performed well in many numerical tests. Shallow water equations have been extensively investigated as a prototype example. In this paper, we develop well-balanced discontinuous Galerkin methods for the shallow water system, which preserve not only the still water at rest steady state, but also the more general moving water equilibrium. The key idea is the recovery of well-balanced states, a special source term approximation, and the approximation of the numerical fluxes based on a generalized hydrostatic reconstruction. We also study the extension of the positivity-preserving limiter presented in [40] in this framework. Numerical examples are provided at the end to verify the well-balanced property and good resolution for smooth and discontinuous solutions.

1. Introduction

Hyperbolic systems often involve source terms arising from geometrical, reactive, biological or other considerations. They are often referred as balance laws and have wide applications in different fields including chemistry, biology, fluid dynamics, astrophysics, and meteorology. In one dimension, they usually take the form

$$U_t + f(U,x)_x = s(U,x), \tag{1.1}$$

where U is the solution vector, $f(U,x)$ is the flux and $s(U,x)$ is the source term. Such balance laws admit steady state solutions which are usually non-trivial and often carry important physical meaning. A straightforward numerical scheme may fail to preserve exactly these steady states. The well-balanced schemes are introduced to preserve exactly, at a discrete level, some of these equilibrium solutions. One major advantage of the well-balanced schemes is that they can accurately resolve small perturbations to such steady state solutions with relatively coarse meshes. Indeed, if a scheme cannot balance the effects of fluxes and source terms, it may introduce spurious oscillations near steady state. In order to reduce these, the grid must be refined more than necessary. On the other hand, well-balanced schemes promise to be efficient near steady state.

☆ Research is sponsored by the National Science Foundation grant DMS-1216454, ORNL's Laboratory Directed Research and Development funds, and the U.S. Department of Energy, Office of Advanced Scientific Computing Research. The work was partially performed at ORNL, which is managed by UT-Battelle, LLC, under Contract No. DE-AC05-00OR22725.

* Correspondence to: Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, United States. Fax: +1 865 241 4811.
E-mail address: xingy@math.utk.edu.

0021-9991/$ – see front matter
http://dx.doi.org/10.1016/j.jcp.2013.10.010

A prototype example investigated extensively in the literature is shallow water equations with a non-flat bottom topography, which are used to model flows in rivers and coastal areas, and have wide applications in ocean, hydraulic engineering, and atmospheric modeling. This system describes the flow as a conservation law with an additional source term due to the bottom topography. In one space dimension, it takes the form

$$\begin{cases} h_t + (hu)_x = 0, \\ (hu)_t + \left(hu^2 + \frac{1}{2}gh^2\right)_x = -ghb_x, \end{cases} \tag{1.2}$$

where h denotes the water height, u is the velocity of the fluid, b represents the bottom topography and g is the gravitational constant. Only the source term due to the bottom topography is taken into account in this system, but other terms could also be added in order to include effects such as friction on the bottom as well as variations of the channel width.

Research on well-balanced numerical methods for the shallow water system has attracted many attentions in the past two decades. Many researchers have developed a large number of well-balanced methods to exactly preserve the still water at rest steady state

$$u = 0 \quad \text{and} \quad h + b = \text{const}, \tag{1.3}$$

see, e.g. [1,2,6,5,14–16,19–21,25,26,34,35], and the recent review paper [23] for more details. Shallow water equations (1.2) also admit the general moving water equilibrium, given by

$$m := hu = \text{const} \quad \text{and} \quad E := \frac{1}{2}u^2 + g(h + b) = \text{const}, \tag{1.4}$$

where m, E are the moving water equilibrium variables. Still water steady state is simply a special case of this, when the velocity reduces to zero. Most well-balanced methods for the still water steady state cannot preserve the moving water equilibrium, and it is significantly more difficult to obtain well-balanced schemes for such equilibrium.

In a recent paper [39], we have shown several numerical examples to demonstrate the advantage of moving-water well-balanced schemes over still-water well-balanced schemes for the shallow water equations. Those numerical examples clearly demonstrate the importance of utilizing moving-water well-balanced methods for solutions near a moving-water equilibrium. There have been a few attempts in developing well-balanced methods for the moving water equilibrium. In [14], Gosse developed a class of first order accurate flux-vector-splitting schemes based on the theory of non-conservative products, which is well-balanced for general steady states including moving water equilibria. Jin and Wen proposed the interface type method [16–18,33] to capture the general equilibria with second order accuracy. Russo [27] developed well-balanced central schemes on staggered grids which are second order accurate. Bouchut and Morales [4] proposed numerical methods based on local subsonic steady state reconstruction, which are exactly well-balanced for subsonic moving equilibria. A few high order accurate well-balanced methods for the moving water equilibrium have been introduced recently. In [22], well-balanced finite volume weighted essentially non-oscillatory (WENO) methods are designed for arbitrary equilibria of the shallow water equations. The key component there is a special way to recover the moving water equilibrium and a well-balanced quadrature rule of the source term. Russo and Khe [28] presented well-balanced central WENO schemes, and recently, Castro et al. [7] proposed a new well-balanced finite volume WENO scheme in the framework of path-conservative methods.

In recent years, high order accurate numerical schemes, including finite difference/volume WENO schemes, spectral methods and discontinuous Galerkin (DG) methods, have been developed to reduce the number of computational cells and minimize the computational time to achieve the desired resolution. Among these methods, DG method is a class of finite element methods using discontinuous piecewise polynomial space as the solution and test function spaces (see [8] for a historic review). It combines advantages of both finite element and finite volume methods, and has been successfully applied to a wide range of applications. Several advantages of the DG method, including its accuracy, high parallel efficiency, flexibility for hp-adaptivity and arbitrary geometry and meshes, make it particularly suited for the shallow water equations [12,11,13,36].

However, high order moving-water well-balanced WENO methods developed in [22] and [7] cannot be generalized to DG methods directly. The first difficulty encountered comes from the fact that the solution function space in DG methods is the piecewise polynomial space, while the water height h in a moving water equilibrium is in general not a polynomial. Other difficulty includes that the approximation of the source term now becomes the integral of the source term multiplied by the test function, therefore the quadrature rule idea in [22] cannot be applied directly. The main objective of this paper is to develop positivity-preserving high order accurate well-balanced DG methods for the shallow water equations with moving water equilibrium. This will be the first paper to achieve this goal, to our best knowledge. We first present the transformation between the conservative variables and equilibrium variables following the techniques presented in [22]. The recovery of well-balanced states is obtained through the computation of the reference states. The numerical solution is then decomposed into the equilibrium part and the remaining part. Then, a special, well-balanced source term approximation can be derived based on this decomposition and we will show it balances the effect of numerical fluxes which are computed by a generalized hydrostatic reconstruction. All together this leads to a rather transparent formulation of our well-balanced DG scheme. Another important difficulty often encountered in the simulations of the shallow water equations is the appearance

of dry areas where no water is present. Many shallow water applications involve rapidly moving interfaces between wet and dry areas, and if no special attention is paid, standard numerical methods may produce unacceptable negative water height. Here we consider the positivity-preserving limiter proposed in [41,40,38], which preserves the high order accuracy without losing local mass and momentum conservation. A careful examination of its validity within the proposed well-balanced methods will be studied.

The novel contribution of this paper is the special well-balanced source term approximation presented in Section 3.3, as well as the choice of b^* in the hydrostatic reconstruction, which is determined (for the first time) by the positivity-preserving approach in Section 4 (see Remark 3.1 for more explanation). Although our new schemes are designed for the shallow water equations, the same idea can be generalized to obtain high order well-balanced numerical methods for any other hyperbolic balance law of the form (1.1). This paper is organized as follows. In Section 2, we present the standard notation and a simple well-balanced DG method for the still water. In Section 3, the novel high order DG methods which preserve exactly the moving water equilibrium of the shallow water equations are presented. We will also show that, when the steady state solution reduces to still water (i.e. $u = 0$), the proposed DG method becomes the one presented in Section 2 for still water. In Section 4, they are coupled with the positivity-preserving technique presented in [40], which keeps the water height non-negative. Section 5 contains extensive numerical simulation results to demonstrate the behavior of our DG methods for one-dimensional shallow water equations, verifying high order accuracy, the well-balanced property, and good resolution for smooth and discontinuous solutions. Concluding remarks are given in Section 6.

2. Well-balanced DG methods for still water

Several well-balanced DG methods for the shallow water equations with still water at rest steady state solutions (1.3) have been developed in the literature, see for example [40] for a list of references. In this section, we present a new approach to achieve well-balanced methods, by simply combining two existing approaches developed by us in [36] and [37]. The key idea to achieve well-balanced property is a non-standard source term treatment (taken from [36]) and numerical flux modification based on hydrostatic reconstruction (taken from [37]). The same structure will be generalized in Section 3 to develop well-balanced methods for the moving water equilibrium. As shown in Remark 3.3, the novel well-balanced methods for moving water, when applied to problems related to still water steady state solutions, reduce to the methods presented in this section.

We start by presenting the standard notations. For the ease of presentation, we denote the shallow water equations (1.2) by

$$U_t + f(U)_x = s(h, b),$$

where $U = (h, hu)^T$ are the conservative variables, $f(U)$ is the flux and $s(h, b)$ is the source term. The moving water equilibrium variables are denoted by

$$V = (m, E)^T = \left(hu, \frac{1}{2}u^2 + g(h + b)\right)^T. \tag{2.1}$$

We discretize the computational domain into cells $I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$, and denote the size of the j-th cell by Δx_j and the maximum mesh size by $\Delta x = \max_j \Delta x_j$. In a high order DG method, we seek an approximation, still denoted by U with an abuse of notation, which belongs to the finite dimensional space

$$V_{\Delta x} = V_{\Delta x}^k \equiv \left\{w:\ w\big|_{I_j} \in P^k(I_j),\ j = 1, \ldots, J\right\}, \tag{2.2}$$

where $P^k(I)$ denotes the space of polynomials of degree at most k and J is the total number of computational cells. We project the bottom function b into the same space $V_{\Delta x}$, to obtain an approximation which is still denoted by b, again with an abuse of notation. The standard DG method is given by

$$\int_{I_j} \partial_t U v\, dx - \int_{I_j} f(U)\partial_x v\, dx + \widehat{f}_{j+\frac{1}{2}} v\left(x_{j+\frac{1}{2}}^-\right) - \widehat{f}_{j-\frac{1}{2}} v\left(x_{j-\frac{1}{2}}^+\right) = \int_{I_j} s(h, b) v\, dx, \tag{2.3}$$

where $v(x)$ is a test function from the test space $V_{\Delta x}$, $\widehat{f}_{j+\frac{1}{2}} = F(U(x_{j+\frac{1}{2}}^-, t), U(x_{j+\frac{1}{2}}^+, t))$ and $F(a_1, a_2)$ is a numerical flux. We could, for example, use the simple Lax–Friedrichs flux

$$F(a_1, a_2) = \frac{1}{2}\big(f(a_1) + f(a_2) - \alpha(a_2 - a_1)\big), \tag{2.4}$$

where $\alpha = \max(|u| + \sqrt{gh})$ and the maximum is taken over the whole region.

