J Sci Comput (2017) 71:994–1034
DOI 10.1007/s10915-016-0329-z

A Positivity-Preserving Well-Balanced Central Discontinuous Galerkin Method for the Nonlinear Shallow Water Equations

Maojun Li[1,2] · **Philippe Guyenne**[3] · **Fengyan Li**[4] · **Liwei Xu**[1,2,5]

Received: 13 January 2016 / Revised: 14 November 2016 / Accepted: 22 November 2016 /
Published online: 8 December 2016

Abstract In this paper, we consider the development of central discontinuous Galerkin methods for solving the nonlinear shallow water equations over variable bottom topography in one and two dimensions. A reliable numerical scheme for these equations should preserve still-water stationary solutions and maintain the non-negativity of the water depth. We propose a high-order technique which exactly balances the flux gradients and source terms in the still-water stationary case by adding correction terms to the base scheme, meanwhile ensures the non-negativity of the water depth by using special approximations to the bottom together with a positivity-preserving limiter. Numerical tests are presented to illustrate the accuracy and validity of the proposed schemes.

Keywords Central discontinuous Galerkin methods · High-order accuracy · Nonlinear shallow water equations · Positivity-preserving property · Well-balanced schemes

✉ Liwei Xu
xul@cqu.edu.cn

Maojun Li
limj@cqu.edu.cn

Philippe Guyenne
guyenne@udel.edu

Fengyan Li
lif@rpi.edu

1 College of Mathematics and Statistics, Chongqing University, Chongqing 401331, People's Republic of China

2 Institute of Computing and Data Sciences, Chongqing University, Chongqing 400044, People's Republic of China

3 Department of Mathematical Sciences, University of Delaware, Newark, DE 19716-2553, USA

4 Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180-3590, USA

5 Present Address: School of Mathematical Sciences, University of Electronic Science and Technology of China, Sichuan 611731, People's Republic of China

Mathematics Subject Classification 65M60 · 76M10

1 Introduction

The nonlinear shallow water equations (NSWEs) over a variable bottom are widely adopted to model free-surface flows in rivers and coastal regions. This system is a hyperbolic conservation law with an additional source term due to the bottom topography, and thus is also called a balance law. In the one-dimensional (1D) case, it takes the form

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

where h is the total water depth, u is the vertically averaged horizontal velocity in the x-direction, g is the gravitational constant, and b is the bottom topography. The subscripts t and x denote the partial derivatives with respect to time t and the coordinate x, respectively. In the two-dimensional (2D) case, it reads

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

where $(u, v)^\top$ denotes the vertically averaged horizontal velocity vector in the x- and y-directions, and the subscript y represents the partial derivative with respect to the coordinate y.

These equations admit still-water stationary solutions for which the non-zero flux gradients are exactly balanced by the source terms. The still-water stationary solution to system (1) reads

$$h + b = \text{constant}, \qquad hu = 0, \tag{3}$$

and that to system (2) is given by

$$h + b = \text{constant}, \qquad hu = 0, \qquad hv = 0. \tag{4}$$

However, standard numerical methods usually fail to preserve such solutions, and thus much effort has been devoted to developing so-called well-balanced schemes for this purpose. In the past few decades, a number of well-balanced schemes have been proposed. LeVeque [20] developed a quasi-steady wave-propagation algorithm by introducing a Riemann problem in the center of each grid cell whose flux difference exactly cancels the source term. Xing and Shu [34–36] proposed well-balanced schemes based on WENO and discontinuous Galerkin (DG) methods. Other well-balanced finite-difference and finite-volume methods for NSWEs include [1,15,17,18,21,30,33,37,39,41] to name a few.

Another major difficulty in the simulation of shallow water waves has to do with the appearance of dry areas where no water is present. Many applications involve rapidly moving interfaces between wet and dry areas, such as wave run-up on beaches or man-made structures, dam break and flood. If no special care is taken, spurious oscillations may arise and the numerical solution may produce unphysical phenomena such as negative water depth, possibly leading to computation breakdown. Recently, a number of positivity-preserving schemes for NSWEs have been developed to maintain the non-negativity of the water depth, such as the

mesh adaptation technique [3], the slope modification technique [11], the thin layer technique [4] and the positivity-preserving limiter [38]. For related work, see also [2,5–7,12,16,19,26].

In this paper, we first apply high-order central discontinuous Galerkin (CDG) methods to NSWEs, then we present a high-order well-balanced scheme and a high-order positivity-preserving scheme based on CDG methods for NSWEs. The main ingredient to achieve the well-balanced property is to modify the source term through a special decomposition similar to that proposed in [36]. The resulting discretization has the features of being high-order accurate for general solutions while exactly preserving the still-water stationary solution. To ensure the positivity-preserving property, limiters are used by following the developments of Xing et al. [38] and Cheng et al. [8]. Xing et al. [38] introduced a simple positivity-preserving limiter in the context of DG methods for NSWEs, while Cheng et al. [8] considered positivity-preserving DG and CDG methods for ideal MHD equations. Such limiters allow for the local conservation of physical variables while maintaining high-order accuracy. Finally, to achieve the positivity-preserving and well-balanced properties simultaneously in our numerical schemes, we use special approximations to the bottom topography in the well-balanced scheme and apply the positivity-preserving limiter to the water depth.

The underlying CDG methods are a family of high-order numerical schemes defined on overlapping meshes, which were originally introduced for hyperbolic conservation laws [27,28]. Being a variant of DG methods, CDG methods evolve two copies of numerical solution without using any numerical flux at element interfaces as in Godunov schemes, and they provide new opportunities for designing accurate and stable schemes, such as for Hamilton–Jacobi equations which are not in divergence form [22], ideal MHD equations which involve a divergence-free magnetic field [23,24], and fully nonlinear weakly dispersive Green–Naghdi models which are not hyperbolic [25]. These methods prove again in the present work that they are a good candidate to solve the NSWEs. It is clear that CDG methods entail more storage space than DG methods. On the other hand, quantitative Fourier analysis and numerical tests in [28] show that CDG methods could be more accurate in some cases. The study in [31] further implies that the time step allowed for linear stability of CDG methods is typically larger than that of DG methods when the order of accuracy is higher than one. It is also worthwhile to mention some recent effort in [10] to renovate original CDG methods in order to reduce the computational cost. Note that Li et al. [25] only discussed well-balanced CDG methods for Green–Naghdi models of 1D water waves, without examining the positivity-preserving property and higher dimensions. The latter two aspects along with the well-balanced property are investigated here in the context of NSWEs, and a number of discriminating tests are carried out to assess the performance of the proposed methods in this regard.

The remainder of the paper is organized as follows. In Sect. 2, we first introduce the standard CDG methods to solve NSWEs. We then present the well-balanced CDG methods, positivity-preserving CDG methods and the combined positivity-preserving well-balanced CDG methods. Although there are similarities between the 1D and 2D cases, we find it more convenient to first describe the 1D schemes because technical calculations such as the proofs of the well-balanced and positivity-preserving properties can be more clearly detailed in this case. Along the way, we also contrast the proposed numerical methods with standard CDG strategies to highlight their differences. These results are then generalized to the 2D problem. Section 3 shows a wide range of numerical tests to illustrate the accuracy and reliability of our methods. Both 1D and 2D tests are discussed. Finally, concluding remarks are given in Sect. 4.

2 Numerical Schemes

2.1 One-Dimensional Case

In this section, we develop the numerical methods for the solution of 1D NSWEs. For ease of presentation, we rewrite (1) as

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{S}(\mathbf{U}; b), \tag{5}$$

where $\mathbf{U} = (h, hu)^\top$,

$$\mathbf{F}(\mathbf{U}) = \left(hu, hu^2 + \frac{1}{2}gh^2\right)^\top$$

is the flux, and

$$\mathbf{S}(\mathbf{U}; b) = \left(0, -ghb_x\right)^\top$$

is the source term.

Let $\{x_j\}_j$ be a partition of the computational domain $\Omega = [x_{\min}, x_{\max}]$. With $x_{j+\frac{1}{2}} = \frac{1}{2}(x_j + x_{j+1})$, $I_j = (x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}})$ and $I_{j+\frac{1}{2}} = (x_j, x_{j+1})$, we define two discrete function spaces associated with the overlapping meshes, the primal mesh $\{I_j\}_j$ and the dual mesh $\{I_{j+\frac{1}{2}}\}_j$, to approximate $\mathbf{U}$,

$$\mathcal{V}_h^C = \mathcal{V}_h^{C,k} = \left\{\mathbf{v} : \mathbf{v}|_{I_j} \in \left[P^k\left(I_j\right)\right]^2, \forall j\right\},$$
$$\mathcal{V}_h^D = \mathcal{V}_h^{D,k} = \left\{\mathbf{v} : \mathbf{v}|_{I_{j+\frac{1}{2}}} \in \left[P^k\left(I_{j+\frac{1}{2}}\right)\right]^2, \forall j\right\},$$

where $P^k(I)$ denotes the space of polynomials in I with degree of at most k, and $[P^k(I)]^2 = \{\mathbf{v} = (v_1, v_2)^\top : v_i \in P^k(I), i = 1, 2\}$ is its vector version. The mesh is assumed to be uniform, with Δx as the mesh size.

For simplicity, we present the numerical schemes in the case of the forward Euler method as time discretization. Higher-order time discretizations will be discussed afterwards. The proposed methods evolve two copies of numerical solution, which are assumed to be available at $t = t_n$, denoted by $\mathbf{U}_h^{n,\star} = (h_h^{n,\star}, (hu)_h^{n,\star})^\top \in \mathcal{V}_h^\star$, and we want to find the solutions at $t = t_{n+1} = t_n + \Delta t_n$. Hereinafter, the symbol $\star$ denotes the copy C or D. For convenience, we only present the procedure to update $\mathbf{U}_h^{n+1,C}$ as the one for $\mathbf{U}_h^{n+1,D}$ is similar. Regarding the bottom topography function b, we project it into $P^k(I_j)$ on I_j (resp. $P^k(I_{j+\frac{1}{2}})$ on $I_{j+\frac{1}{2}}$) in the L^2 sense, and obtain an approximation b_h^C (resp. b_h^D) throughout the domain Ω. A different strategy is used in Sect. 2.1.4 to approximate b.

2.1.1 Standard CDG Methods

To get $\mathbf{U}_h^{n+1,C} = (h_h^{n+1,C}, (hu)_h^{n+1,C})^\top$, we apply to (5) the CDG methods of Liu et al. [27] for space discretization and the forward Euler method for time discretization. That is, we look for $\mathbf{U}_h^{n+1,C} \in \mathcal{V}_h^{C,k}$ such that for any $\mathbf{V} \in \mathcal{V}_h^{C,k}|_{I_j}$ with any j,

$$\begin{aligned}\int_{I_j}\mathbf{U}_h^{n+1,C}\cdot\mathbf{V}dx=&\int_{I_j}\left(\theta_n\mathbf{U}_h^{n,D}+(1-\theta_n)\,\mathbf{U}_h^{n,C}\right)\cdot\mathbf{V}dx\\&+\Delta t_n\int_{I_j}\mathbf{F}\left(\mathbf{U}_h^{n,D}\right)\cdot\mathbf{V}_x dx\\&-\Delta t_n\left[\mathbf{F}\left(\mathbf{U}_h^{n,D}\left(x_{j+\frac{1}{2}}\right)\right)\cdot\mathbf{V}\left(x_{j+\frac{1}{2}}^-\right)\right.\\&\left.-\mathbf{F}\left(\mathbf{U}_h^{n,D}\left(x_{j-\frac{1}{2}}\right)\right)\cdot\mathbf{V}\left(x_{j-\frac{1}{2}}^+\right)\right]\\&+\Delta t_n\int_{I_j}\mathbf{S}\left(\mathbf{U}_h^{n,D};b_h^D\right)\cdot\mathbf{V}dx.\end{aligned}\tag{6}$$

Here $\theta_n=\Delta t_n/\tau_n\in[0,1]$ with τ_n being the maximal time step allowed by the CFL restriction at t_n [27], and $w(x^{\pm})=\lim_{\epsilon\to0+}w(x\pm\epsilon)$ for any scalar or vector function w. In general, the numerical scheme (6) does not maintain the still-water stationary solution and may produce solutions with spurious oscillations when they are in or close to a stationary state, such as small perturbations of the still-water surface.

2.1.2 Well-Balanced CDG Methods

To achieve the well-balanced property, we add terms to the right-hand side of (6) in the following way

$$\begin{aligned}\int_{I_j}\mathbf{U}_h^{n+1,C}\cdot\mathbf{V}dx=&\int_{I_j}\left(\theta_n\mathbf{U}_h^{n,D}+(1-\theta_n)\,\mathbf{U}_h^{n,C}\right)\cdot\mathbf{V}dx+\widetilde{\mathbf{U}}\left(b_h^C,b_h^D,\mathbf{V}\right)\\&+\Delta t_n\int_{I_j}\mathbf{F}\left(\mathbf{U}_h^{n,D}\right)\cdot\mathbf{V}_x dx\\&-\Delta t_n\left[\mathbf{F}\left(\mathbf{U}_h^{n,D}\left(x_{j+\frac{1}{2}}\right)\right)\cdot\mathbf{V}\left(x_{j+\frac{1}{2}}^-\right)\right.\\&\left.-\mathbf{F}\left(\mathbf{U}_h^{n,D}\left(x_{j-\frac{1}{2}}\right)\right)\cdot\mathbf{V}\left(x_{j-\frac{1}{2}}^+\right)\right]\\&+\Delta t_n\left[\int_{I_j}\mathbf{S}\left(\mathbf{U}_h^{n,D};b_h^D\right)\cdot\mathbf{V}dx+\widetilde{\mathbf{S}}\left(\mathbf{U}_h^{n,D};b_h^D\right)\cdot\mathbf{V}\left(x_j\right)\right],\end{aligned}\tag{7}$$

where the newly added terms are given by

$$\widetilde{\mathbf{U}}\left(b_h^C,b_h^D,\mathbf{V}\right)=\theta_n\int_{I_j}\left(b_h^D-b_h^C,0\right)^{\top}\cdot\mathbf{V}dx,\tag{8}$$

$$\begin{aligned}\widetilde{\mathbf{S}}\left(U_h^{n,D};b_h^D\right)=&\left(0,\frac{g}{2}\left(\left(b_h^D\left(x_j^+\right)\right)^2-\left(b_h^D\left(x_j^-\right)\right)^2\right)\right.\\&\left.-\gamma_j^{n,D}g\left(b_h^D\left(x_j^+\right)-b_h^D\left(x_j^-\right)\right)\right)^{\top},\end{aligned}\tag{9}$$

and

$$\gamma_j^{n,D}=\frac{1}{2}\left(h_h^{n,D}\left(x_{j+\frac{1}{2}}\right)+b_h^D\left(x_{j+\frac{1}{2}}\right)+h_h^{n,D}\left(x_{j-\frac{1}{2}}\right)+b_h^D\left(x_{j-\frac{1}{2}}\right)\right).$$

To construct a well-balanced scheme, the crucial property of $\gamma_j^{n,D}$ is that in the case of the still-water stationary solution, $\gamma_j^{n,D}$ should be a constant with respect to both j and n. In

particular, it should be equal to the constant water surface in (3). Alternatively, we can take

$$\gamma_j^{n,D} = \frac{1}{|I_j|}\int_{I_j}\left(h_h^{n,D} + b_h^D\right)dx.$$

Note that the modification by adding the terms $\widetilde{\mathbf{U}}$ and $\widetilde{\mathbf{S}}$ is of order $O(\Delta x^{k+1})$ when the bottom topography function b is sufficiently smooth, and hence it will not affect the (formal) spatial accuracy of standard CDG schemes in (6).

Inclusion of the term $\widetilde{\mathbf{S}}$ is motivated by the following decomposition

$$\begin{aligned}\mathbf{S}\left(\mathbf{U}_h^{n,D}; b_h^D\right) = &\left(0, \frac{g}{2}\left(b_h^D\right)^2 - \gamma_j^{n,D} g b_h^D\right)_x^\top \\ &- \left(0, g\left(h_h^{n,D} + b_h^D - \gamma_j^{n,D}\right)\left(b_h^D\right)_x\right)^\top. \end{aligned} \tag{10}$$

With this, the terms related to the source in (7) become

$$\begin{aligned}&\int_{I_j}\mathbf{S}\left(\mathbf{U}_h^{n,D}; b_h^D\right)\cdot\mathbf{V}dx + \widetilde{\mathbf{S}}\left(\mathbf{U}_h^{n,D}; b_h^D\right)\cdot\mathbf{V}\left(x_j\right) \\ &= \left(0, \frac{g}{2}\left(b_h^D\left(x_{j+\frac{1}{2}}\right)\right)^2 - \gamma_j^{n,D} g b_h^D\left(x_{j+\frac{1}{2}}\right)\right)^\top\cdot\mathbf{V}\left(x_{j+\frac{1}{2}}^-\right) \\ &\quad - \left(0, \frac{g}{2}\left(b_h^D\left(x_{j-\frac{1}{2}}\right)\right)^2 - \gamma_j^{n,D} g b_h^D\left(x_{j-\frac{1}{2}}\right)\right)^\top\cdot\mathbf{V}\left(x_{j-\frac{1}{2}}^+\right) \\ &\quad - \int_{I_j}\left(0, \frac{g}{2}\left(b_h^D\right)^2 - \gamma_j^{n,D} g b_h^D\right)^\top\cdot\mathbf{V}_x dx \\ &\quad - \int_{I_j}\left(0, g\left(h_h^{n,D} + b_h^D - \gamma_j^{n,D}\right)\left(b_h^D\right)_x\right)^\top\cdot\mathbf{V}dx, \end{aligned} \tag{11}$$

which is obtained by applying integration by parts to the term containing

$$\left(0, \frac{g}{2}\left(b_h^D\right)^2 - \gamma_j^{n,D} g b_h^D\right)_x^\top,$$

in (10). Given that b_h^D is discontinuous at x_j, integration by parts is indeed applied to two half-elements, $[x_{j-\frac{1}{2}}, x_j]$ and $[x_j, x_{j+\frac{1}{2}}]$ separately. The newly added term $\widetilde{\mathbf{S}}$ contributes to the cancellation of some terms. Equation (11) is used to update $\mathbf{U}_h^{n+1,C}$ in (7). We have the following theorem on the well-balanced property of the proposed scheme.

Theorem 1 *The numerical scheme defined in* (7)*, and its counterpart for* $\mathbf{U}_h^{n+1,D}$*, to solve the 1D NSWEs is well-balanced in the sense that it preserves the still-water stationary solution* (3)*.*

Proof We start with the still-water solution $h + b = C_0, hu = 0$ at $t = 0$, where C_0 is a constant. In the initialization step, it is easy to ensure

$$\mathbf{U}_h^{n,C} + \left(b_h^C, 0\right)^\top = \mathbf{U}_h^{n,D} + \left(b_h^D, 0\right)^\top = (C_0, 0)^\top, \tag{12}$$

for $n = 0$. By induction, assuming condition (12) is true for $n \geq 0$, we want to establish that the numerical solution computed from (7) and its counterpart for $\mathbf{U}_h^{n+1,D}$ satisfy

$$\mathbf{U}_h^{n+1,C} + \left(b_h^C, 0\right)^\top = \mathbf{U}_h^{n+1,D} + \left(b_h^D, 0\right)^\top = (C_0, 0)^\top.$$

By virtue of (12), the first two terms on the right-hand side of (7) become

$$\begin{aligned}&\int_{I_j}\left(\theta_n\mathbf{U}_h^{n,D}+(1-\theta_n)\mathbf{U}_h^{n,C}\right)\cdot\mathbf{V}dx+\widetilde{\mathbf{U}}\left(b_h^C,b_h^D,\mathbf{V}\right)\\&=\int_{I_j}\left(C_0-b_h^C,0\right)^\top\cdot\mathbf{V}dx,\end{aligned}\tag{13}$$

and the flux term is

$$\mathbf{F}\left(\mathbf{U}_h^{n,D}\right)=\left(0,\frac{g}{2}\left(C_0-b_h^D\right)^2\right)^\top=\left(0,\frac{g}{2}C_0^2+\frac{g}{2}\left(b_h^D\right)^2-gC_0b_h^D\right)^\top.\tag{14}$$

With this, all the terms in (7) related to the flux take the form

$$\begin{aligned}&\int_{I_j}\mathbf{F}\left(\mathbf{U}_h^{n,D}\right)\cdot\mathbf{V}_x dx\\&\quad-\mathbf{F}\left(\mathbf{U}_h^{n,D}\left(x_{j+\frac12}\right)\right)\cdot\mathbf{V}\left(x_{j+\frac12}^-\right)+\mathbf{F}\left(\mathbf{U}_h^{n,D}\left(x_{j-\frac12}\right)\right)\cdot\mathbf{V}\left(x_{j-\frac12}^+\right)\\&=\int_{I_j}\left(0,\frac{g}{2}\left(b_h^D\right)^2-gC_0b_h^D\right)^\top\cdot\mathbf{V}_x dx\\&\quad-\left(0,\frac{g}{2}\left(b_h^D\left(x_{j+\frac12}\right)\right)^2-gC_0b_h^D\left(x_{j+\frac12}\right)\right)^\top\cdot\mathbf{V}\left(x_{j+\frac12}^-\right)\\&\quad+\left(0,\frac{g}{2}\left(b_h^D\left(x_{j-\frac12}\right)\right)^2-gC_0b_h^D\left(x_{j-\frac12}\right)\right)^\top\cdot\mathbf{V}\left(x_{j-\frac12}^+\right).\end{aligned}\tag{15}$$

Now we can combine (7), (11), (13), (15) and use the fact that $\gamma_j^{n,D}=C_0\ (\forall j,n)$ to get

$$\int_{I_j}\mathbf{U}_h^{n+1,C}\cdot\mathbf{V}dx=\int_{I_j}\left(C_0-b_h^C,0\right)^\top\cdot\mathbf{V}dx.\tag{16}$$

By further taking $\mathbf{V}=\mathbf{U}_h^{n+1,C}-\left(C_0-b_h^C,0\right)^\top\in\mathcal{V}_h^{C,k}$, we obtain $\mathbf{U}_h^{n+1,C}=\left(C_0-b_h^C,0\right)^\top$. Similarly, we can establish $\mathbf{U}_h^{n+1,D}=\left(C_0-b_h^D,0\right)^\top$. This completes the proof. □

2.1.3 Positivity-Preserving CDG Methods

In this section, we propose a positivity-preserving CDG method for the 1D NSWEs, with which the cell averages of the computed water depth are non-negative at any discrete time $t=t_n$.

Recall that for the exact solution $\mathbf{U}=(h,hu)^\top$, when h is zero at some (x,t), hu will vanish simultaneously. This however may not hold for numerical solutions. To prepare for our positivity-preserving strategy at $t=t_n$, we pre-process the numerical solutions $\mathbf{U}_h^{n,C}=(h_h^{n,C},(hu)_h^{n,C})^\top$ and $\mathbf{U}_h^{n,D}=(h_h^{n,D},(hu)_h^{n,D})^\top$, and modify them into $\widetilde{\mathbf{U}}_h^{n,C}=(h_h^{n,C},\widetilde{(hu)}_h^{n,C})^\top$ and $\widetilde{\mathbf{U}}_h^{n,D}=(h_h^{n,D},\widetilde{(hu)}_h^{n,D})^\top$, where $\widetilde{(hu)}_h^{n,C}$ satisfies

$$\begin{aligned}&\min_{\widetilde{(hu)}_h^{n,C}|_{I_j}\in P^k(I_j)}\ \|\widetilde{(hu)}_h^{n,C}-(hu)_h^{n,C}\|_{L^2(I_j)},\\&\text{subject to}\qquad\widetilde{(hu)}_h^{n,C}\left(x_j\right)=0,\end{aligned}\tag{17}$$

if $h_h^{n,C}(x_j) = 0$, and $\widetilde{(hu)}_h^{n,C}|_{I_j} = (hu)_h^{n,C}|_{I_j}$ if $h_h^{n,C}(x_j) \neq 0, \forall j$, while $\widetilde{(hu)}_h^{n,D}$ satisfies

$$\min_{\widetilde{(hu)}_h^{n,D}|_{I_{j-\frac{1}{2}}} \in P^k\left(I_{j-\frac{1}{2}}\right)} \| \widetilde{(hu)}_h^{n,D} - (hu)_h^{n,D} \|_{L^2\left(I_{j-\frac{1}{2}}\right)},$$
$$\text{subject to} \qquad \widetilde{(hu)}_h^{n,D}\left(x_{j-\frac{1}{2}}\right) = 0, \tag{18}$$

if $h_h^{n,D}(x_{j-\frac{1}{2}}) = 0$, and $\widetilde{(hu)}_h^{n,D}|_{I_{j-\frac{1}{2}}} = (hu)_h^{n,D}|_{I_{j-\frac{1}{2}}}$ if $h_h^{n,D}(x_{j-\frac{1}{2}}) \neq 0, \forall j$.