We are interested in preserving the still water stationary solution (1.3) exactly. As mentioned in [37], a well-balanced numerical scheme takes the form:

$$\int_{I_j} \partial_t U^n v\,dx - \int_{I_j} f(U^n)\partial_x v\,dx + \widehat{f}^l_{j+\frac{1}{2}} v\left(x^-_{j+\frac{1}{2}}\right) - \widehat{f}^r_{j-\frac{1}{2}} v\left(x^+_{j-\frac{1}{2}}\right) = \int_{I_j} s(h^n, b) v\,dx, \tag{2.5}$$

where the numerical fluxes $\widehat{f}^{l,r}_{j+\frac{1}{2}}$ are given by (2.8)–(2.9). After computing boundary values $U^{\pm}_{j+\frac{1}{2}}$, we follow the hydrostatic reconstruction [1] to set

$$h^{*,\pm}_{j+\frac{1}{2}} = \max\left(0, h^{\pm}_{j+\frac{1}{2}} + b^{\pm}_{j+\frac{1}{2}} - \max\left(b^+_{j+\frac{1}{2}}, b^-_{j+\frac{1}{2}}\right)\right) \tag{2.6}$$

and redefine the left and right values of U as:

$$U^{*,\pm}_{j+\frac{1}{2}} = \begin{pmatrix} h^{*,\pm}_{j+\frac{1}{2}} \\ h^{*,\pm}_{j+\frac{1}{2}} u^{\pm}_{j+\frac{1}{2}} \end{pmatrix}. \tag{2.7}$$

Then the left and right fluxes $\widehat{f}^l_{j+\frac{1}{2}}$ and $\widehat{f}^r_{j-\frac{1}{2}}$ are given by:

$$\widehat{f}^l_{j+\frac{1}{2}} = F\left(U^{*,-}_{j+\frac{1}{2}}, U^{*,+}_{j+\frac{1}{2}}\right) + \begin{pmatrix} 0 \\ \frac{g}{2}(h^-_{j+\frac{1}{2}})^2 - \frac{g}{2}(h^{*,-}_{j+\frac{1}{2}})^2 \end{pmatrix}, \tag{2.8}$$

$$\widehat{f}^r_{j-\frac{1}{2}} = F\left(U^{*,-}_{j-\frac{1}{2}}, U^{*,+}_{j-\frac{1}{2}}\right) + \begin{pmatrix} 0 \\ \frac{g}{2}(h^+_{j-\frac{1}{2}})^2 - \frac{g}{2}(h^{*,+}_{j-\frac{1}{2}})^2 \end{pmatrix}. \tag{2.9}$$

As explained in [37,40], $\widehat{f}_{j+\frac{1}{2}} - \widehat{f}^l_{j+\frac{1}{2}}$ and $\widehat{f}_{j-\frac{1}{2}} - \widehat{f}^r_{j-\frac{1}{2}}$ are both high order correction terms at the level of $O(\Delta x^{k+1})$ regardless of the smoothness of the solution U. Therefore, the scheme (2.5) can be rewritten in the form whose left side is the traditional RKDG method and right side contains high order approximation to the source term. It can be shown to be a spatially $(k+1)$-th order conservative scheme and converge to the weak solution.

The source term approximation can be handled by the techniques presented in [36]. We separate the source term integral into:

$$\begin{aligned}
&-\int_{I_j} ghb_x v\,dx \\
&= -g\overline{(h+b)}_j \int_{I_j} b_x v\,dx + g\int_{I_j}\left(\frac{b^2}{2}\right)_x v\,dx - \int_{I_j} g\left(h+b-\overline{(h+b)}_j\right) b_x v\,dx \\
&= -g\overline{(h+b)}_j\left((bv)^-_{j+\frac{1}{2}} - (bv)^+_{j-\frac{1}{2}} - \int_{I_j} bv_x dx\right) + \frac{g}{2}\left(\left(b^2v\right)^-_{j+\frac{1}{2}} - \left(b^2v\right)^+_{j-\frac{1}{2}} - \int_{I_j} b^2 v_x dx\right) \\
&\quad - \int_{I_j} g\left(h+b-\overline{(h+b)}_j\right) b_x v\,dx \\
&= \frac{g}{2}\left(-2\overline{(h+b)}_j bv + b^2 v\right)^-_{j+\frac{1}{2}} - \frac{g}{2}\left(-2\overline{(h+b)}_j bv + b^2 v\right)^+_{j-\frac{1}{2}} - \int_{I_j} \frac{g}{2}\left(-2\overline{(h+b)}_j b + b^2\right) v_x dx \\
&\quad - \int_{I_j} g\left(h+b-\overline{(h+b)}_j\right) b_x v\,dx \\
&= \left(\frac{g}{2}\left(\overline{(h+b)}_j - b\right)^2 v\right)^-_{j+\frac{1}{2}} - \left(\frac{g}{2}\left(\overline{(h+b)}_j - b\right)^2 v\right)^+_{j-\frac{1}{2}} - \int_{I_j} \frac{g}{2}\left(\overline{(h+b)}_j - b\right)^2 v_x dx \\
&\quad - \int_{I_j} g\left(h+b-\overline{(h+b)}_j\right) b_x v\,dx,
\end{aligned} \tag{2.10}$$

where the second equality follows integration by parts and the last equality is obtained by adding and subtracting $\int_{I_j} \frac{g}{2}\overline{(h+b)}_j^2 v_x dx$. It is easy to prove that the above methods (2.5), combined with the choice of fluxes (2.8)–(2.9) and

source term approximation (2.10), are well-balanced for the still water steady state of the shallow water equations. We refer to [36,37] for the details of proof.

Total variation diminishing (TVD) high order Runge–Kutta time discretization [31] is used in practice for stability and to increase temporal accuracy. For example, the third order TVD Runge–Kutta method is used in the simulation in this paper:

$$\begin{aligned} U^{(1)} &= U^n + \Delta t \mathcal{F}\big(U^n\big), \\ U^{(2)} &= \frac{3}{4}U^n + \frac{1}{4}\big(U^{(1)} + \Delta t \mathcal{F}\big(U^{(1)}\big)\big), \\ U^{n+1} &= \frac{1}{3}U^n + \frac{2}{3}\big(U^{(2)} + \Delta t \mathcal{F}\big(U^{(2)}\big)\big), \end{aligned} \tag{2.11}$$

where $\mathcal{F}(U)$ is the spatial operator.

3. Well-balanced DG methods for moving water

In this section, we design high order finite element DG methods for the shallow water equations (1.2), with the objective to maintain the general moving steady state (1.4). We will concentrate on the one-dimensional case. The basic framework of the well-balanced scheme follows the one introduced in Section 2. However, extra attention is required to handle the flux and source term approximation due to the complexity of moving water equilibrium.

3.1. Conservative and equilibrium variables

In order to construct our well-balanced scheme, it is essential to transform the conservative variables U into the equilibrium variables V and vice versa. Here we follow the technique presented by us in [22], and a similar variable transform has also been used in [17,18,27].

Given conservative variables U and a bottom function b, the energy E, and hence the equilibrium variables $V = V(U)$, can be easily computed by (2.1). The main difficulty comes from the inverse transform $U = U(V)$. As in [22], we first label the different flow regimes by defining the sign function

$$\sigma := \operatorname{sign}(Fr - 1), \qquad Fr := |u|/\sqrt{gh}, \tag{3.1}$$

and a state is called sonic, sub- or supersonic if σ is zero, negative or positive.

With given $V = (m, E)$, b and σ, one can recover the conservative variable h and establish the transform $U = U(V)$. If $m = 0$, the transformation is trivial. Assume m is a fixed nonzero parameter, we define the function φ by

$$\varphi(h) := \frac{m^2}{2h^2} + gh, \tag{3.2}$$

which achieves its minimum value φ_0 at h_0:

$$\varphi_0 = \frac{3}{2}\big(g|m|\big)^{2/3}, \qquad gh_0 = \big(g|m|\big)^{2/3}. \tag{3.3}$$

Therefore, we have the following result (see [22] for details):

Lemma 3.1. *Let m be given, and suppose either $E = \varphi_0 + gb$ if $\sigma = 0$ or $E > \varphi_0 + gb$ if $\sigma = \pm 1$. Then there exists a unique solution*

$$h = h(E, b, \sigma) \tag{3.4}$$

such that

$$\begin{cases} h < h_0 & \text{for } \sigma = 1 \text{ (supersonic flow)}, \\ h = h_0 & \text{for } \sigma = 0 \text{ (sonic flow)}, \\ h > h_0 & \text{for } \sigma = -1 \text{ (subsonic flow)}. \end{cases} \tag{3.5}$$

Given the set (E, b, σ) satisfying the condition in Lemma 3.1, it is straightforward to compute the corresponding solution h by Newton's method, and we refer to [22] for the strategy.

3.2. Recovery of well-balanced states

The first difficulty encountered in designing well-balanced methods is the recovery of well-balanced states from the provided initial equilibrium data. Let's assume the initial data are in perfect equilibrium, i.e. $V(x) \equiv \bar{V}$ for some equilibrium state $\bar{V}$. However, the initial condition $U_j(x)$ and $b_j(x)$ (in the DG setting, they are the L^2 projection of the initial data into the piecewise polynomial space $V_{\Delta x}$) may not be in equilibrium any more. Therefore, the cell boundary values $U^{\pm}_{j+\frac{1}{2}}$, which will be used to evaluate the numerical fluxes at the cell boundary, may not be in equilibrium, and this will increase the complexity in designing well-balanced methods.