The proof of the upcoming Theorem 2 requires that h and hu be consistent at point x_j on the primal mesh (resp. $x_{j-\frac{1}{2}}$ on the dual mesh), namely if $h = 0$, then $hu = 0$. Therefore, in Eq. (17) (resp. Eq. (18)), we only need to check point x_j for the solution on the primal mesh (resp. $x_{j-\frac{1}{2}}$ for the solution on the dual mesh). The consistency of h and hu at other quadrature points affects neither the proof nor the positivity-preserving property.

After the pre-processing step at $t = t_n$, the solutions at the next time $t = t_{n+1}$ are updated as follows using CDG methods,

$$\begin{aligned}\int_{I_j} \mathbf{U}_h^{n+1,C} \cdot \mathbf{V} dx = &\int_{I_j} \left(\theta_n \widetilde{\mathbf{U}}_h^{n,D} + (1-\theta_n)\widetilde{\mathbf{U}}_h^{n,C}\right) \cdot \mathbf{V} dx \\ &+\Delta t_n \int_{I_j} \mathbf{F}\left(\widetilde{\mathbf{U}}_h^{n,D}\right) \cdot \mathbf{V}_x dx \\ &-\Delta t_n \left[\mathbf{F}\left(\widetilde{\mathbf{U}}_h^{n,D}\left(x_{j+\frac{1}{2}}\right)\right) \cdot \mathbf{V}\left(x_{j+\frac{1}{2}}^-\right)\right. \\ &\left. - \mathbf{F}\left(\widetilde{\mathbf{U}}_h^{n,D}\left(x_{j-\frac{1}{2}}\right)\right) \cdot \mathbf{V}\left(x_{j-\frac{1}{2}}^+\right)\right] \\ &+\Delta t_n \int_{I_j} \mathbf{S}\left(\widetilde{\mathbf{U}}_h^{n,D}; b_h^D\right) \cdot \mathbf{V} dx. \end{aligned} \tag{19}$$

Again we only present the numerical scheme with the forward Euler method for temporal integration, with higher-order time discretizations to be discussed later (see Sect. 2.3).

Next, we consider the equation satisfied by the cell average of the CDG solution h_h^C, and take $\mathbf{V} = (\frac{1}{\Delta x}, 0)^\top$ in (19),

$$\begin{aligned}\bar{h}_j^{n+1,C} = &(1-\theta_n)\bar{h}_j^{n,C} + \frac{\theta_n}{\Delta x}\int_{I_j} h_h^{n,D} dx \\ &-\frac{\Delta t_n}{\Delta x}\left[\widetilde{(hu)}_h^{n,D}\left(x_{j+\frac{1}{2}}\right) - \widetilde{(hu)}_h^{n,D}\left(x_{j-\frac{1}{2}}\right)\right]. \end{aligned} \tag{20}$$

Here $\bar{h}_j^{n,C}$ is the cell average of h_h^C over I_j at time t_n.

In order to provide a sufficient condition that ensures $\bar{h}_j^{n+1,C} \geq 0, \forall j$, let $\widehat{L}_j^{1,x} = \{\widehat{x}_j^{1,\beta}, \beta = 1, 2, \ldots, \widehat{N}\}$ and $\widehat{L}_j^{2,x} = \{\widehat{x}_j^{2,\beta}, \beta = 1, 2, \ldots, \widehat{N}\}$ be the Legendre Gauss–Lobatto quadrature points on $[x_{j-\frac{1}{2}}, x_j]$ and $[x_j, x_{j+\frac{1}{2}}]$ respectively. The corresponding quadrature rule is exact for the integration of polynomials of degree up to $2\widehat{N}-3$. We choose $\widehat{N}$ such that $2\widehat{N}-3 \geq k$. Let $\{\widehat{\omega}_\beta, \beta = 1, 2, \ldots, \widehat{N}\}$ be the corresponding quadrature weights for the reference interval $[-\frac{1}{2}, \frac{1}{2}]$. Note that $\sum_{\beta=1}^{\widehat{N}} \widehat{\omega}_\beta = 1$ with $\widehat{\omega}_1 = \widehat{\omega}_{\widehat{N}}$, and $x_{j-\frac{1}{2}} = \widehat{x}_j^{1,1} = \widehat{x}_{j-1}^{2,\widehat{N}}$, $x_j = \widehat{x}_j^{1,\widehat{N}} = \widehat{x}_j^{2,1}, \forall j$.

For any given j, we define

$$u_h^{n,D}\left(x_{j-\frac{1}{2}}\right)=\begin{cases}\dfrac{\widetilde{(hu)}_h^{n,D}\left(x_{j-\frac{1}{2}}\right)}{h_h^{n,D}\left(x_{j-\frac{1}{2}}\right)} & \text{if } h_h^{n,D}\left(x_{j-\frac{1}{2}}\right)>0,\\ 0 & \text{if } h_h^{n,D}\left(x_{j-\frac{1}{2}}\right)=0.\end{cases}\tag{21}$$

Theorem 2 *For any given $n\geq 0$, we assume $\bar{h}_j^{n,C}\geq 0$ and $\bar{h}_j^{n,D}\geq 0,\forall j$. Consider the numerical scheme in (20) and its counterpart for $\bar{h}_j^{n+1,D}$, if $h_h^C(x,t_n)\geq 0, h_h^D(x,t_n)\geq 0,\forall x\in\widehat{L}_j^{l,x},\forall j$ and $l=\{1,2\}$, then $\bar{h}_j^{n+1,C}\geq 0$ and $\bar{h}_{j-\frac{1}{2}}^{n+1,D}\geq 0,\forall j$, under the CFL condition*

$$\lambda_x a_x\leq\frac{1}{2}\theta_n\widehat{\omega}_1,\tag{22}$$

where $\lambda_x=\Delta t_n/\Delta x$ and $a_x=\max(\|u_h^{n,C}\|_\infty,\|u_h^{n,D}\|_\infty)$.

Proof Using the Legendre Gauss–Lobatto quadrature rule, one has

$$\frac{1}{\Delta x}\int_{I_j}h_h^{n,D}dx=\frac{1}{2}\left(\sum_{\beta=1}^{\widehat{N}}\widehat{\omega}_\beta h_{1,\beta}^{n,D}+\sum_{\beta=1}^{\widehat{N}}\widehat{\omega}_\beta h_{2,\beta}^{n,D}\right),\tag{23}$$

where $h_{l,\beta}^{n,D}=h_h^{n,D}(\widehat{x}_j^{l,\beta})$. Now, substituting (23) into (20) yields

$$\begin{aligned}\bar{h}_j^{n+1,C}&=(1-\theta_n)\,\bar{h}_j^{n,C}+\frac{\theta_n}{2}\left(\sum_{\beta=1}^{\widehat{N}}\widehat{\omega}_\beta h_{1,\beta}^{n,D}+\sum_{\beta=1}^{\widehat{N}}\widehat{\omega}_\beta h_{2,\beta}^{n,D}\right)\\&\quad-\lambda_x\left[\widetilde{(hu)}_h^{n,D}\left(x_{j+\frac{1}{2}}\right)-\widetilde{(hu)}_h^{n,D}\left(x_{j-\frac{1}{2}}\right)\right],\\&=(1-\theta_n)\,\bar{h}_j^{n,C}+\frac{\theta_n}{2}\left(\sum_{\beta=2}^{\widehat{N}}\widehat{\omega}_\beta h_{1,\beta}^{n,D}+\sum_{\beta=1}^{\widehat{N}-1}\widehat{\omega}_\beta h_{2,\beta}^{n,D}\right)+\mathcal{H}_j,\end{aligned}\tag{24}$$

with

$$\begin{aligned}\mathcal{H}_j&=\left(\frac{1}{2}\theta_n\widehat{\omega}_1 h_{j-\frac{1}{2}}^{n,D}+\lambda_x\widetilde{(hu)}_{j-\frac{1}{2}}^{n,D}\right)+\left(\frac{1}{2}\theta_n\widehat{\omega}_{\widehat{N}}h_{j+\frac{1}{2}}^{n,D}-\lambda_x\widetilde{(hu)}_{j+\frac{1}{2}}^{n,D}\right),\\&=\begin{cases}h_{j-\frac{1}{2}}^{n,D}\left(\frac{1}{2}\theta_n\widehat{\omega}_1+\lambda_x u_{j-\frac{1}{2}}^{n,D}\right)+h_{j+\frac{1}{2}}^{n,D}\left(\frac{1}{2}\theta_n\widehat{\omega}_1-\lambda_x u_{j+\frac{1}{2}}^{n,D}\right) & \text{if } h_{j\pm\frac{1}{2}}^{n,D}>0,\\ h_{j+\frac{1}{2}}^{n,D}\left(\frac{1}{2}\theta_n\widehat{\omega}_1-\lambda_x u_{j+\frac{1}{2}}^{n,D}\right) & \text{if } h_{j-\frac{1}{2}}^{n,D}=0,\ h_{j+\frac{1}{2}}^{n,D}>0,\\ h_{j-\frac{1}{2}}^{n,D}\left(\frac{1}{2}\theta_n\widehat{\omega}_1+\lambda_x u_{j-\frac{1}{2}}^{n,D}\right) & \text{if } h_{j-\frac{1}{2}}^{n,D}>0,\ h_{j+\frac{1}{2}}^{n,D}=0,\\ 0 & \text{if } h_{j\pm\frac{1}{2}}^{n,D}=0.\end{cases}\end{aligned}$$

Here we have used $\widehat{\omega}_1=\widehat{\omega}_{\widehat{N}}$, $h_{1,1}^{n,D}=h_{j-\frac{1}{2}}^{n,D}$, $h_{2,\widehat{N}}^{n,D}=h_{j+\frac{1}{2}}^{n,D}$ and $\widetilde{(hu)}_{j\pm\frac{1}{2}}^{n,D}=\widetilde{(hu)}_h^{n,D}(x_{j\pm\frac{1}{2}})$.

A few observations can be made. Firstly, the CFL condition (22) implies $\mathcal{H}_j \geq 0$. Secondly, $\bar{h}_j^{n+1,C}$ is a linear combination of $\bar{h}_j^{n,C}$, $h_{1,\beta}^{n,D}$ $(\beta = 2, 3, \ldots, \widehat{N})$, $h_{2,\beta}^{n,D}$ $(\beta = 1, 2, \ldots, \widehat{N}-1)$ and $\mathcal{H}_j$, which are all non-negative and the corresponding coefficients are also non-negative, therefore $\bar{h}_j^{n+1,C} \geq 0, \forall j$. Similarly, one can show $\bar{h}_{j-\frac{1}{2}}^{n+1,D} \geq 0, \forall j$. □

Next, under the assumption $\bar{h}_j^{n,C} \geq 0$ and $\bar{h}_j^{n,D} \geq 0, \forall j$, we will give a positivity-preserving limiter which modifies the CDG solution polynomials h_h^C and h_h^D at time t_n into $\widetilde{h}_h^C$ and $\widetilde{h}_h^D$ such that they satisfy the sufficient condition in Theorem 2, while maintaining accuracy and local conservation. In fact, this limiter is almost the same as the one in [8,38] as long as it is applied to h_h^C and h_h^D separately.

Let K denote an element from the primal mesh or the dual mesh. Let $\widehat{L}_K$ represent the set of relevant quadrature points in K, namely $\widehat{L}_K = \widehat{L}_j^{1,x} \cup \widehat{L}_j^{2,x}$ when $K = I_j$, and $\widehat{L}_K = \widehat{L}_{j-1}^{2,x} \cup \widehat{L}_j^{1,x}$ if $K = I_{j-\frac{1}{2}}$. Following [8,38], the positivity-preserving limiter is given as follows: on each mesh element K, we modify the water depth $h_h^{n,\star}$ into $\widetilde{h}_h^{n,\star} = \alpha_K(h_h^{n,\star} - \bar{h}_h^{n,\star}) + \bar{h}_h^{n,\star}$ with

$$\alpha_K = \min_{x \in \widehat{L}_K} \left\{ 1, \left| \frac{\bar{h}_h^{n,\star}}{\bar{h}_h^{n,\star} - h_h^{n,\star}(x)} \right| \right\}.$$

It can be seen that the total number of quadrature points involved in the positivity-preserving limiter for CDG methods is twice (or $2d$ with d being the spatial dimension) that of standard DG methods.

Remark 1 In actual implementation, at time $t = t_n$, the positivity-preserving limiter is first applied to obtain $\widetilde{h}_h^{n,\star}$ and then, using this newly limited water height, we further pre-process the momentum and get $\widetilde{(hu)}_h^{n,\star}$ as described at the beginning of this subsection.

2.1.4 Positivity-Preserving Well-Balanced CDG Methods

In this section, we will propose a CDG method which is simultaneously positivity-preserving and well-balanced for the 1D NSWEs. This is achieved by proposing a positivity-preserving limiter for the well-balanced CDG method in (7) such that the cell averages of the computed water depth on each mesh are non-negative at any discrete time t_n.

To this end, we consider the equation satisfied by the cell average of the well-balanced CDG solution h_h^C, and by taking $\mathbf{V} = (\frac{1}{\Delta x}, 0)^\top$ in (7) and replacing $(hu)_h^{n,D}$ with $\widetilde{(hu)}_h^{n,D}$ as defined in Sect. 2.1.3, we obtain

$$\begin{aligned} \bar{h}_j^{n+1,C} = {} & (1-\theta_n)\,\bar{h}_j^{n,C} + \frac{\theta_n}{\Delta x}\int_{I_j} h_h^{n,D}\,dx \\ & -\lambda_x\left[\widetilde{(hu)}_h^{n,D}\left(x_{j+\frac{1}{2}}\right) - \widetilde{(hu)}_h^{n,D}\left(x_{j-\frac{1}{2}}\right)\right] \\ & +\theta_n\left(\frac{1}{\Delta x}\int_{I_j} b_h^D\,dx - \bar{b}_j^C\right). \end{aligned} \tag{25}$$

Theorem 3 *For any given $n \geq 0$, we assume $\bar{h}_j^{n,C} \geq 0$ and $\bar{h}_j^{n,D} \geq 0, \forall j$. Consider the numerical scheme in (25) and its counterpart for $\bar{h}_j^{n+1,D}$, if*

$$\bar{b}_j^C = \frac{1}{\Delta x}\int_{I_j} b_h^D dx, \qquad \bar{b}_{j+\frac{1}{2}}^D = \frac{1}{\Delta x}\int_{I_{j+\frac{1}{2}}} b_h^C dx, \quad \forall j, \tag{26}$$

and $h_h^C(x,t_n) \geq 0, h_h^D(x,t_n) \geq 0, \forall x \in \widehat{L}_j^{l,x}, \forall j$ with $l = \{1,2\}$, then $\bar{h}_j^{n+1,C} \geq 0$ and $\bar{h}_{j-\frac{1}{2}}^{n+1,D} \geq 0, \forall j$, under the CFL condition (22).

Proof Using the relations in (26), Eq. (25) becomes

$$\begin{aligned}\bar{h}_j^{n+1,C} = (1-\theta_n)\,\bar{h}_j^{n,C} + \frac{\theta_n}{\Delta x}\int_{I_j} h_h^{n,D} dx \\ -\lambda_x\left[\widetilde{(hu)}_h^{n,D}\left(x_{j+\frac{1}{2}}\right) - \widetilde{(hu)}_h^{n,D}\left(x_{j-\frac{1}{2}}\right)\right],\end{aligned} \tag{27}$$

which is identical to (20), thus we get $\bar{h}_j^{n+1,C} \geq 0$ and $\bar{h}_{j-\frac{1}{2}}^{n+1,D} \geq 0, \forall j$ following the proof of Theorem 2. □

To enforce the sufficient condition in Theorem 3, we first obtain new approximations to the bottom topography $b(x)$, still denoted as b_h^C and b_h^D, by solving the following constrained minimization problem

$$\begin{aligned}\min_{b_h^C \in V_h^C,\ b_h^D \in V_h^D} \quad & \|b_h^C - b\|_{L^2} + \|b_h^D - b\|_{L^2}, \\ \text{subject to} \quad & \bar{b}_j^C = \tfrac{1}{\Delta x}\textstyle\int_{I_j} b_h^D dx, \quad \bar{b}_{j+\frac{1}{2}}^D = \tfrac{1}{\Delta x}\textstyle\int_{I_{j+\frac{1}{2}}} b_h^C dx,\end{aligned} \tag{28}$$

where $V_h^C = \{v : v|_{I_j} \in P^k(I_j), \forall j\}$ and $V_h^D = \{v : v|_{I_{j+\frac{1}{2}}} \in P^k(I_{j+\frac{1}{2}}), \forall j\}$. We then employ a positivity-preserving limiter which is the same as the one described in Sect. 2.1.3. The constrained minimization problem (28) is solved using the Lagrange multiplier method. Even though there is no rigorous proof, our numerical tests (see Sect. 3.1.1) suggest that both b_h^C and b_h^D are still high-order approximations to $b(x)$ when it is smooth. If we need to maintain the well-balanced and positivity-preserving properties simultaneously, we compute $(h+b)_h^C$ and $(h+b)_h^D$ by standard L^2 projection at the initial time, and set $h_h^C = (h+b)_h^C - b_h^C$ and $h_h^D = (h+b)_h^D - b_h^D$. With this, (12) is satisfied. Meanwhile, the bottom approximations b_h^C and b_h^D are computed as in (28).

2.2 Two-Dimensional Case

We now extend the CDG methods proposed in Sect. 2.1 to the 2D NSWEs. The governing equations (2) can be written compactly as

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x + \mathbf{G}(\mathbf{U})_y = \mathbf{S}(\mathbf{U}; b), \tag{29}$$

where $\mathbf{U} = (h, hu, hv)^\top$,

$$\mathbf{F}(\mathbf{U}) = \left(hu, hu^2 + \frac{1}{2}gh^2, huv\right)^\top, \qquad \mathbf{G}(\mathbf{U}) = \left(hv, huv, hv^2 + \frac{1}{2}gh^2\right)^\top$$

are the fluxes, and

$$\mathbf{S}(\mathbf{U}; b) = \left(0, -ghb_x, -ghb_y\right)^\top$$

is the source term.

Let $\mathcal{T}_h^C = \{C_{ij}, \forall i, j\}$ and $\mathcal{T}_h^D = \{D_{ij}, \forall i, j\}$ define two overlapping meshes for the computational domain $\Omega = [x_{\min}, x_{\max}]\times[y_{\min}, y_{\max}]$, with $C_{ij} = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]\times[y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}]$, $D_{ij} = [x_{i-1}, x_i]\times[y_{j-1}, y_j]$, $x_{i+\frac{1}{2}} = \frac{1}{2}(x_i + x_{i+1})$ and $y_{j+\frac{1}{2}} = \frac{1}{2}(y_j + y_{j+1})$, where $\{x_i\}_i$ and $\{y_j\}_j$ are partitions of $[x_{\min}, x_{\max}]$ and $[y_{\min}, y_{\max}]$ respectively. Associated with each mesh, we define the following discrete spaces

$$\mathcal{W}_h^C = \mathcal{W}_h^{C,k} = \left\{\mathbf{v} : \mathbf{v}|_{C_{ij}} \in \left[P^k\left(C_{ij}\right)\right]^3, \forall i, j\right\},$$
$$\mathcal{W}_h^D = \mathcal{W}_h^{D,k} = \left\{\mathbf{v} : \mathbf{v}|_{D_{ij}} \in \left[P^k\left(D_{ij}\right)\right]^3, \forall i, j\right\}.$$

As in the 1D discussion, we only present the numerical schemes with the forward Euler method for time discretization. Higher-order time discretizations will be discussed in Sect. 2.3. The proposed methods evolve two copies of numerical solution, which are assumed to be available at $t = t_n$, denoted by $\mathbf{U}_h^{n,\star} = (h_h^{n,\star}, (hu)_h^{n,\star}, (hv)_h^{n,\star})^\top \in \mathcal{W}_h^\star$, and we want to find the solutions at $t = t_{n+1} = t_n + \Delta t_n$. Only the procedure to update $\mathbf{U}_h^{n+1,C}$ will be described. Except in Sect. 2.2.4, we project the bottom topography function b into $P^k(C_{ij})$ on C_{ij} (resp. $P^k(D_{ij})$ on D_{ij}) in the L^2 sense, and obtain an approximation b_h^C (resp. b_h^D) throughout the computational domain.

2.2.1 Standard CDG Methods

Here again, to highlight the proposed improvements, we first apply to (29) the standard CDG methods of Liu et al. [27] for space discretization and the forward Euler method for time discretization. That is, we look for $\mathbf{U}_h^{n+1,C} = (h_h^{n+1,C}, (hu)_h^{n+1,C}, (hv)_h^{n+1,C})^\top \in \mathcal{W}_h^{C,k}$ such that for any $\mathbf{V} \in \mathcal{W}_h^{C,k}|_{C_{ij}}$ with any i and j,

$$\begin{aligned}
\int_{C_{ij}} \mathbf{U}_h^{n+1,C}\cdot\mathbf{V}dxdy = &\int_{C_{ij}} \left(\theta_n\mathbf{U}_h^{n,D} + (1-\theta_n)\,\mathbf{U}_h^{n,C}\right)\cdot\mathbf{V}dxdy\\
&+\Delta t_n\int_{C_{ij}}\left[\mathbf{F}\left(\mathbf{U}_h^{n,D}\right)\cdot\mathbf{V}_x + \mathbf{G}\left(\mathbf{U}_h^{n,D}\right)\cdot\mathbf{V}_y\right]dxdy\\
&-\Delta t_n\int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}}\left[\mathbf{F}\left(\mathbf{U}_h^{n,D}\left(x_{i+\frac{1}{2}}, y\right)\right)\cdot\mathbf{V}\left(x_{i+\frac{1}{2}}^-, y\right)\right.\\
&\left.-\mathbf{F}\left(\mathbf{U}_h^{n,D}\left(x_{i-\frac{1}{2}}, y\right)\right)\cdot\mathbf{V}\left(x_{i-\frac{1}{2}}^+, y\right)\right]dy\\
&-\Delta t_n\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}}\left[\mathbf{G}\left(\mathbf{U}_h^{n,D}\left(x, y_{j+\frac{1}{2}}\right)\right)\cdot\mathbf{V}\left(x, y_{j+\frac{1}{2}}^-\right)\right.\\
&\left.-\mathbf{G}\left(\mathbf{U}_h^{n,D}\left(x, y_{j-\frac{1}{2}}\right)\right)\cdot\mathbf{V}\left(x, y_{j-\frac{1}{2}}^+\right)\right]dx\\
&+\Delta t_n\int_{C_{ij}}\mathbf{S}\left(\mathbf{U}_h^{n,D}; b_h^D\right)\cdot\mathbf{V}dxdy. \qquad (30)
\end{aligned}$$

This numerical scheme does not usually maintain the still-water stationary solution (4).

2.2.2 Well-Balanced CDG Methods

As in (7), the well-balanced property is accounted for by adding terms to scheme (30),

$$\begin{aligned}\int_{C_{ij}} \mathbf{U}_h^{n+1,C}\cdot\mathbf{V}dxdy &= \text{right-hand side of (30)} + \widetilde{\mathbf{U}}\left(b_h^C, b_h^D, \mathbf{V}\right)\\ &+\Delta t_n \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} \left[\widetilde{\mathbf{S}}_1\left(\mathbf{U}_h^{n,D}; b_h^D\left(x_i^+, y\right)\right) - \widetilde{\mathbf{S}}_1\left(\mathbf{U}_h^{n,D}; b_h^D\left(x_i^-, y\right)\right)\right]\cdot\mathbf{V}\left(x_i, y\right)dy\\ &+\Delta t_n \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \left[\widetilde{\mathbf{S}}_2\left(\mathbf{U}_h^{n,D}; b_h^D\left(x, y_j^+\right)\right) - \widetilde{\mathbf{S}}_2\left(\mathbf{U}_h^{n,D}; b_h^D\left(x, y_j^-\right)\right)\right]\cdot\mathbf{V}\left(x, y_j\right)dx,\end{aligned} \tag{31}$$

where the correction terms are given by

$$\widetilde{\mathbf{U}}\left(b_h^C, b_h^D, \mathbf{V}\right) = \theta_n \int_{C_{ij}} \left(b_h^D - b_h^D, 0, 0\right)^\top \cdot \mathbf{V}dxdy, \tag{32}$$

$$\widetilde{\mathbf{S}}_1\left(\mathbf{U}_h^{n,D}; b_h^D\right) = \left(0, \frac{g}{2}\left(b_h^D\right)^2 - g\gamma_{ij}^{n,D} b_h^D, 0\right)^\top, \tag{33}$$

$$\widetilde{\mathbf{S}}_2\left(\mathbf{U}_h^{n,D}; b_h^D\right) = \left(0, 0, \frac{g}{2}\left(b_h^D\right)^2 - g\gamma_{ij}^{n,D} b_h^D\right)^\top, \tag{34}$$

and $\gamma_{ij}^{n,D}$ is a special constant which represents the average value of the water surface $\eta_h^{n,D} = h_h^{n,D} + b_h^D$ over the element C_{ij}. More specifically, we take

$$\begin{aligned}\gamma_{ij}^{n,D} = \frac{1}{4}\Big[&\eta_h^{n,D}\left(x_{i+\frac{1}{2}}, y_{j+\frac{1}{2}}\right) + \eta_h^{n,D}\left(x_{i-\frac{1}{2}}, y_{j+\frac{1}{2}}\right)\\ &+ \eta_h^{n,D}\left(x_{i+\frac{1}{2}}, y_{j-\frac{1}{2}}\right) + \eta_h^{n,D}\left(x_{i-\frac{1}{2}}, y_{j-\frac{1}{2}}\right)\Big].\end{aligned}$$