A strategy has been proposed in [22] to recover the well-balanced states. Here we summarize it briefly and refer the readers to [22] for the details. Assume that the cell is in equilibrium with constant values (m, E). Therefore, $m_j(x) = \bar{m}_j = m$, where $\bar{m}_j$ is the cell average of $m_j(x)$ in cell I_j. From (3.3), the minimal possible value of E at x is $E_0(x) = \frac{3}{2}(g|m|)^{2/3} + gb(x)$, and we denote its cell average to be $E_j^{min} := \frac{3}{2}(g|\bar{m}_j|)^{2/3} + g\bar{b}_j$. We then define two more average heights related to cell I_j, one the maximal supersonic average height h_j^+ and the other the minimal subsonic average height h_j^- via

$$h_j^+ := \frac{1}{\Delta x_j}\int_{I_j} h\big(\bar{m}_j, E_j^{min}, b(x), 1\big)\,dx, \tag{3.6}$$

$$h_j^- := \frac{1}{\Delta x_j}\int_{I_j} h\big(\bar{m}_j, E_j^{min}, b(x), -1\big)\,dx. \tag{3.7}$$

Note that $h_j^+ < h_0 < h_j^-$, where h_0 is the critical, or sonic, point. Let $\bar{h}_j$ be the cell average of $h_j(x)$ in cell I_j. Now we define a reference equilibrium value $\bar{E}_j$ as follows:

Definition 3.1.

(i) If $\bar{h}_j < h_j^+$, define $E = \bar{E}_j$ to be the unique solution of

$$\bar{h}_j = \frac{1}{\Delta x_j}\int_{I_j} h\big(\bar{m}_j, E, b(x), 1\big)\,dx. \tag{3.8}$$

(ii) If $\bar{h}_j > h_j^-$, define $E = \bar{E}_j$ to be the unique solution of

$$\bar{h}_j = \frac{1}{\Delta x_j}\int_{I_j} h\big(\bar{m}_j, E, b(x), -1\big)\,dx. \tag{3.9}$$

(iii) If $h_j^+ \leqslant \bar{h}_j \leqslant h_j^-$, set

$$\bar{E}_j := E_j^{min}. \tag{3.10}$$

In this case, the equilibrium solution becomes sonic for some point $x^* \in I_j$, and the sonic point x^* is located at the maximum of b over the cell I_j.

One can show that $\bar{V}_j = (\bar{m}_j, \bar{E}_j)$ recovers the equilibrium state. We refer to [22] for the detailed proof, as well as the Newton iteration to compute $\bar{E}_j$ numerically.

3.3. Source term approximation

The main structure of well-balanced methods for moving water equilibrium (1.4) follows the one (2.5) for still water. One important ingredient in designing well-balanced methods is the approximation of the source term integral in (2.5). In this subsection, let us discuss the well-balanced approach to compute the source term.

By Definition 3.1, one can compute the reference equilibrium values $\bar{V}_j$, which lead to the reference equilibrium functions $U(\bar{V}_j, b(x))$. Since they may not be polynomials, we consider their L^2 projection into the finite element space $V_{\Delta x}$ which was introduced in (2.2), and denote it by

$$U_j^e(x) = \big(h_j^e(x), m_j^e(x)\big) = PU\big(\bar{V}_j, b(x)\big) \tag{3.11}$$

in each cell I_j, where P denotes the L^2 projection operator. Therefore, the numerical solutions U, which are piecewise polynomials, can be decomposed as

$$U = U^e + U^r, \tag{3.12}$$

where $U^r = U - U^e \in V_{\Delta x}$. The source term approximation now becomes

$$\int s(h,b)v\,dx = \int s\big(h^e,b\big)v\,dx + \int s\big(h^r,b\big)v\,dx, \tag{3.13}$$

since $s(h,b) = -ghb_x$ is linear with respect to h. The second term on the right hand side of (3.13) can be computed by the standard quadrature rule. Next, let us discuss how to approximate the first term numerically. Noticing the fact that $U(\bar V_j, b) = (h(\bar V_j, b), \bar m_j)^T$ is the equilibrium state, we have the relation

$$\int_{I_j} s\big(h(\bar V_j,b),b\big)v\,dx = -\int_{I_j} f\big(U(\bar V_j,b)\big)v_x\,dx + f\big(U\big(\bar V_j, b^-_{j+\frac12}\big)\big)v^-_{j+\frac12} - f\big(U\big(\bar V_j, b^+_{j-\frac12}\big)\big)v^+_{j-\frac12}. \tag{3.14}$$

Since U^e is the L^2 projection of $U(\bar V_j, b)$, we conclude that

$$\int_{I_j} s\big(h^e,b\big)v\,dx + O\big(\Delta x^{k+1}\big) = -\int_{I_j} f\big(U^e\big)v_x\,dx + f\big(U^{e,-}_{j+\frac12}\big)v^-_{j+\frac12} - f\big(U^{e,+}_{j-\frac12}\big)v^+_{j-\frac12} \tag{3.15}$$

and can approximate the source term integral (3.13) by

$$\int_{I_j} s(h,b)v\,dx \approx -\int_{I_j} f\big(U^e\big)v_x\,dx + f\big(U^{e,-}_{j+\frac12}\big)v^-_{j+\frac12} - f\big(U^{e,+}_{j-\frac12}\big)v^+_{j-\frac12} + \int s\big(h^r,b\big)v\,dx. \tag{3.16}$$

Since U^e is always smooth inside a cell, the relation (3.15) is always true regardless of the smoothness of the solution U. Therefore, numerical methods with this source term approximation (3.16) will satisfy the Lax–Wendroff theorem and converge to the weak solution.

3.4. Numerical fluxes via hydrostatic reconstruction

In this subsection, we discuss an important and last piece of our method, namely the well-balanced numerical fluxes. They are computed by a generalized hydrostatic reconstruction. The idea of hydrostatic reconstruction was first introduced by Audusse et al. in [1] to provide stable well-balanced methods for the still water steady state, and was later used in many well-balanced methods [21,37], including some recent well-balanced methods for moving water [22,7]. Here, we consider its extension in our well-balanced DG methods.

At each time step t^n, one can compute the cell boundary values $U^{\pm}_{j+\frac12}$ from the solution $U(x)$. But in the case of moving water equilibrium, suppose $U(x)$ are computed from the exact solution, these cell boundary values $U^{\pm}_{j+\frac12}$ do not equal the exact solution value at the same point, as $U(x)$ is the projection of the exact solution into the polynomial space and this projection does not preserve the equilibrium state. To overcome this problem, we redefine an updated boundary value as:

$$\tilde U^{\pm}_{j+\frac12} = U\big(\bar V_j, b^{\pm}_{j+\frac12}\big) + U^{r,\pm}_{j+\frac12}, \tag{3.17}$$

where U^r is defined in (3.12). One can easily verify that $\tilde U^{\pm}_{j+\frac12} = U(\bar V_j, b^{\pm}_{j+\frac12})$ in the case of moving water equilibrium.

Next, we follow the idea of hydrostatic reconstruction to compute the numerical fluxes. We define

$$\tilde V^{\pm}_{j+\frac12} = V\big(\tilde U^{\pm}_{j+\frac12}, b^{\pm}_{j+\frac12}\big), \tag{3.18}$$

and

$$b^*_{j+\frac12} = \begin{cases} \max(b^+_{j+\frac12}, b^-_{j+\frac12}), & \text{if } \sigma_{i+\frac12} = -1, 0, \\ \min(b^+_{j+\frac12}, b^-_{j+\frac12}), & \text{if } \sigma_{i+\frac12} = 1. \end{cases} \tag{3.19}$$

The cell boundary values (used to evaluate the numerical fluxes) are defined by:

$$U^{*,\pm}_{j+\frac12} = \big(\max\big(0, h\big(\tilde V^{\pm}_{j+\frac12}, b^*_{j+\frac12}\big)\big), \tilde m^{\pm}_{j+\frac12}\big)^T = \big(\max\big(0, h\big(\tilde V^{\pm}_{j+\frac12}, b^*_{j+\frac12}\big)\big), m^{\pm}_{j+\frac12}\big)^T, \tag{3.20}$$

as one can easily observe that $\tilde m^{\pm}_{j+\frac12} = m^{\pm}_{j+\frac12}$. At the end, the left and right fluxes $\widehat f^l_{j+\frac12}$, $\widehat f^r_{j-\frac12}$ are given by:

$$\begin{aligned} \widehat f^l_{j+\frac12} &= F\big(U^{*,-}_{j+\frac12}, U^{*,+}_{j+\frac12}\big) + f\big(U^-_{j+\frac12}\big) - f\big(U^{*,-}_{j+\frac12}\big), \\ \widehat f^r_{j-\frac12} &= F\big(U^{*,-}_{j-\frac12}, U^{*,+}_{j-\frac12}\big) + f\big(U^+_{j-\frac12}\big) - f\big(U^{*,+}_{j-\frac12}\big). \end{aligned} \tag{3.21}$$

Remark 3.1. In the original hydrostatic reconstruction [1], the maximum is taken when evaluating the b^* in (3.19). The well-balanced WENO methods in [22] used the minimal value. If the purpose is to achieve well-balanced property, both maximum and minimal values would work well and there is no unified way to define them. In (3.19), a different choice of b^* for sub/supersonic flow is presented. We would like to comment that this choice is motivated by the purpose of positivity preserving. We refer to Section 4, more specifically, the proof of Lemma 4.1 for more details.

3.5. Summary of the well-balanced scheme

We now summarize the complete procedure of our high order well-balanced DG methods for solving the shallow water equation (1.2) with moving water equilibrium. The methods are given by

$$\int_{I_j} \partial_t U^n v\,dx - \int_{I_j} f\big(U^n\big)\partial_x v\,dx + \widehat{f}^l_{j+\frac{1}{2}} v\big(x^-_{j+\frac{1}{2}}\big) - \widehat{f}^r_{j-\frac{1}{2}} v\big(x^+_{j-\frac{1}{2}}\big) = \int_{I_j} s\big(h^n, b\big) v\,dx, \tag{3.22}$$

where the numerical fluxes $\widehat{f}^l$ and $\widehat{f}^r$ are computed in (3.21), and the source term is approximated by (3.16). The scheme is completed by a temporal TVD Runge–Kutta discretization (2.11).