With the following decomposition of the source term

$$\begin{aligned}\mathbf{S}\left(\mathbf{U}_h^{n,D}; b_h^D\right) &= \left(0, -g\left(\eta_h^{n,D} - b_h^D\right)\left(b_h^D\right)_x, -g\left(\eta_h^{n,D} - b_h^D\right)\left(b_h^D\right)_y\right)^\top\\ &= \left(0, \frac{g}{2}\left(b_h^D\right)^2 - g\gamma_{ij}^{n,D} b_h^D, 0\right)_x^\top + \left(0, 0, \frac{g}{2}\left(b_h^D\right)^2 - g\gamma_{ij}^{n,D} b_h^D\right)_y^\top\\ &\quad - \left(0, g\left(\eta_h^{n,D} - \gamma_{ij}^{n,D}\right)\left(b_h^D\right)_x, g\left(\eta_h^{n,D} - \gamma_{ij}^{n,D}\right)\left(b_h^D\right)_y\right)^\top,\end{aligned} \tag{35}$$

scheme (32) is rewritten as

$$\begin{aligned}&\int_{C_{ij}} \mathbf{U}_h^{n+1,C}\cdot\mathbf{V}dxdy\\ &= \int_{C_{ij}} \left(\theta_n\mathbf{U}_h^{n,D} + (1-\theta_n)\mathbf{U}_h^{n,C}\right)\cdot\mathbf{V}dxdy + \widetilde{\mathbf{U}}\left(b_h^C, b_h^D, \mathbf{V}\right)\\ &\quad +\Delta t_n \int_{C_{ij}} \left[\mathbf{F}\left(\mathbf{U}_h^{n,D}\right)\cdot\mathbf{V}_x + \mathbf{G}\left(\mathbf{U}_h^{n,D}\right)\cdot\mathbf{V}_y\right]dxdy\\ &\quad -\Delta t_n \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} \left[\mathbf{F}\left(\mathbf{U}_h^{n,D}\left(x_{i+\frac{1}{2}}, y\right)\right)\cdot\mathbf{V}\left(x_{i+\frac{1}{2}}^-, y\right)\right.\end{aligned}$$

$$
\begin{aligned}
&-\mathbf{F}\left(\mathbf{U}_h^{n,D}\left(x_{i-\frac{1}{2}},y\right)\right)\cdot\mathbf{V}\left(x_{i-\frac{1}{2}}^{+},y\right)\Big]dy\\
&-\Delta t_n\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}}\Big[\mathbf{G}\left(\mathbf{U}_h^{n,D}\left(x,y_{j+\frac{1}{2}}\right)\right)\cdot\mathbf{V}\left(x,y_{j+\frac{1}{2}}^{-}\right)\\
&-\mathbf{G}\left(\mathbf{U}_h^{n,D}\left(x,y_{j-\frac{1}{2}}\right)\right)\cdot\mathbf{V}\left(x,y_{j-\frac{1}{2}}^{+}\right)\Big]dx\\
&-\Delta t_n\int_{C_{ij}}\left(0,g\left(\eta_h^{n,D}-\gamma_{ij}^{n,D}\right)\left(b_h^D\right)_x,g\left(\eta_h^{n,D}-\gamma_{ij}^{n,D}\right)\left(b_h^D\right)_y\right)^{\top}\cdot\mathbf{V}dxdy\\
&-\Delta t_n\int_{C_{ij}}\left(0,\frac{g}{2}\left(b_h^D\right)^2-g\gamma_{ij}^{n,D}\right)b_h^D,0\Big)^{\top}\cdot\mathbf{V}_x dxdy\\
&-\Delta t_n\int_{C_{ij}}\left(0,0,\frac{g}{2}\left(b_h^D\right)^2-g\gamma_{ij}^{n,D}\right)b_h^D\Big)^{\top}\cdot\mathbf{V}_y dxdy\\
&+\Delta t_n\int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}}\left(0,\frac{g}{2}\left(b_h^D\left(x_{i+\frac{1}{2}},y\right)\right)^2-g\gamma_{ij}^{n,D}\right)b_h^D\left(x_{i+\frac{1}{2}},y\right),0\Big)^{\top}\cdot\mathbf{V}\left(x_{i+\frac{1}{2}}^{-},y\right)dy\\
&-\Delta t_n\int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}}\left(0,\frac{g}{2}\left(b_h^D\left(x_{i-\frac{1}{2}},y\right)\right)^2-g\gamma_{ij}^{n,D}\right)b_h^D\left(x_{i-\frac{1}{2}},y\right),0\Big)^{\top}\cdot\mathbf{V}\left(x_{i-\frac{1}{2}}^{+},y\right)dy\\
&+\Delta t_n\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}}\left(0,0,\frac{g}{2}\left(b_h^D\left(x,y_{j+\frac{1}{2}}\right)\right)^2-g\gamma_{ij}^{n,D}\right)b_h^D\left(x,y_{j+\frac{1}{2}}\right)\Big)^{\top}\cdot\mathbf{V}\left(x,y_{j+\frac{1}{2}}^{-}\right)dx\\
&-\Delta t_n\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}}\left(0,0,\frac{g}{2}\left(b_h^D\left(x,y_{j-\frac{1}{2}}\right)\right)^2-g\gamma_{ij}^{n,D}\right)b_h^D\left(x,y_{j-\frac{1}{2}}\right)\Big)^{\top}\cdot\mathbf{V}(x,y_{j-\frac{1}{2}}^{+})dx.
\end{aligned}
\tag{36}
$$

Theorem 4 *The numerical scheme defined in* (36)*, and its counterpart for* $\mathbf{U}_h^{n+1,D}$*, to solve the 2D NSWEs is well-balanced in the sense that it preserves the still-water stationary solution* (4).

Proof The proof is a direct 2D generalization of that for Theorem 1 and thus is omitted here. □

2.2.3 Positivity-Preserving CDG Methods

We now extend the positivity-preserving CDG methods introduced in Sect. 2.1.3 for the 1D NSWEs to the 2D NSWEs.

Firstly, let $\widehat{L}_i^{1,x}=\{\widehat{x}_i^{1,\beta},\beta=1,2,\ldots,\widehat{N}\}$ and $\widehat{L}_i^{2,x}=\{\widehat{x}_i^{2,\beta},\beta=1,2,\ldots,\widehat{N}\}$ be the Legendre Gauss–Lobatto quadrature points on $[x_{i-\frac{1}{2}},x_i]$ and $[x_i,x_{i+\frac{1}{2}}]$ respectively, while $\widehat{L}_j^{1,y}=\{\widehat{y}_j^{1,\beta},\beta=1,2,\ldots,\widehat{N}\}$ and $\widehat{L}_j^{2,y}=\{\widehat{y}_j^{2,\beta},\beta=1,2,\ldots,\widehat{N}\}$ represent the Legendre Gauss–Lobatto quadrature points on $[y_{j-\frac{1}{2}},y_j]$ and $[y_j,y_{j+\frac{1}{2}}]$ respectively, $\forall i,j$. The corresponding quadrature weights on the reference element $[-\frac{1}{2},\frac{1}{2}]$ are $\{\widehat{\omega}_\beta,\beta=1,2,\ldots,\widehat{N}\}$, and $\widehat{N}$ is chosen such that $2\widehat{N}-3\geq k$. In addition, let $L_i^{1,x}=\{x_i^{1,\alpha},\alpha=1,2,\ldots,N\}$ and $L_i^{2,x}=\{x_i^{2,\alpha},\alpha=1,2,\ldots,N\}$ denote the Gaussian quadrature points on $[x_{i-\frac{1}{2}},x_i]$

and $[x_i, x_{i+\frac{1}{2}}]$ respectively, while $L_j^{1,y} = \{y_j^{1,\alpha}, \alpha = 1, 2, \ldots, N\}$ and $L_j^{2,y} = \{y_j^{2,\alpha}, \alpha = 1, 2, \ldots, N\}$ represent the Gaussian quadrature points on $[y_{j-\frac{1}{2}}, y_j]$ and $[y_j, y_{j+\frac{1}{2}}]$ respectively. The corresponding quadrature weights $\{\omega_\alpha, \alpha = 1, 2, \ldots, N\}$ are distributed on the interval $[-\frac{1}{2}, \frac{1}{2}]$ and N is chosen such that the Gaussian quadrature is exact for the integration of univariate polynomials of degree $2k+1$. Then denote $L_{i,j}^{l,m} = (L_i^{l,x} \otimes \widehat{L}_j^{m,y}) \cup (\widehat{L}_i^{l,x} \otimes L_j^{m,y})$ with $l, m = \{1, 2\}$.

Similar to the 1D case, we define $\widetilde{(hu)}_h^{n,C}$, $\widetilde{(hv)}_h^{n,C}$, $\widetilde{(hu)}_h^{n,D}$ and $\widetilde{(hv)}_h^{n,D}$ by

$$\min_{\widetilde{(hu)}_h^{n,C} | C_{ij} \in P^k(C_{ij})} \quad \|\widetilde{(hu)}_h^{n,C} - (hu)_h^{n,C}\|_{L^2(C_{ij})},$$

$$\text{subject to } \widetilde{(hu)}_h^{n,C}(x, y) = 0, \forall (x, y) \in \mathcal{Q}_C, \tag{37}$$

$$\min_{\widetilde{(hv)}_h^{n,C} | C_{ij} \in P^k(C_{ij})} \quad \|\widetilde{(hv)}_h^{n,C} - (hv)_h^{n,C}\|_{L^2(C_{ij})},$$

$$\text{subject to } \widetilde{(hv)}_h^{n,C}(x, y) = 0, \forall (x, y) \in \mathcal{Q}_C, \tag{38}$$

$$\min_{\widetilde{(hu)}_h^{n,D} | D_{ij} \in P^k(D_{ij})} \quad \|\widetilde{(hu)}_h^{n,D} - (hu)_h^{n,D}\|_{L^2(D_{ij})},$$

$$\text{subject to } \widetilde{(hu)}_h^{n,D}(x, y) = 0, \forall (x, y) \in \mathcal{Q}_D, \tag{39}$$

$$\min_{\widetilde{(hv)}_h^{n,D} | D_{ij} \in P^k(D_{ij})} \quad \|\widetilde{(hv)}_h^{n,D} - (hv)_h^{n,D}\|_{L^2(D_{ij})},$$

$$\text{subject to } \widetilde{(hv)}_h^{n,D}(x, y) = 0, \forall (x, y) \in \mathcal{Q}_D, \tag{40}$$

with $\mathcal{Q}_C = \{(x_i, y) : h_h^{n,C}(x_i, y) = 0, y \in L_j^{1,y} \cup L_j^{2,y}\} \cup \{(x, y_j) : h_h^{n,C}(x, y_j) = 0, x \in L_i^{1,x} \cup L_i^{2,x}\}$ and $\mathcal{Q}_D = \{(x_{i-\frac{1}{2}}, y) : h_h^{n,C}(x_{i-\frac{1}{2}}, y) = 0, y \in L_{j-1}^{2,y} \cup L_j^{1,y}\} \cup \{(x, y_{j-\frac{1}{2}}) : h_h^{n,C}(x, y_{i-\frac{1}{2}}) = 0, x \in L_{i-1}^{2,x} \cup L_i^{1,x}\}$ for any given i, j. If $\mathcal{Q}_C$ is an empty set, then $\widetilde{(hu)}_h^{n,C} = (hu)_h^{n,C}$ and $\widetilde{(hv)}_h^{n,C} = (hv)_h^{n,C}$. If $\mathcal{Q}_D$ is empty, then $\widetilde{(hu)}_h^{n,D} = (hu)_h^{n,D}$ and $\widetilde{(hv)}_h^{n,D} = (hv)_h^{n,D}$.

We modify the CDG scheme (30) by replacing $\mathbf{U}_h^{n,C}$ and $\mathbf{U}_h^{n,D}$ with $\widetilde{\mathbf{U}}_h^{n,C} = (h_h^{n,C}, \widetilde{(hu)}_h^{n,C}, \widetilde{(hv)}_h^{n,C})^\top$ and $\widetilde{\mathbf{U}}_h^{n,D} = (h_h^{n,D}, \widetilde{(hu)}_h^{n,D}, \widetilde{(hv)}_h^{n,D})^\top$ respectively, and by taking the test function $\mathbf{V} = (\frac{1}{\Delta x \Delta y}, 0, 0)^\top$. This leads to the equation satisfied by the cell average of h_h^C,

$$\begin{aligned}
\bar{h}_{ij}^{n+1,C} = (1 - \theta_n)\, \bar{h}_{ij}^{n,C} + \frac{\theta_n}{\Delta x \Delta y} \int_{C_{ij}} h_h^{n,D} dx dy \\
- \frac{\Delta t_n}{\Delta x \Delta y} \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} \left[\widetilde{(hu)}_h^{n,D}\left(x_{i+\frac{1}{2}}, y\right) - \widetilde{(hu)}_h^{n,D}\left(x_{i-\frac{1}{2}}, y\right) \right] dy \\
- \frac{\Delta t_n}{\Delta x \Delta y} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \left[\widetilde{(hv)}_h^{n,D}\left(x, y_{j+\frac{1}{2}}\right) - \widetilde{(hv)}_h^{n,D}\left(x, y_{j-\frac{1}{2}}\right) \right] dx,
\end{aligned} \tag{41}$$

where $\bar{h}_{ij}^{n,C}$ denotes the cell average of the CDG solution h_h^C over C_{ij} at time t_n.

If the integrals along mesh interfaces in (41) are evaluated by applying the Gaussian quadrature rule described above to each half of the interval (see also Section 3.2 in [8]), then we have the following theorem.

Theorem 5 *For any given $n \geq 0$, we assume $\bar{h}_{ij}^{n,C} \geq 0$ and $\bar{h}_{ij}^{n,D} \geq 0, \forall i, j$. Consider the numerical scheme in (41) and its counterpart for $\bar{h}_{ij}^{n+1,D}$, if $h_h^C(x, y, t_n) \geq 0$ and $h_h^D(x, y, t_n) \geq 0, \forall (x, y) \in L_{i,j}^{l,m}, \forall i, j$ with $l, m = \{1, 2\}$, then $\bar{h}_{ij}^{n+1,C} \geq 0$ and $\bar{h}_{ij}^{n+1,D} \geq 0, \forall i, j$, under the CFL condition*

$$\lambda_x a_x + \lambda_y a_y \leq \frac{1}{2}\theta_n \widehat{\omega}_1, \tag{42}$$

where λ_x and a_x are the same as in Theorem 2, while $\lambda_y = \Delta t_n/\Delta y$ and $a_y = \max(\|v_h^{n,C}\|_\infty, \|v_h^{n,D}\|_\infty)$.

Proof The proof is a direct 2D generalization of that for Theorem 2 and thus is omitted here. □

Next, we will give positivity-preserving limiters which modify the CDG solution polynomials h_h^C and h_h^D into $\widetilde{h}_h^C$ and $\widetilde{h}_h^D$ such that they satisfy the sufficient condition given in Theorem 5. In fact, these limiters are the same as in Sect. 2.1.3 for the 1D case, as long as the notations K and $\widehat{L}_K$ are re-defined as follows: on the primal mesh, K denotes a mesh element C_{ij} and $\widehat{L}_K$ represents the set of relevant quadrature points in K, namely $\widehat{L}_K = \cup_{l,m=1}^2 L_{i,j}^{l,m}$. On the dual mesh, K denotes a mesh element D_{ij} and $\widehat{L}_K$ represents the set of relevant quadrature points in K, namely $\widehat{L}_K = L_{i,j}^{1,1} \cup L_{i,j-1}^{1,2} \cup L_{i-1,j}^{2,1} \cup L_{i-1,j-1}^{2,2}$. In actual implementation, the comments of Remark 1 in Sect. 2.1.3 are incorporated.

2.2.4 Positivity-Preserving Well-Balanced CDG Methods

Finally, we extend the combined positivity-preserving and well-balanced CDG method proposed in Sect. 2.1.4 to the 2D NSWEs. We start with the well-balanced CDG scheme satisfied by the cell average of the numerical solution h_h^C, which is obtained by taking the test function $\mathbf{V} = (\frac{1}{\Delta x \Delta y}, 0, 0)^\top$, and by replacing $\mathbf{U}_h^{n,C}$ and $\mathbf{U}_h^{n,D}$ with $\widetilde{\mathbf{U}}_h^{n,C} = (h_h^{n,C}, \widetilde{(hu)}_h^{n,C}, \widetilde{(hv)}_h^{n,C})^\top$ and $\widetilde{\mathbf{U}}_h^{n,D} = (h_h^{n,D}, \widetilde{(hu)}_h^{n,D}, \widetilde{(hv)}_h^{n,D})^\top$ in scheme (36),

$$\begin{aligned}
\bar{h}_{ij}^{n+1,C} = {} & (1-\theta_n)\,\bar{h}_{ij}^{n,C} + \frac{\theta_n}{\Delta x \Delta y}\int_{C_{ij}} h_h^{n,D} dxdy \\
& - \frac{\Delta t_n}{\Delta x \Delta y}\int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} \left[\widetilde{(hu)}_h^{n,D}\left(x_{i+\frac{1}{2}}, y\right) - \widetilde{(hu)}_h^{n,D}\left(x_{i-\frac{1}{2}}, y\right)\right] dy \\
& - \frac{\Delta t_n}{\Delta x \Delta y}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \left[\widetilde{(hv)}_h^{n,D}\left(x, y_{j+\frac{1}{2}}\right) - \widetilde{(hv)}_h^{n,D}\left(x, y_{j-\frac{1}{2}}\right)\right] dx \\
& + \theta_n\left(\frac{1}{\Delta x \Delta y}\int_{C_{ij}} b_h^D \cdot dxdy - \bar{b}_{ij}^C\right),
\end{aligned} \tag{43}$$

where $\bar{b}_{ij}^C$ (resp. $\bar{b}_{ij}^D$) denotes the cell average of the bottom topography b_h^C (resp. b_h^D) over the element C_{ij} (resp. D_{ij}).

Theorem 6 *For any given $n \geq 0$, we assume $\bar{h}_{ij}^{n,C} \geq 0$ and $\bar{h}_{ij}^{n,D} \geq 0$, $\forall i, j$. Consider the numerical scheme in (43) and its counterpart for $\bar{h}_{ij}^{n+1,D}$, if*

$$\bar{b}_{ij}^{C} = \frac{1}{\Delta x \Delta y} \int_{C_{ij}} b_h^D dxdy, \qquad \bar{b}_{ij}^{D} = \frac{1}{\Delta x \Delta y} \int_{D_{ij}} b_h^C dxdy, \quad \forall i, j, \tag{44}$$

and $h_h^C(x, y, t_n) \geq 0, h_h^D(x, y, t_n) \geq 0, \forall (x, y) \in L_{i,j}^{l,m}, \forall i, j$ with $l, m = \{1, 2\}$, then $\bar{h}_{ij}^{n+1,C} \geq 0$ and $\bar{h}_{ij}^{n+1,D} \geq 0$, $\forall i, j$, under the CFL condition

$$\lambda_x a_x + \lambda_y a_y \leq \frac{1}{2} \theta_n \widehat{\omega}_1. \tag{45}$$

Proof The proof is a direct consequence of Theorem 5 and condition (44). □

To enforce the sufficient condition in Theorem 6, we also need to modify the approximations to the bottom topography and use a positivity-preserving limiter as in Sect. 2.1.4.

2.3 High-Order Time Discretizations and Nonlinear Limiters

To achieve better accuracy in time, strong stability preserving (SSP) high-order time discretizations will be used in the numerical simulations [14]. Such discretizations can be written as a convex combination of the forward Euler method, and therefore the resulting SSP schemes are also well-balanced and positivity-preserving.

When CDG methods are applied to nonlinear problems, nonlinear limiters are often needed to prevent numerical instabilities. In this work, we observed that nonlinear limiters are needed when the solution develops sharp gradients. If so, we use the total variation bounded (TVB) minmod slope limiter with parameter $M = 1$ and it is implemented in local characteristic fields [9]. This limiter is applied to $(h + b, hu)^\top$ and $(h + b, hu, hv)^\top$ in the 1D and 2D methods, respectively. It is used prior to application of the positivity-preserving limiter.

3 Numerical Examples

In this section, numerical experiments are presented to demonstrate the performance of the proposed methods for solving the 1D and 2D NSWEs. All simulations were performed with both P^1 and P^2 approximations. We only show P^2 results, except for the accuracy tests with varying mesh sizes, and all reported results are from numerical solutions on the primal mesh $\{I_j\}_j$ in 1D and $\{C_{ij}\}_{i,j}$ in 2D. In all simulations, we used a uniform mesh with constant mesh sizes Δx and Δy in the x- and y-directions respectively. All errors reported here are absolute errors in the L^2 or L^∞ sense.

A third-order TVD Runge–Kutta algorithm is employed for time discretization [9], with the time step being adaptively determined by

$$\Delta t_n = \frac{C_{\text{cfl}}}{\frac{a_x}{\Delta x}} \quad \text{(1D case)} \qquad \text{or} \qquad \Delta t_n = \frac{C_{\text{cfl}}}{\frac{a_x}{\Delta x} + \frac{a_y}{\Delta y}} \quad \text{(2D case)},$$

where C_{cfl} denotes the CFL number. We chose $C_{\text{cfl}} = 0.3$ in 1D and $C_{\text{cfl}} = 0.5$ in 2D for both P^1 and P^2 approximations when the standard or well-balanced CDG methods were used, otherwise $C_{\text{cfl}} = 1/12$ when the positivity-preserving CDG methods or combined positivity-preserving well-balanced CDG methods were applied. We also set $\theta_n = 1$, and non-dimensionalize the equations in such a way that the gravitational constant g is unity.

When h and hu (or hv) are both small but $u = hu/h$ (or $v = hv/h$) is very large, this may lead to prohibitively small time steps. To overcome this difficulty, we always set $u = 0$ (or $v = 0$) if $0 \le h < \varepsilon_0$ where ε_0 is a small positive constant. In our implementation, we choose $\varepsilon_0 = 10^{-12}$ for example 3.1.7 (flat bottom) and $\varepsilon_0 = (\Delta x)^{k+1}$ for examples 3.1.8, 3.1.9 and 3.2.4 (variable bottom). Otherwise, the choice $\varepsilon_0 = 10^{-12}$ would produce a large velocity u and thus the time step would become very small in the case of a variable bottom.

In addition to accuracy tests, we present various applications where the well-balanced and positivity-preserving properties are examined separately. The standard and well-balanced CDG schemes are comparable in terms of computational cost. It is observed that the CPU time required by the well-balanced scheme is only 2% more than that for the standard scheme. Using the same CFL number ($C_{\text{cfl}} = 1/12$), we compared the computational cost between CDG schemes with and without positivity-preserving limiter, and found that the positivity-preserving CDG scheme is about 9% more costly than the standard scheme.

It is also pointed out that, although the presence of overlapping meshes in the positivity-preserving CDG method implies a CFL number $C_{\text{cfl}} = 1/12$ which is smaller than the one required by DG and standard CDG methods, in practice one can adopt a dynamical strategy by starting with the usual CFL number and, if the computed water depth h happens to be negative, the simulation returns to the previous discrete time and uses a smaller CFL number ($C_{\text{cfl}} = 1/12$) for a few steps, after that the usual CFL number is re-activated. In this way, the computational cost can be significantly reduced.

3.1 Examples for 1D NSWEs

3.1.1 Accuracy Tests for Bottom Approximation

As the first example, we test the high-order accuracy of the bottom approximation obtained from the constrained minimization problem (28) by varying the mesh size. We choose the following smooth function

$$b(x) = \sin^2(x),$$

for the bottom profile. The computational domain is $[0, 2\pi]$. We present L^2 errors and orders of accuracy for b_h^C and b_h^D in Table 1. We can observe that the results are $(k+1)$st-order accurate for P^k with $k = \{1, 2\}$. We further consider a discontinuous bottom,

$$b(x) = \begin{cases} 1.0, & 0.5 \le x \le 1.5, \\ 0, & \text{otherwise.} \end{cases}$$

Figure 1 shows the approximate solution b_h^C of (28) for varying mesh size. Although high-order accuracy is not guaranteed in this case, we can see that the approximation converges to the exact discontinuous profile as the mesh is refined.