A detailed implementation of this algorithm consists of the following steps:

1. Compute the L^2 projection of U and bottom b based on the initial data.
2. At each time step, compute the reference value $\bar{V}_j$ as the implicit solution of Eqs. (3.8)–(3.10).
3. Decompose the solution U as the sum of U^e and U^r following (3.11)–(3.12).
4. Compute the numerical fluxes following (3.17)–(3.21).
5. Evaluate the source term approximation (3.16).
6. Apply a TVD Runge–Kutta method (2.11) to advance in time.

We have the following well-balanced property of our methods.

Proposition 3.2. *The RKDG schemes described above maintain smooth moving water equilibrium (1.4) exactly, if the stationary shocks separating the smooth regions are all located at cell boundaries, and computed by Roe's numerical flux function.*

Proof. Suppose that the initial data are moving steady state equilibria, i.e. $V(x) \equiv \bar{V}$. Then Lemma 3.1 implies that all reference values $\bar{V}_i$ coincide with $\bar{V}$. Therefore U^e computed from $\bar{V}_j$ is equivalent to U, and we have $U^r = 0$ in (3.12). The source term approximation (3.16) then becomes

$$\int_{I_j} s(h,b) v\,dx = -\int_{I_j} f(U) v_x\,dx + f\big(U^-_{j+\frac{1}{2}}\big) v^-_{j+\frac{1}{2}} - f\big(U^+_{j-\frac{1}{2}}\big) v^+_{j-\frac{1}{2}}. \tag{3.23}$$

For the numerical fluxes, we have $\tilde{U}^{\pm}_{j+\frac{1}{2}} = U(\bar{V}_j, b^{\pm}_{j+\frac{1}{2}})$ in (3.17), which leads to $\tilde{V}^{\pm}_{j+\frac{1}{2}} = \bar{V}$, and consequently, $U^{*,-}_{j+\frac{1}{2}} = U^{*,+}_{j+\frac{1}{2}}$ in (3.19). Imposing the results into (3.21), we have

$$\widehat{f}^l_{j+\frac{1}{2}} = f\big(U^-_{j+\frac{1}{2}}\big), \qquad \widehat{f}^r_{j-\frac{1}{2}} = f\big(U^+_{j-\frac{1}{2}}\big). \tag{3.24}$$

One can easily observe that numerical fluxes exactly balance the source term approximation (3.23), and that is the desired well-balanced property. □

At the end of this section, we would like to provide a few remarks about the proposed well-balanced RKDG methods.

Remark 3.2. When the bottom function b is flat, our well-balanced DG methods become the traditional DG methods. It can be seen based on the following facts. First of all, when $b = 0$, the equilibrium part U^e becomes a constant in each cell. So the source term approximation in (3.16) is equal to zero. Second, one can easily verify that $\tilde{U}^{\pm}_{j+\frac{1}{2}}$ defined in (3.17) equals $U^{\pm}_{j+\frac{1}{2}}$, and therefore, it equals $U^{*,\pm}_{j+\frac{1}{2}}$. So the well-balanced numerical fluxes defined in (3.21) return to the traditional DG numerical fluxes.

Remark 3.3. Our well-balanced methods are designed to preserve the moving water equilibrium (1.4). When applied to still water steady state (1.3), they become the existing well-balanced methods presented in Section 2. This fact can be easily verified. In the case of still water steady state, $\bar{V}_j$ becomes $(\overline{(h+b)}_j, 0)^T$, and the decomposition of U in (3.11) takes the form of

$$h = h^e + h^r = \big(\overline{(h+b)}_j - b\big) + \big(h + b - \overline{(h+b)}_j\big), \qquad hu = (hu)^e + (hu)^r = 0 + hu.$$

Therefore, the source term approximation (3.16) is exactly the same as the approximation in (2.10). Simple calculation leads to the fact that $\tilde{U}^{\pm}_{j+\frac{1}{2}}$ defined in (3.17) equals $U^{\pm}_{j+\frac{1}{2}}$, so the numerical fluxes defined in (3.21) are also equivalent to the ones defined in (2.8)–(2.9).

Remark 3.4. Another important ingredient for the DG methods is that a slope limiter procedure might be needed after each inner stage in the Runge–Kutta time stepping, when the solution contains discontinuities. We use the characteristic-wise total variation bounded (TVB) limiter presented in [10,30]. For the shallow water system, we perform the limiting in the local characteristic variables. However, this limiter procedure might destroy the preservation of the moving water equilibrium state (1.4). In order to achieve the well-balanced property, we follow the idea presented in [40] and propose the following way to perform the TVB limiter. Note that the TVB limiter procedure actually involves two steps: the first one is to check whether any limiting is needed in a specific cell; and, if the answer is yes, the second step is to apply the TVB limiter on the variables in this cell. So we propose to first check if the limiting is needed based on the equilibrium variables V. If a certain cell is flagged by this procedure needing limiting, then the actual TVB limiter is implemented on U. Note that in an equilibrium region, we first check if the limiting is needed based on $V = \text{const}$, which demonstrates that limiting is not needed in this cell. Therefore the water height h will not be affected by the limiter procedure and the well-balanced property is maintained.

4. A positivity-preserving limiter

A simple positivity-preserving limiter has been proposed and implemented for the shallow water equations in [40] for the DG method and in [38] for the WENO scheme. It was shown to be able to keep the water height non-negative under suitable CFL condition, preserve the mass conservation and at the same time does not affect the high order accuracy for the general solutions. In this section, we will explore how to couple this limiter with well-balanced DG methods presented in Section 3. We will refer to [40] for the detailed explanation of this limiter, and only highlight the main difference in the following part.

We first consider the scheme satisfied by the cell averages of the water height in the well-balanced DG methods with the Euler forward discretization in time:

$$\bar{h}^{n+1}_j = \bar{h}^n_j - \lambda\big[\widehat{F}\big(h^{*,-}_{j+\frac{1}{2}}, m^-_{j+\frac{1}{2}}; h^{*,+}_{j+\frac{1}{2}}, m^+_{j+\frac{1}{2}}\big) - \widehat{F}\big(h^{*,-}_{j-\frac{1}{2}}, m^-_{j-\frac{1}{2}}; h^{*,+}_{j-\frac{1}{2}}, m^+_{j-\frac{1}{2}}\big)\big], \tag{4.1}$$

where

$$\widehat{F}\big(h^{*,-}_{j+\frac{1}{2}}, m^-_{j+\frac{1}{2}}; h^{*,+}_{j+\frac{1}{2}}, m^+_{j+\frac{1}{2}}\big) = \frac{1}{2}\big(m^-_{j+\frac{1}{2}} + m^+_{j+\frac{1}{2}} - \alpha\big(h^{*,+}_{j+\frac{1}{2}} - h^{*,-}_{j+\frac{1}{2}}\big)\big), \tag{4.2}$$

with $h^{*,\pm}_{j+\frac{1}{2}}$ defined in (3.19), and $\lambda = \Delta t/\Delta x$. As shown in [40,38,41], one building block of the positivity-preserving limiter is to show the positivity of the first order version of the proposed numerical methods.

Lemma 4.1. *Under the CFL condition $\lambda\alpha \leqslant 1$, with $\alpha = \max(|u| + \sqrt{gh},\ |hu|/h^{*,\pm} + \sqrt{gh})$, consider the following scheme*

$$h^{n+1}_j = h^n_j - \lambda\big[\widehat{F}\big(h^{*,+}_j, m^n_j; h^{*,-}_{j+1}, m^n_{j+1}\big) - \widehat{F}\big(h^{*,+}_{j-1}, m^n_{j-1}; h^{*,-}_j, m^n_j\big)\big] \tag{4.3}$$

with $\widehat{F}$ defined in (4.2),

$$b^*_{j+\frac{1}{2}} = \begin{cases} \max(b_{j+1}, b_j), & \text{if } \sigma_{j+\frac{1}{2}} = -1, 0, \\ \min(b_{j+1}, b_j), & \text{if } \sigma_{j+\frac{1}{2}} = 1, \end{cases} \tag{4.4}$$

and

$$h^{*,+}_j = \max\big(0, h\big(\tilde{V}_j, b^*_{j+\frac{1}{2}}\big)\big), \qquad h^{*,-}_j = \max\big(0, h\big(\tilde{V}_j, b^*_{j-\frac{1}{2}}\big)\big), \tag{4.5}$$

where $\tilde{V}_j = m^2_j/h_i + gh^2_j/2$. If h^n_j, $h^n_{j\pm1}$ are non-negative, then h^{n+1}_j is also non-negative.

Proof. First of all, let's show that $h^{*,\pm}_j \leqslant h^n_j$. In a subsonic flow region, i.e. $\sigma = -1$, one knows that $\varphi(h)$, as defined in (3.2), is a decreasing function in h. Therefore, since $b^*_{j+\frac{1}{2}} \geqslant b_j$, we have $h^{*,+}_j \leqslant h^n_j$. Similarly, we have the same conclusion in a supersonic flow region.