3.1.2 Accuracy Tests for Well-Balanced CDG Solution

In this example, we test the high-order accuracy of the well-balanced CDG method by varying the mesh size. We choose the following smooth functions

$$h(x, 0) = 5 + e^{\cos(2\pi x)}, \qquad (hu)(x, 0) = \sin(\cos(2\pi x)) \tag{46}$$

for the initial conditions, and

$$b(x) = \sin^2(\pi x) \tag{47}$$

Table 1 L^2 errors and orders of accuracy for b_h^C and b_h^D as obtained from (28)

Mesh	b_h^C		b_h^D	
	L^2 error	Order	L^2 error	Order
P^1				
50	1.85E−03	–	1.87E−03	–
100	4.65E−04	1.99	4.67E−04	2.00
200	1.16E−04	2.00	1.16E−04	2.01
400	2.90E−05	2.00	2.91E−05	2.00
800	7.26E−06	2.00	7.27E−06	2.00
P^2				
50	5.83E−05	–	5.77E−05	–
100	7.30E−06	3.00	7.26E−06	2.99
200	9.13E−07	3.00	9.11E−07	2.99
400	1.14E−07	3.00	1.14E−07	3.00
800	1.42E−08	3.00	1.42E−08	3.00

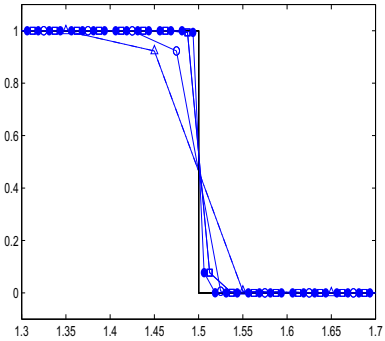


Fig. 1 Zoomed-in bottom approximation b_h^C as given by (28) for varying mesh size. *Black solid line* exact profile. *Triangles* 20 cells; *circles* 40 cells; *squares* 80 cells; *dots* 160 cells

for the bottom profile [35]. The computational domain is $[0, 1]$ with periodic boundary conditions. The final time is $t = 0.2$ at which the solution is still smooth (indeed it tends to steepen in time). Since an exact solution is unknown for this problem, we first compute the numerical solution on a very fine mesh with 8000 elements, and then use it as the reference solution to evaluate errors and orders of accuracy for coarser resolutions. The standard CDG method in (6) and its well-balanced version in (7) are used to simulate this problem. We present L^2 errors and orders of accuracy for h and hu in Table 2 by the standard CDG method and in Table 3 by the well-balanced one. We can observe that both methods are $(k + 1)$st-order accurate for P^k with $k = \{1, 2\}$ and therefore the well-balanced treatment does not affect the accuracy of the standard CDG method.

3.1.3 Stationary Solution

In this test, we validate the well-balanced feature, i.e. the ability to preserve still-water steady states, of the proposed well-balanced CDG method in cases of continuous (non-smooth), discontinuous and smooth bottoms. The initial conditions (which correspond to a stationary solution) are given by

Table 2 L^2 errors and orders of accuracy for (h, hu) at $t = 0.2$ by the standard CDG method

Mesh	h		hu	
	L^2 error	Order	L^2 error	Order
P^1				
25	5.65E−03	–	1.38E−02	–
50	1.05E−03	2.41	2.58E−03	2.42
100	1.87E−04	2.49	4.47E−04	2.52
200	3.84E−05	2.29	8.88E−05	2.33
400	8.93E−06	2.10	2.03E−05	2.12
800	2.18E−06	2.03	4.96E−06	2.03
P^2				
25	2.85E−04	–	7.34E−04	–
50	2.70E−05	3.40	6.94E−05	3.40
100	3.23E−06	3.06	8.31E−06	3.06
200	4.02E−07	3.00	1.03E−06	3.00
400	5.02E−08	3.00	1.29E−07	3.00
800	6.24E−09	3.00	1.60E−08	3.00

Table 3 L^2 errors and orders of accuracy for (h, hu) at $t = 0.2$ by the well-balanced CDG method

Mesh	h		hu	
	L^2 error	Order	L^2 error	Order
P^1				
25	5.67E−03	–	1.39E−02	–
50	1.06E−03	2.41	2.59E−03	2.42
100	1.88E−04	2.49	4.51E−04	2.52
200	3.85E−05	2.29	8.98E−05	2.32
400	8.95E−06	2.10	2.06E−05	2.12
800	2.18E−06	2.03	5.02E−06	2.03
P^2				
25	2.86E−04	–	7.35E−04	–
50	2.71E−05	3.40	6.94E−05	3.40
100	3.24E−06	3.07	8.32E−06	3.06
200	4.03E−07	3.01	1.04E−06	3.00
400	5.03E−08	3.00	1.29E−07	3.00
800	6.25E−09	3.01	1.61E−08	3.01

$$hu = 0, \qquad h + b = 1. \tag{48}$$

The continuous bottom profile is defined by

$$b(x) = \begin{cases} 4(x - 0.2), & 0.2 \le x \le 0.4, \\ 0.8, & 0.4 \le x \le 0.6, \\ -4(x - 0.8), & 0.6 \le x \le 0.8, \\ 0, & \text{otherwise}, \end{cases}$$

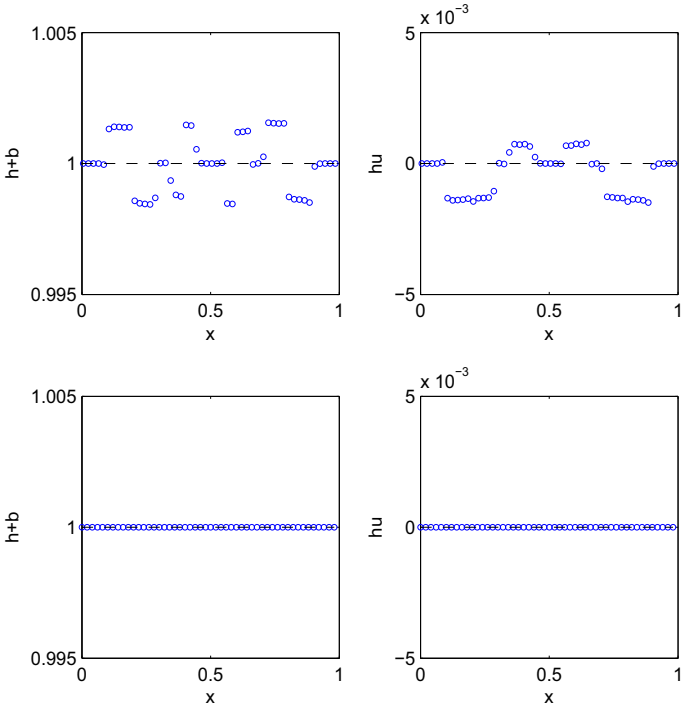


Fig. 2 Numerical results for the stationary solution (48) over a continuous bottom at $t = 0.5$. *Circles* numerical solution; *dashed line* exact solution. *Upper panel* standard CDG method; *lower panel* well-balanced CDG method

the discontinuous one is given by

$$b(x) = \begin{cases} 0.5, & 0.3 \le x \le 0.7, \\ 0, & \text{otherwise}, \end{cases}$$

while the smooth one is given by

$$b(x) = 0.5 \sin(2\pi x), \quad 0 \le x \le 1.$$

We choose [0, 1] as the computational domain, divided into 100 elements, and use outgoing boundary conditions for the continuous and discontinuous bottoms and periodic boundary condition for the smooth bottom. We compute the solution up to $t = 0.5$ by both the standard and well-balanced CDG methods. The resulting water level $h+b$ and discharge hu are plotted in Figs. 2, 3 and 4 for the continuous, discontinuous and smooth bottoms, respectively. We see that the standard CDG method fails to maintain the stationary solution (48) regardless of the bottom profile being continuous or discontinuous, smooth or non-smooth. We also observe that the errors related to smooth bottom are smaller than the ones related to non-smooth bottoms. Note however that after refining the mesh, we can observe the convergence of the computed solution. For the well-balanced CDG method, we can see that the stationary solution is well preserved, and this is also true on a coarser mesh (say, with 10 elements). To further demonstrate that the well-balanced scheme indeed preserves the still-water stationary solution exactly (i.e. up to machine precision), we perform the computation in both single and double precision. The corresponding L^2 and L^∞ errors on $h + b$ and hu are listed in Table 4

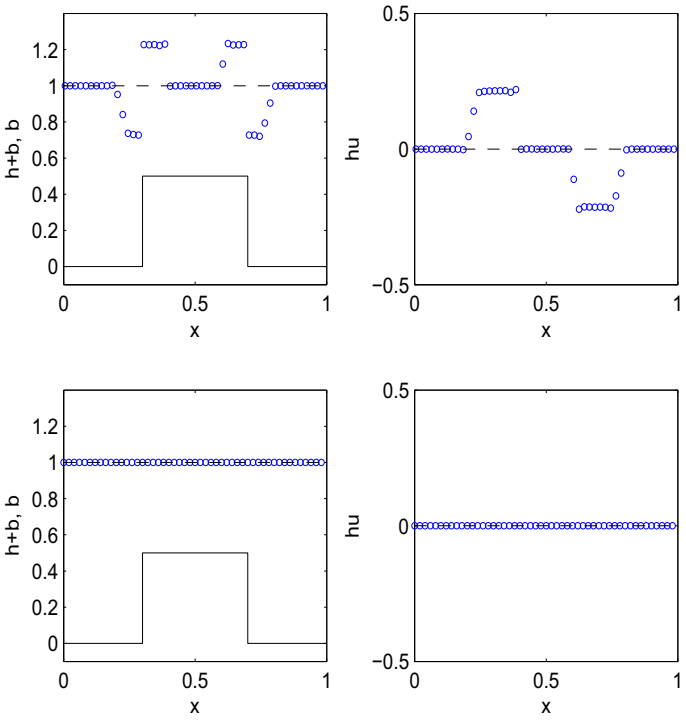


Fig. 3 Numerical results for the stationary solution (48) over a discontinuous bottom at $t = 0.5$. *Circles* numerical solution; *dashed line* exact solution; *solid line* bottom profile. *Upper panel* standard CDG method, *lower panel* well-balanced CDG method

for these topographies. Their values have orders of magnitude consistent with machine single and double precision, thus verifying the well-balanced property.

3.1.4 Perturbation of a Stationary Solution

We consider small perturbations of a stationary solution over a hump [20,36], with the bottom topography given by

$$b(x) = \begin{cases} 0.25(\cos(10\pi(x - 1.5)) + 1), & 1.4 \le x \le 1.6, \\ 0, & \text{otherwise.} \end{cases}$$

The initial conditions are

$$(hu)(x, 0) = 0 \quad \text{and} \quad h(x, 0) = \begin{cases} 1 - b(x) + \varepsilon, & 1.1 \le x \le 1.2, \\ 1 - b(x), & \text{otherwise,} \end{cases}$$

where ε is a non-zero perturbation parameter. We examine two cases: $\varepsilon = 0.2$ (big pulse) and $\varepsilon = 0.001$ (small pulse). Theoretically, for small ε, this disturbance should split into two waves, one propagating to the left at the characteristic speed $-\sqrt{gh}$ and the other propagating to the right at the characteristic speed $\sqrt{gh}$.

We set $[0, 2]$ as the computational domain and use outgoing boundary conditions. We simulate the problem up to $t = 0.63$ by both the standard and well-balanced CDG methods

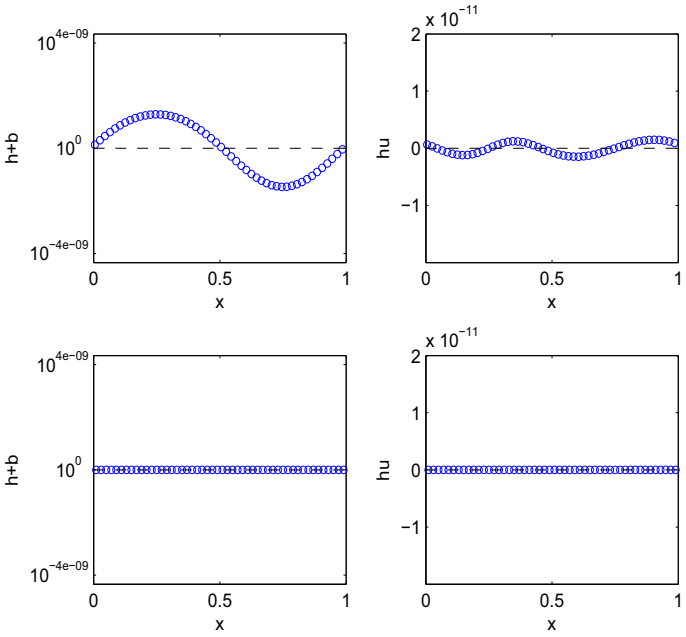


Fig. 4 Numerical results for the stationary solution (48) over a smooth bottom at $t = 0.5$. *Circles* numerical solution; *dashed line* exact solution. *Upper panel* standard CDG method, *lower panel* well-balanced CDG method

on two different meshes (with 200 and 2000 elements). At that moment, the downstream-traveling water pulse has already passed the hump.