Next, the scheme (4.3) can be written as

$$h_j^{n+1} = \left[1 - \frac{1}{2}\lambda\left(\alpha + \frac{m_j^n}{h_j^{*,+}}\right)\frac{h_j^{*,+}}{h_j^n} - \frac{1}{2}\lambda\left(\alpha - \frac{m_j^n}{h_j^{*,-}}\right)\frac{h_j^{*,-}}{h_j^n}\right]h_j^n + \left[\frac{1}{2}\lambda\left(\alpha + \frac{m_{j-1}^n}{h_{j-1}^{*,+}}\right)\frac{h_{j-1}^{*,+}}{h_{j-1}^n}\right]h_{j-1}^n + \left[\frac{1}{2}\lambda\left(\alpha - \frac{m_{j+1}^n}{h_{j+1}^{*,-}}\right)\frac{h_{j+1}^{*,-}}{h_{j+1}^n}\right]h_{j+1}^n.$$

Therefore, h_j^{n+1} is a linear combination of h_{j-1}^n, h_j^n and h_{j+1}^n and all the coefficients are non-negative since $0 \leqslant h_j^{*,\pm} \leqslant h_j^n$. Thus, $h_j^{n+1} \geqslant 0$. □

We now consider high order schemes. From now on, everything follows exactly the same as in [40], and only the main idea is presented here. We introduce the N-point (with $2N-3 \geqslant k$) Legendre Gauss–Lobatto quadrature rule on the interval $I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$, and denote these quadrature points

$$S_j = \left\{x_{j-\frac{1}{2}} = \hat{x}_j^1, \hat{x}_j^2, \ldots, \hat{x}_j^{N-1}, \hat{x}_j^N = x_{j+\frac{1}{2}}\right\},$$

with the corresponding quadrature weights $\widehat{w}_r$ for the interval $[-1/2, 1/2]$ satisfying $\sum_{r=1}^N \widehat{w}_r = 1$. Following the approaches in [24,41,40,38], we have the following result (see [40] for the details of the proof):

Proposition 4.2. *Consider the scheme (4.1) satisfied by the cell averages of the water height in our DG method. Let $h_j^n(x)$ be the DG polynomial for the water height in the cell I_j. If $h_{j-\frac{1}{2}}^-$, $h_{j+\frac{1}{2}}^+$ and $h_j^n(\hat{x}_j^r)$ $(r = 1, \ldots, N)$ are all non-negative, then $\bar{h}_j^{n+1}$ is also non-negative under the CFL condition*

$$\lambda\alpha \leqslant \widehat{w}_1. \tag{4.6}$$

To enforce the conditions of this proposition, we need to modify $h_j^n(x)$ by employing the following limiter [41,40] on the DG polynomial $U_j^n(x) = (h_j^n(x), (hu)_j^n(x))^T$, which is essentially a linear scaling around its cell average:

$$\widetilde{U}_j^n(x) = \theta\left(U_j^n(x) - \overline{U}_j^n\right) + \overline{U}_j^n, \quad \theta = \min\left\{1, \frac{\bar{h}_j^n}{\bar{h}_j^n - m_j}\right\}, \tag{4.7}$$

with

$$m_j = \min_{x\in S_j} h_j^n(x) = \min_{r=1,\ldots N} h_j^n\left(\hat{x}_j^r\right). \tag{4.8}$$

It is easy to observe that $\widetilde{h}_j^n(\hat{x}_j^r) \geqslant 0$ $(r = 1, \ldots, N)$. We compute the modified polynomial $\widetilde{U}_j^n(x)$ and use $\widetilde{U}_j^n(x)$ instead of $U_j^n(x)$ in the scheme (2.5). Hence by this proposition, $\bar{h}_j^{n+1}$ at time level $n+1$ is also non-negative, therefore (4.7) is indeed a positivity-preserving limiter for the well-balanced DG methods. Moreover, it can also be shown that this limiter preserves the local mass conservation, and does not destroy the high order accuracy. We refer to [40,38] for the proofs, as well as other comments regarding this positivity-preserving limiter.

5. Numerical examples

In this section we present numerical results of our well-balanced positivity-preserving DG methods for the one-dimensional shallow water equations (1.2). We implemented the third order finite element RKDG method (i.e. $k = 2$), coupled with the third order TVD Runge–Kutta time discretization (2.11) in these examples. The CFL number is taken as 0.16, less than $\widehat{w}_1 = 1/6$ in the corresponding CFL condition (4.6) when $k = 2$ is used. Recall that the CFL condition for linear stability for the DG methods is $\lambda\alpha \leqslant 1/5$ for $k = 2$, which are comparable to our CFL restriction. The TVB constant M in the TVB limiter is taken as 0 in most numerical examples, unless the accuracy test in which a nonzero M is used to avoid the accuracy order reduction near the extrema. The gravitation constant g is fixed as 9.812 m/s^2. To measure the extra cost of our well-balancing methods, we have compared runtimes of our methods and the positivity-preserving DG methods presented in [40], for the first numerical test below. The CPU time of the new method is about twice of that of the DG method in [40], and most of the extra time is spent on the Newton iteration to compute water height h from the equilibrium variables E and m.

Table 1
L^1 and L^∞ errors for three test problems in Section 5.1.

Test	L^1 error		L^∞ error	
	h	hu	h	hu
(a)	4.43E−14	2.62E−14	1.01E−13	4.17E−13
(b)	5.74E−14	1.97E−14	3.94E−12	2.38E−13
(c)	1.46E−14	1.54E−14	2.80E−12	2.77E−12

5.1. Test for the well-balanced property

We test our proposed DG schemes on a moving water steady state problem with a non-flat bottom, to verify the well-balanced property. These steady state problems are classical test cases for transcritical and subcritical flows, and have been widely used to test numerical schemes for shallow water equations [32,34,22].

The bottom function is given by:

$$b(x) = \begin{cases} 0.2 - 0.05(x-10)^2 & \text{if } 8 \leqslant x \leqslant 12, \\ 0 & \text{otherwise} \end{cases} \tag{5.1}$$

for a channel of length 25 meters. Three steady states, subcritical or transcritical flow with or without a steady shock will be investigated. As proven in Proposition 3.2, these equilibrium states should be exactly preserved, and we will verify them numerically.

(a) *Subcritical flow*. The initial condition is given by:

$$E = 22.06605, \qquad m = 4.42, \tag{5.2}$$

together with the boundary condition of $m = 4.42$ at the upstream, and $h = 2$ at the downstream.

(b) *Transcritical flow without a shock*. The initial condition is given by:

$$E = \frac{1.53^2}{2 \times 0.66^2} + 9.812 \times 0.66, \qquad m = 1.53, \tag{5.3}$$

together with the boundary condition of $m = 1.53$ at the upstream, and $h = 0.66$ at the downstream when the flow is subsonic.

(c) *Transcritical flow with a shock*. The initial condition is given by:

$$E = \begin{cases} \frac{3}{2}(9.812 \times 0.18)^{\frac{2}{3}} + 9.812 \times 0.2 & \text{if } x \leqslant 11.665504281554291, \\ \frac{0.18^2}{2 \times 0.33^2} + 9.812 \times 0.33 & \text{otherwise}, \end{cases} \qquad m = 0.18, \tag{5.4}$$

together with the boundary condition of $m = 0.18$ at the upstream, and $h = 0.33$ at the downstream. As illustrated in Proposition 3.2, we only discuss the case when the shock is exactly located at the cell boundary, and need to use the Roe's flux to compute the approximate Riemann problem. Hence we shift the computational domain to put the shock at the cell boundary in this computation.

For all these three cases, the steady states should be exactly preserved. We compute their solutions until $t = 5$ using $N = 200$ uniform mesh points. In order to demonstrate that the steady state is maintained up to round-off error, we show the L^1 and L^∞ errors for the water height h and the discharge hu (note: h in this case is a constant or polynomial function!) in Table 1, where we can clearly see that the L^1 and L^∞ errors are at the level of round-off errors for all three cases. This verifies the desired well-balanced property for the moving water equilibrium.

5.2. Small perturbation tests of a moving water equilibrium

The test cases in this subsection are chosen to demonstrate the capability of these schemes for computations on the small perturbation of moving water steady state solutions. We presented three moving water equilibrium solutions in the previous subsection, and showed that our numerical schemes did maintain them exactly. The main setup is kept the same here, but with an initial perturbation on the height. The same tests have been shown in [39] to demonstrate the advantage of moving-water well-balanced WENO methods over the still-water well-balanced methods.

Our initial conditions are given by imposing a small perturbation of size 0.05 on the height of these steady states in the interval $[5.75, 6.25]$. Theoretically, this disturbance should split into two waves, propagating to the left and right respectively. Many numerical methods have difficulty with the calculations involving such small perturbations of the water surface. We run the tests with 200 uniform cells with simple transmissive boundary conditions. The stopping time T is set as 1.5 for the subcritical flow and transcritical flow without a shock, 3 for the transcritical flow with a shock. At this time, the downstream-traveling water pulse has already passed the bump. The solutions of water height at the final stopping time, compared with the results using refined 1000 uniform cells, are shown in Figs. 1–3 for these three flows. We can clearly observe that there are no spurious numerical oscillations and the propagated small perturbation is very well captured by the proposed well-balanced methods.

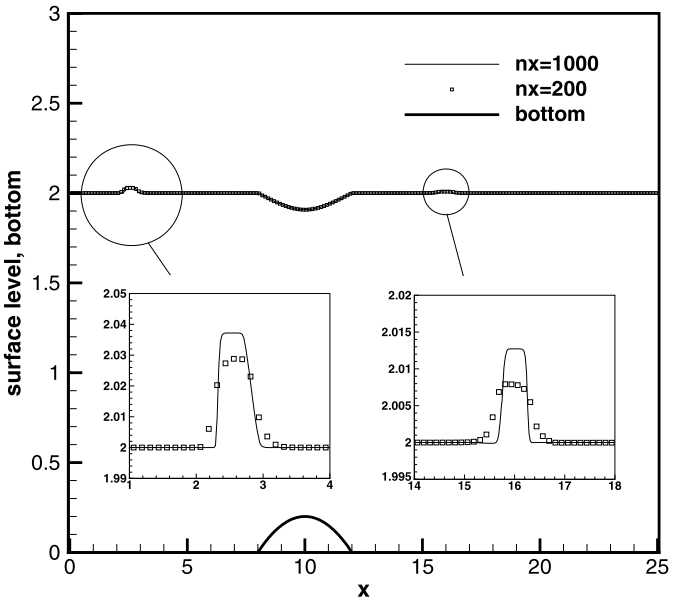

Fig. 1. Small perturbation of the subcritical flow.

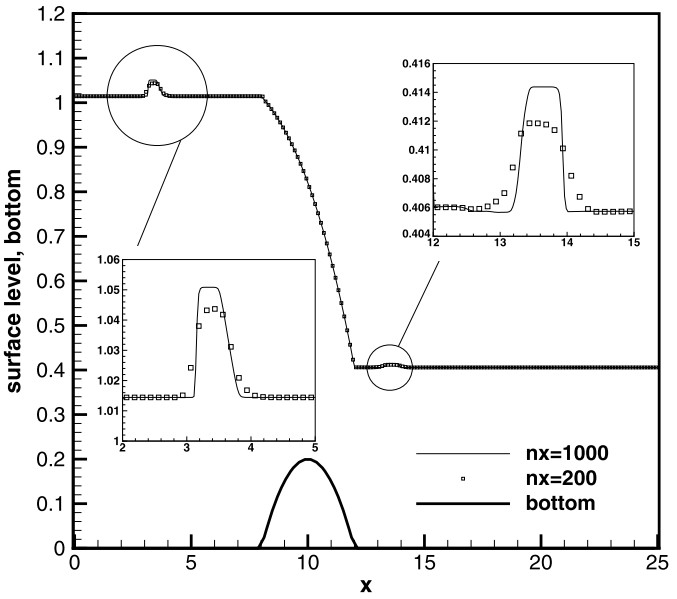

Fig. 2. Small perturbation of the transcritical flow without a shock.

5.3. Accuracy test

The high order accuracy of our proposed schemes will be tested for a smooth solution. Following the setup in [37], we have chosen the following bottom topography and initial conditions

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

with periodic boundary conditions in the domain $[0,1]$. Since the exact solution is not known explicitly for this case, we use the fifth order finite volume WENO scheme from [36] with 12,800 cells to compute a reference solution, and treat this reference solution as the exact solution in computing the numerical errors. The TVB constant M is taken a nonzero constant here, to avoid the accuracy order reduction near the extrema. In general, M should be chosen proportional to the size of the second derivative of the solution near smooth extrema, see [9]. In this example, the TVB constant $M = 480$ is picked based on the following strategy: we first observe that the accuracy is 3 when no limiter is used. If the TVB limiter with $M = 0$ is used, the accuracy decreased to 2. We then manually tune M to find the minimal value of M (which is 480 in this

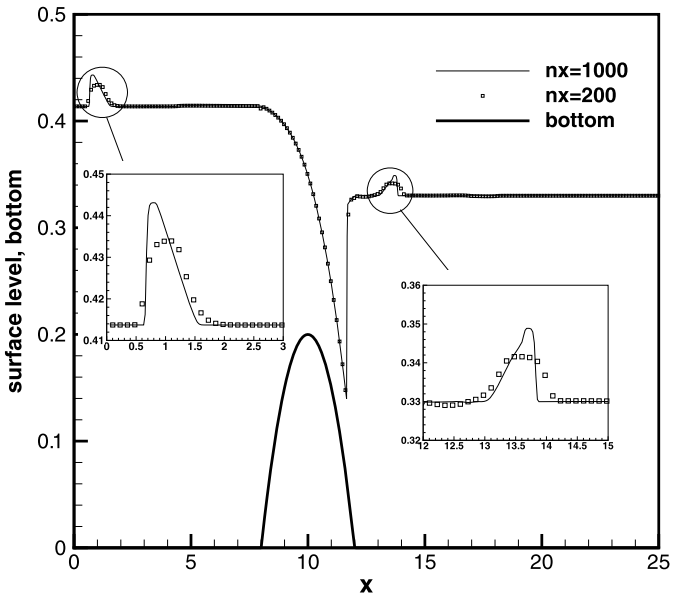

Fig. 3. Small perturbation of the transcritical flow with a shock.

Table 2
L^1 errors and numerical orders of accuracy for the example in Section 5.3.

No. of cells	h		hu	
	L^1 error	order	L^1 error	order
25	8.39E−04		6.38E−03	
50	5.13E−05	4.03	4.43E−04	3.85
100	6.19E−06	3.05	5.34E−05	3.05
200	3.87E−07	4.00	3.26E−06	4.03
400	4.70E−08	3.04	3.97E−07	3.04
800	5.81E−09	3.02	4.71E−08	3.07

case) such that the accuracy returned to 3. We compute up to $t = 0.1$ when the solution is still smooth (shocks develop later in time for this problem). Table 2 contains the L^1 errors for the cell averages and numerical orders of accuracy for the WENO scheme. We can clearly see that third order accuracy is achieved asymptotically, which verifies the high order accuracy property.

5.4. Riemann problem over a flat bottom

In this subsection, we utilize a simple Riemann problem over a flat bottom (i.e. $b(x) \equiv 0$) to demonstrate the effectiveness of the positivity-preserving limiter presented in Section 4. We set the computational domain as $[-300, 300]$, and consider the following initial conditions which contain dry region on the right half domain:

$$hu(x,0) = 0 \quad \text{and} \quad h(x,0) = \begin{cases} 10 & \text{if } x \leqslant 0, \\ 0 & \text{otherwise.} \end{cases} \tag{5.5}$$

For this Riemann problem, its analytic solution can be easily computed [3], and will be used to check the performance of the numerical solution. Using proposed well-balanced positivity-preserving DG methods with simple transmissive boundary conditions, we run the simulation until time $t = 12$ with 200 uniform cells. The numerical solutions, compared with the exact solutions, at different time $t = 4$, 8 and 12 are shown in Fig. 4. We can observe that the numerical solutions agree well with the exact solutions. The solutions near the wet/dry front are shown in Fig. 5 to provide a zoomed-in comparison.

We have also run this test case using the well-balanced DG methods, but without the positivity-preserving limiter. Shortly after the computation started, negative water height was generated, which caused blow-up immediately unless we enforce them to be 0. This confirms the proposed positivity-preserving property of our method.

5.5. Parabolic bowl

For one-dimensional shallow water equations with a parabolic bottom topography, analytic solutions have been derived by Sampson et al. [29]. This provides a good test case to validate our numerical methods.

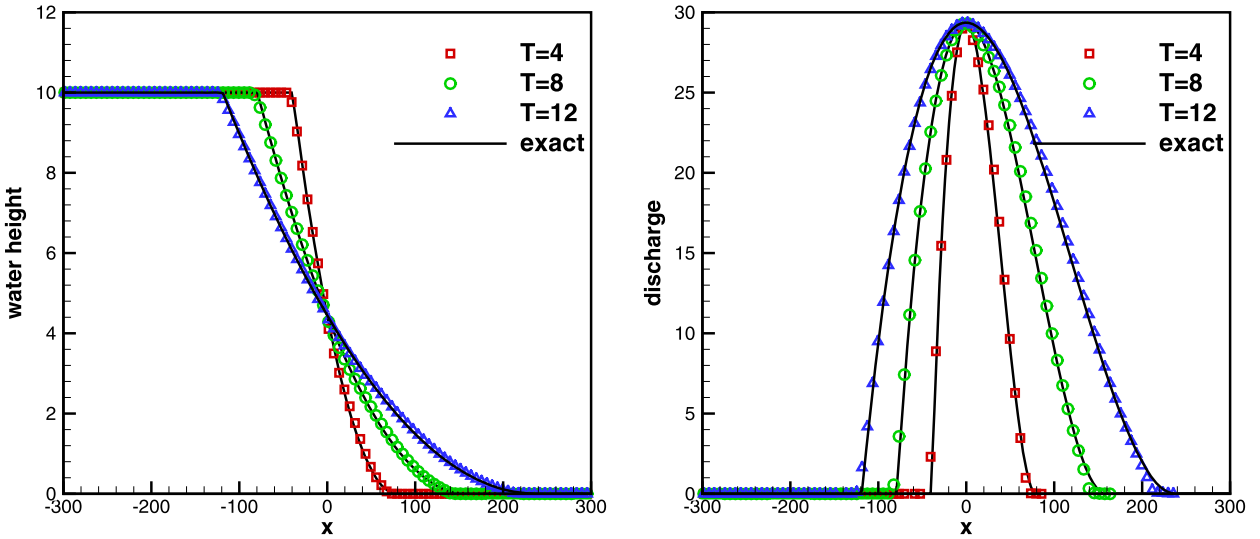

Fig. 4. The numerical and exact solutions of the Riemann problem in Section 5.4 at different time with 200 uniform cells. Left: the water height h; Right: the discharge hu.

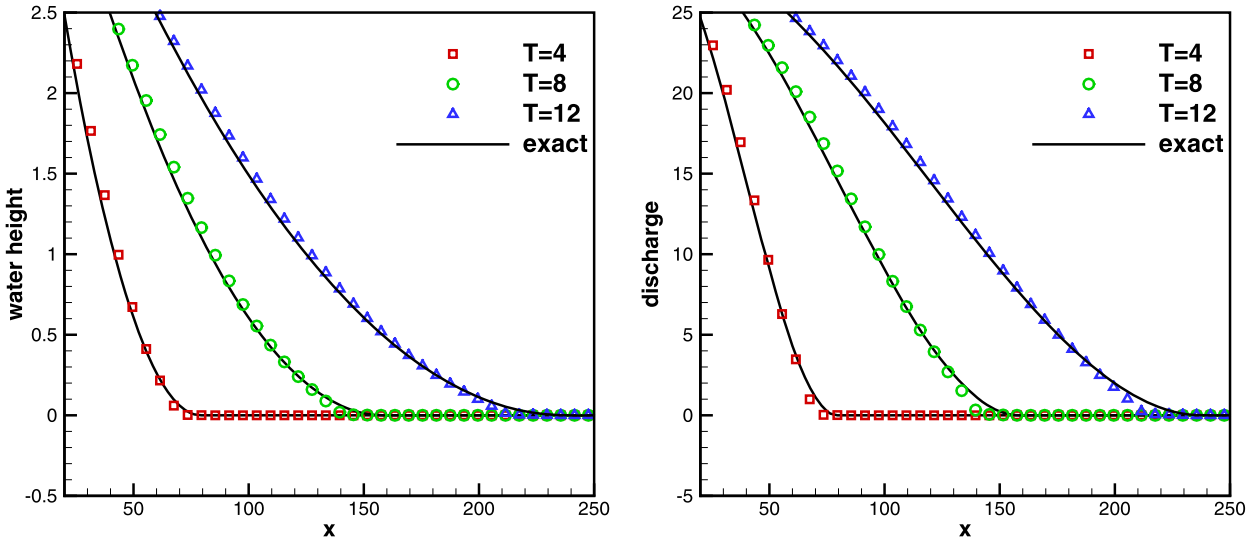

Fig. 5. The numerical and exact solutions of the Riemann problem in Section 5.4 at different time with 200 uniform cells. Zoom-in of the wet/dry front. Left: the water height h; Right: the discharge hu.

On the computational domain $[-5000, 5000]$, we consider the parabolic bottom

$$b(x) = h_0(x/a)^2, \tag{5.6}$$

with constants h_0 and a to be specified later. The analytical water surface with this set of parabolic bottom takes the form of [29]

$$h(x,t) + b(x) = h_0 - \frac{B^2}{4g}\cos(2\omega t) - \frac{B^2}{4g} - \frac{Bx}{2a}\sqrt{\frac{8h_0}{g}}\cos(\omega t), \tag{5.7}$$

where $\omega = \sqrt{2gh_0}/a$ and B is a given constant. The exact location of the wet/dry front is given by

$$x_0 = -\frac{B\omega a^2}{2gh_0}\cos(\omega t) \pm a. \tag{5.8}$$

As studied in [40], these coefficients are chosen to be $a = 3000$, $B = 5$ and $h_0 = 10$. We consider a zero discharge as the initial condition. The initial water height profile can be computed from (5.7), which contains dry regions near two

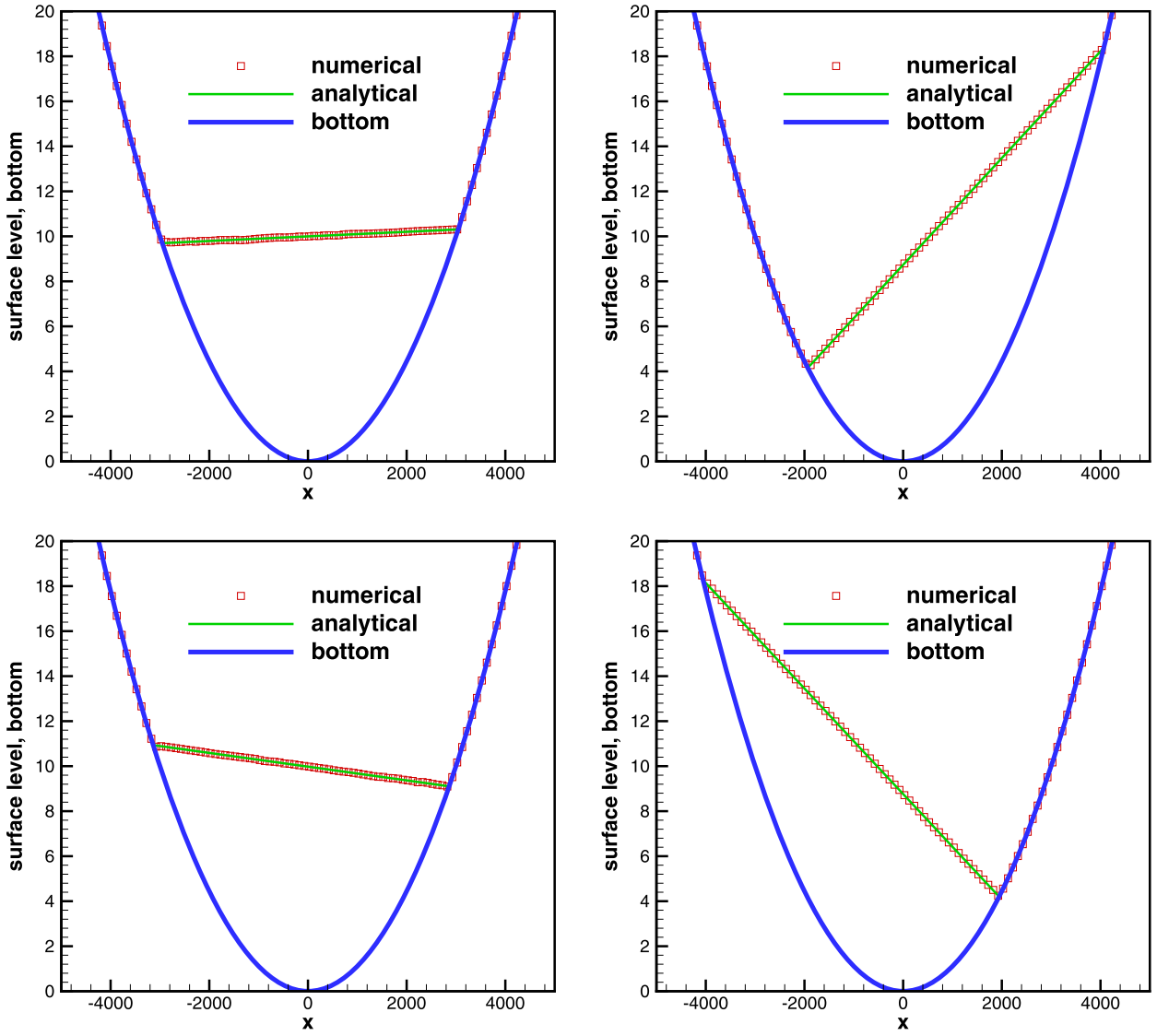

Fig. 6. The water surface level in the parabolic bowl problem at different time. Top left: $t = 1000$; Top right: $t = 2000$; Bottom left: $t = 3000$; Bottom right: $t = 4000$.

boundaries. We solve this problem until $T = 4000$ with 200 uniform cells. The numerical water surface at different times are plotted in Fig. 6. For comparison, analytical solution are also shown in these figures. We can observe a nice agreement can be observed, which confirms the positivity-preserving property of our methods for problem with a non-flat bottom topography.

5.6. The dam-breaking problem over a rectangular bump

In this traditional test case we evaluate the performance of our methods for flows involving discontinuous solutions. We consider this dam-breaking problem over a rectangular bump. It will produce a rapidly varying flow over a discontinuous bottom topography.

We consider the discontinuous bottom topography given by:

$$b(x) = \begin{cases} 8 & \text{if } |x - 750| \leqslant 1500/8, \\ 0 & \text{otherwise}, \end{cases} \tag{5.9}$$

for $x \in [0, 1500]$. The initial conditions are

$$(hu)(x, 0) = 0 \quad \text{and} \quad h(x, 0) = \begin{cases} 20 - b(x) & \text{if } x \leqslant 750, \\ 15 - b(x) & \text{otherwise}. \end{cases} \tag{5.10}$$

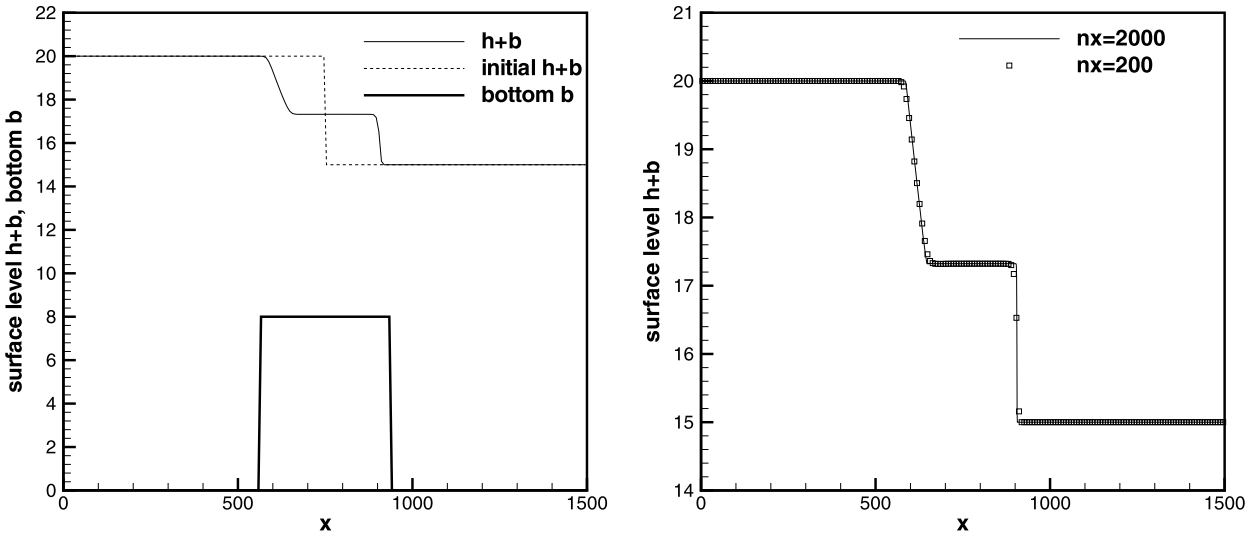

Fig. 7. The surface level $h+b$ for the dam breaking problem at time $t=15$ s. Left: the numerical solution using 200 grid cells, plotted with the initial condition and the bottom topography; Right: the numerical solution using 200 and 2000 grid cells.

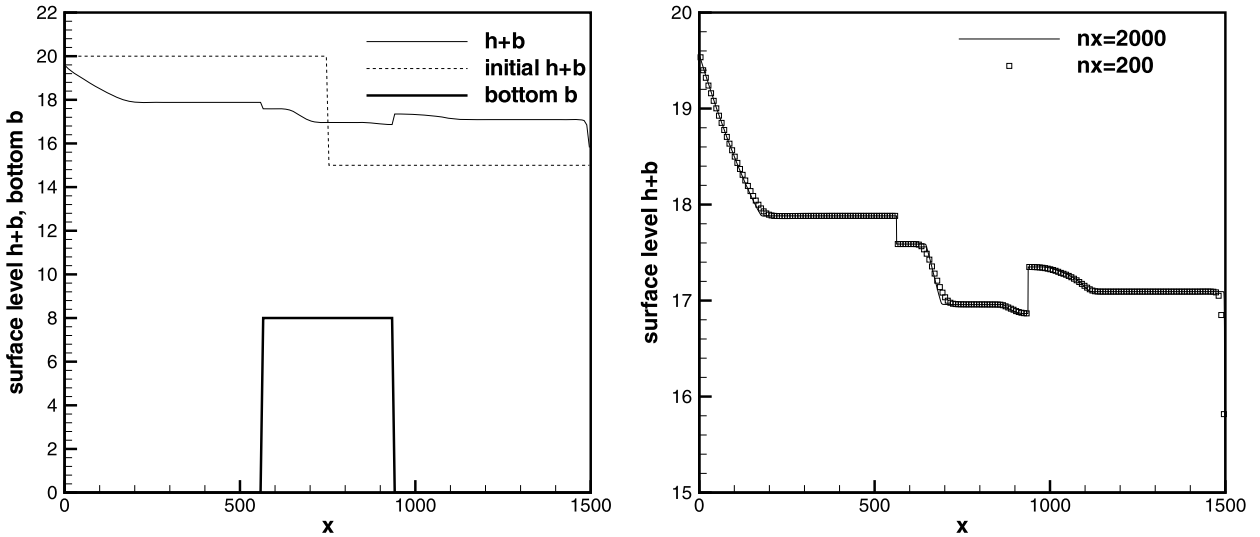

Fig. 8. The surface level $h+b$ for the dam breaking problem at time $t=60$ s. Left: the numerical solution using 200 grid cells, plotted with the initial condition and the bottom topography; Right: the numerical solution using 200 and 2000 grid cells.

Open boundary conditions are employed on both sides. Standard rarefaction and shock waves can be observed in the beginning. We show the numerical results with 200 uniform cells (and a comparison with the refined results using 2000 uniform cells) in Fig. 7 at the ending time $t=15$. These rarefaction and shock waves are the solution of the Riemann problem of the homogeneous shallow water equations. At time $T\approx 17$, the waves reach the discontinuous edges of the bottom. After that, the wave pattern becomes complicated: a part of the wave is transmitted, another part reflected, and a remaining part becomes a standing wave. Later on, this wave system keeps interacting. When the time T reaches 60, there are six waves appearing in our solution. The numerical results with 200 uniform cells (and a comparison with the refined results using 2000 uniform cells) are shown in Fig. 8 at the ending time $t=60$ s.

Note that in both plots of this example, the water height h is discontinuous at the points $x=562.5$ and $x=937.5$. Our scheme works very well for this example, giving well resolved, non-oscillatory solutions using 200 cells which agree with the converged results using 2000 cells.

6. Concluding remarks

In this paper we have constructed well-balanced DG methods of arbitrary order of accuracy for the moving water equilibrium of the shallow water equations. Special attention has been paid to the approximation of the source term, as well as

the construction of well-balanced numerical fluxes. We have also incorporated a simple positivity-preserving limiter into the well-balanced methods. Numerical examples are given to demonstrate the well-balanced property, accuracy, good capturing of the small perturbation to the steady state solutions, and non-oscillatory shock resolution of the proposed numerical methods. This well-balanced approach can be generalized to many other balance laws, and we will explore its extension to two-dimensional shallow water equations. There is no general form of the moving water equilibrium states in two dimensions, and one difficulty lies in how to identify the most interesting equilibrium state.

References

[1] E. Audusse, F. Bouchut, M.-O. Bristeau, R. Klein, B. Perthame, A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows, SIAM J. Sci. Comput. 25 (2004) 2050–2065.
[2] A. Bermudez, M.E. Vazquez, Upwind methods for hyperbolic conservation laws with source terms, Comput. Fluids 23 (1994) 1049–1071.
[3] O. Bokhove, Flooding and drying in discontinuous Galerkin finite-element discretizations of shallow-water equations. Part 1: one dimension, J. Sci. Comput. 22 (2005) 47–82.
[4] F. Bouchut, T. Morales, A subsonic-well-balanced reconstruction scheme for shallow water flows, SIAM J. Numer. Anal. 48 (2010) 1733–1758.
[5] V. Caleffi, A. Valiani, A. Bernini, Fourth-order balanced source term treatment in central WENO schemes for shallow water equations, J. Comput. Phys. 218 (2006) 228–245.
[6] M.J. Castro, J.M. Gallardo, C. Parés, High order finite volume schemes based on reconstruction of states for solving hyperbolic systems with nonconservative products. Applications to shallow-water systems, Math. Comput. 75 (2006) 1103–1134.
[7] M.J. Castro, J.A. López-García, C. Parés, High order exactly well-balanced numerical methods for shallow water systems, J. Comput. Phys. 246 (2013) 242–264.
[8] B. Cockburn, G. Karniadakis, C.-W. Shu, The development of discontinuous Galerkin methods, in: B. Cockburn, G. Karniadakis, C.-W. Shu (Eds.), Discontinuous Galerkin Methods: Theory, Computation and Applications, in: Lecture Notes in Computational Science and Engineering, vol. 11, Springer, 2000, pp. 3–50.
[9] B. Cockburn, C.-W. Shu, TVB Runge–Kutta local projection discontinuous Galerkin finite element method for conservation laws II: general framework, Math. Comput. 52 (1989) 411–435.
[10] B. Cockburn, C.-W. Shu, The Runge–Kutta discontinuous Galerkin method for conservation laws V: multidimensional systems, J. Comput. Phys. 141 (1998) 199–224.
[11] C. Dawson, J. Proft, Discontinuous and coupled continuous/discontinuous Galerkin methods for the shallow water equations, Comput. Methods Appl. Mech. Eng. 191 (2002) 4721–4746.
[12] C. Eskilsson, S.J. Sherwin, A triangular spectral/hp discontinuous Galerkin method for modelling 2D shallow water equations, Int. J. Numer. Methods Fluids 45 (2004) 605–623.
[13] F.X. Giraldo, J.S. Hesthaven, T. Warburton, Nodal high-order discontinuous Galerkin methods for the spherical shallow water equations, J. Comput. Phys. 181 (2002) 499–525.
[14] L. Gosse, A well-balanced flux-vector splitting scheme designed for hyperbolic systems of conservation laws with source terms, Comput. Math. Appl. 39 (2000) 135–159.
[15] J.M. Greenberg, A.Y. LeRoux, A well-balanced scheme for the numerical processing of source terms in hyperbolic equations, SIAM J. Numer. Anal. 33 (1996) 1–16.
[16] S. Jin, A steady-state capturing method for hyperbolic systems with geometrical source terms, Math. Model. Numer. Anal. (M2AN) 35 (2001) 631–645.
[17] S. Jin, X. Wen, An efficient method for computing hyperbolic systems with geometrical source terms having concentrations, J. Comput. Math. 22 (2004) 230–249.
[18] S. Jin, X. Wen, Two interface type numerical methods for computing hyperbolic systems with geometrical source terms having concentrations, SIAM J. Sci. Comput. 26 (2005) 2079–2101.
[19] A. Kurganov, D. Levy, Central-upwind schemes for the Saint-Venant system, Math. Model. Numer. Anal. (M2AN) 36 (2002) 397–425.
[20] R.J. LeVeque, Balancing source terms and flux gradients on high-resolution Godunov methods: the quasi-steady wave-propagation algorithm, J. Comput. Phys. 146 (1998) 346–365.
[21] S. Noelle, N. Pankratz, G. Puppo, J.R. Natvig, Well-balanced finite volume schemes of arbitrary order of accuracy for shallow water flows, J. Comput. Phys. 213 (2006) 474–499.
[22] S. Noelle, Y. Xing, C.-W. Shu, High-order well-balanced finite volume WENO schemes for shallow water equation with moving water, J. Comput. Phys. 226 (2007) 29–58.
[23] S. Noelle, Y. Xing, C.-W. Shu, High-order well-balanced schemes, in: G. Russo, G. Puppo (Eds.), Numerical Methods for Relaxation Systems and Balance Equations, in: Quaderni di Matematica, Dipartimento di Matematica, Seconda Universita di Napoli, Italy, 2009.
[24] B. Perthame, C.-W. Shu, On positivity preserving finite volume schemes for Euler equations, Numer. Math. 273 (1996) 119–130.
[25] B. Perthame, C. Simeoni, A kinetic scheme for the Saint-Venant system with a source term, Calcolo 38 (2001) 201–231.
[26] G. Russo, Central schemes for balance laws, in: Proceedings of the VIII International Conference on Nonlinear Hyperbolic Problems, Magdeburg, 2000.
[27] G. Russo, Central schemes for conservation laws with application to shallow water equations, in: S. Rionero, G. Romano (Eds.), Trends and Applications of Mathematics to Mechanics: STAMM 2002, Springer-Verlag Italia SRL, 2005, pp. 225–246.
[28] G. Russo, A. Khe, High order well balanced schemes for systems of balance laws, in: Hyperbolic Problems: Theory, Numerics and Applications, Part 2, in: Proceedings of Symposia in Applied Mathematics, vol. 67, American Mathematics Society, Providence, RI, 2009, pp. 919–928.
[29] J. Sampson, A. Easton, M. Singh, Moving boundary shallow water flow above parabolic bottom topography, ANZIAM J. 47 (2006) C373–378.
[30] C.-W. Shu, TVB uniformly high-order schemes for conservation laws, Math. Comput. 49 (1987) 105–121.
[31] C.-W. Shu, S. Osher, Efficient implementation of essentially non-oscillatory shock-capturing schemes, J. Comput. Phys. 77 (1988) 439–471.
[32] M.E. Vazquez-Cendon, Improved treatment of source terms in upwind schemes for the shallow water equations in channels with irregular geometry, J. Comput. Phys. 148 (1999) 497–526.
[33] X. Wen, A steady state capturing and preserving method for computing hyperbolic systems with geometrical source terms having concentrations, J. Comput. Phys. 219 (2006) 322–390.
[34] Y. Xing, C.-W. Shu, High order finite difference WENO schemes with the exact conservation property for the shallow water equations, J. Comput. Phys. 208 (2005) 206–227.
[35] Y. Xing, C.-W. Shu, High-order well-balanced finite difference WENO schemes for a class of hyperbolic systems with source terms, J. Sci. Comput. 27 (2006) 477–494.
[36] Y. Xing, C.-W. Shu, High order well-balanced finite volume WENO schemes and discontinuous Galerkin methods for a class of hyperbolic systems with source terms, J. Comput. Phys. 214 (2006) 567–598.

[37] Y. Xing, 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 (2006) 100–134.
[38] Y. Xing, C.-W. Shu, High-order finite volume WENO schemes for the shallow water equations with dry states, Adv. Water Resour. 34 (2011) 1026–1038.
[39] Y. Xing, C.-W. Shu, S. Noelle, On the advantage of well-balanced schemes for moving-water equilibria of the shallow water equations, J. Sci. Comput. 48 (2011) 339–349.
[40] Y. Xing, X. Zhang, C.-W. Shu, Positivity-preserving high order well-balanced discontinuous Galerkin methods for the shallow water equations, Adv. Water Resour. 33 (2010) 1476–1493.
[41] X. Zhang, C.-W. Shu, On maximum-principle-satisfying high order schemes for scalar conservation laws, J. Comput. Phys. 229 (2010) 3091–3120.