We present numerical results for the big and small pulses in Figs. 5 and 6, respectively. It can be seen that when the standard CDG method is applied to simulate the small perturbation pulse, the numerical solution exhibits spurious oscillations of large amplitude on coarse meshes. To remove such oscillations, finer meshes need to be used and this has been confirmed using additional numerical results with varying resolutions (not shown here). On the other hand, the standard CDG method seems to perform well overall in the case of a big pulse, with some oscillations of small amplitude on coarser meshes. For the well-balanced CDG scheme, however, no spurious oscillations are discernible for either the small or big pulse, even when a coarse mesh is used.

3.1.5 Dam Breaking Over a Rectangular Bottom

We investigate the dam-breaking problem over a rectangular bottom [33,35], with initial conditions

$$(hu)(x,0) = 0, \qquad h(x,0) = \begin{cases} 2 - b(x), & x \le 75, \\ 1.5 - b(x), & \text{otherwise}, \end{cases}$$

and bottom topography

$$b(x) = \begin{cases} 0.8, & |x - 75| \le 75/4, \\ 0, & \text{otherwise}. \end{cases}$$

Table 4 L^2 and L^∞ errors on $(h+b, hu)$ for the stationary solution at $t = 0.5$ by the well-balanced CDG method

Precision	$h+b$		hu	
	L^2 error	L^∞ error	L^2 error	L^∞ error
Continuous bottom				
Single	5.79 E−10	2.15 E−09	1.28 E−12	5.86E−12
Double	5.79 E−16	2.22 E−15	3.56 E−16	1.28E−15
Discontinuous bottom				
Single	5.43 E−10	2.15 E−09	1.17 E−11	6.96 E−11
Double	3.07 E−16	1.11 E−15	6.38 E−17	3.62 E−16
Smooth bottom				
Single	9.53 E−08	3.58 E−07	1.22 E−07	5.38 E−07
Double	3.22 E−16	8.88 E−16	1.31 E−17	6.32 E−16

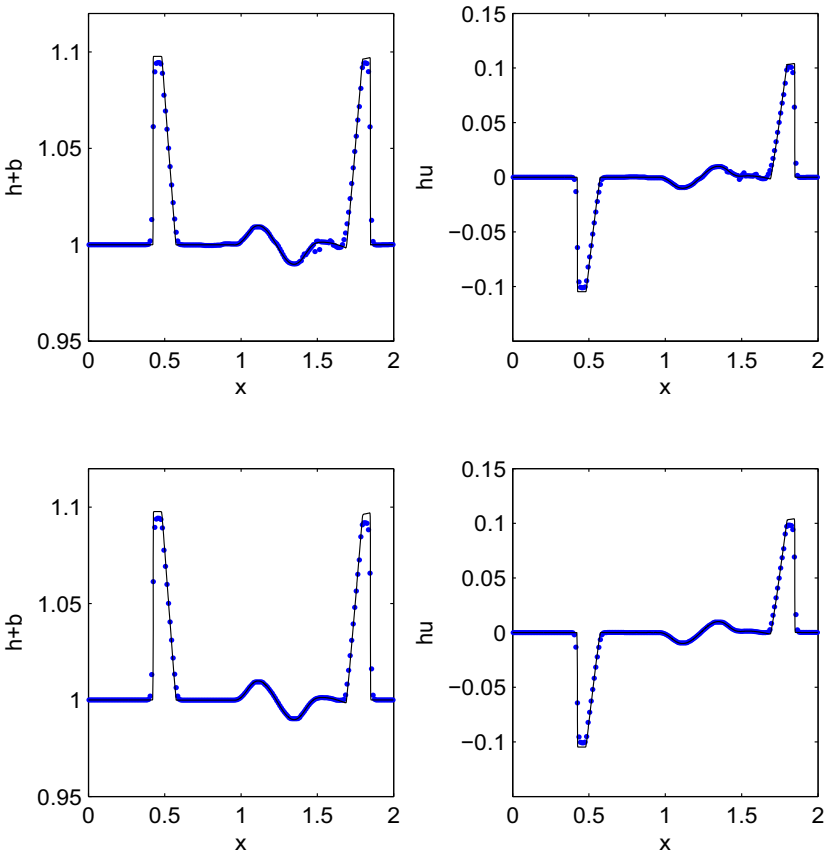


Fig. 5 Numerical results for a stationary solution perturbed with a big pulse at $t = 0.63$. *Dots* 200 elements, *solid line* 2000 elements. *Upper panel* standard CDG method, *lower panel* well-balanced CDG method

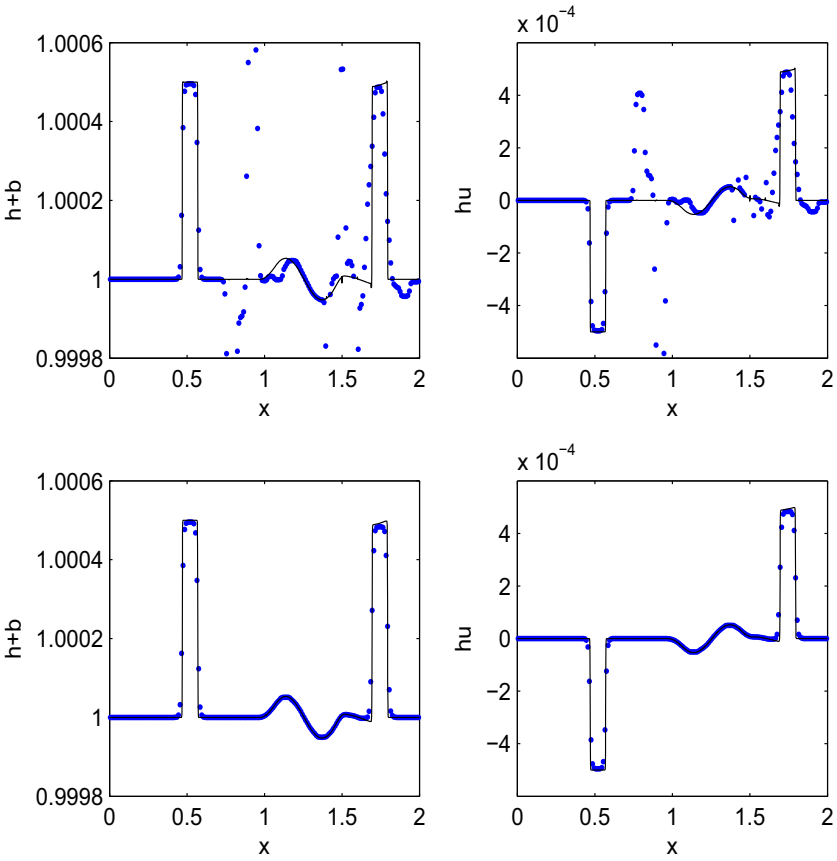


Fig. 6 Numerical results for a stationary solution perturbed with a small pulse at $t = 0.63$. *Dots* 200 elements, *solid line* 2000 elements. *Upper panel* standard CDG method, *lower panel* well-balanced CDG method

We show numerical results by the well-balanced CDG method with 200 elements at times $t = 15$ and $t = 60$ in Figs. 7 and 8, respectively. A better resolved solution (by using 2000 elements) is also plotted for comparison. We can see that both solutions agree well with each other and do not exhibit spurious oscillations. Numerical results by the standard CDG method, however, display parasitic waves near jump points of the bottom topography (not shown here).

3.1.6 Subcritical and Transcritical Steady Flows

This example examines the convergence to subcritical steady flow over a hump [13,35,36]. The computational domain together with the boundary and initial conditions are

$$\begin{cases} h(x,0) = 2.0 - b(x), & 0 \le x \le 25, \\ (hu)(x,0) = 1.411, & 0 \le x \le 25, \\ h(25,t) = 2.0, & 0 \le t \le 626.5, \\ (hu)(0,t) = 1.411, & 0 \le t \le 626.5, \end{cases}$$

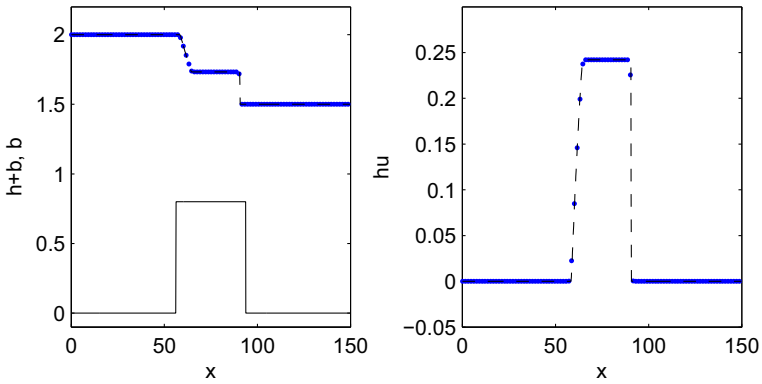


Fig. 7 Numerical results by the well-balanced CDG method for a dam-breaking problem at $t = 15$. *Dots* 200 elements; *dashed line* 2000 elements; *solid line* bottom profile

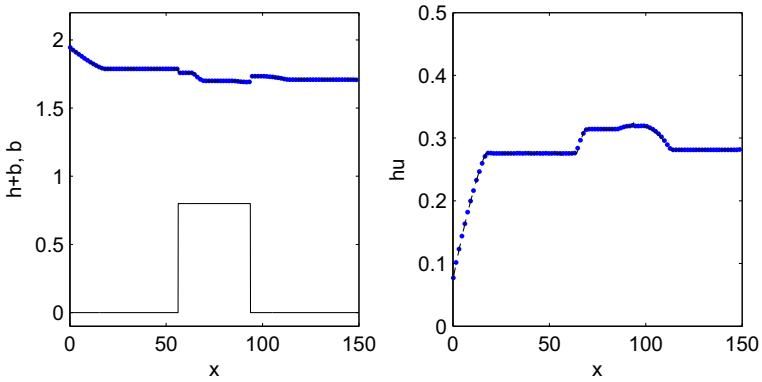


Fig. 8 Numerical results by the well-balanced CDG method for a dam-breaking problem at $t = 60$. *Dots* 200 elements; *dashed line* 2000 elements; *solid line* bottom profile

and the bottom topography is given by

$$b(x) = \begin{cases} 0.2 - 0.05(x-10)^2, & 8 \le x \le 12, \\ 0, & \text{otherwise.} \end{cases} \tag{49}$$

Two different meshes are specified (with 200 and 1000 elements) and this problem is simulated up to $t = 626.5$ using the well-balanced CDG method. The surface level $h + b$ and discharge hu are illustrated in Fig. 9. The lower-resolution computation agrees overall well with the converged (higher-resolution) computation. Spurious oscillations for the discharge on the coarse mesh are found to develop near non-smooth points of the surface level. This is consistent with previous observations [13,35,36] indicating that it is more difficult to accurately capture the discharge than the surface level. Our results, however, are quite acceptable and we checked that errors decrease as the mesh is refined.

Another similar test addresses the problem of transcritical steady flow with a shock, as also investigated in [13,35,36]. The computational domain together with the boundary and initial conditions are

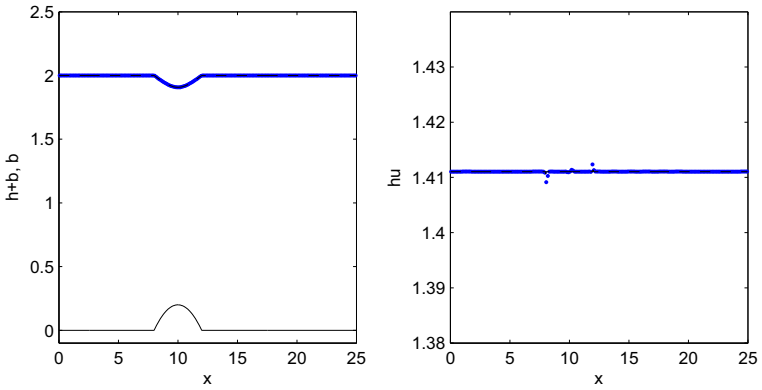


Fig. 9 Numerical results by the well-balanced CDG method for a subcritical steady flow at $t = 626.5$. *Dots* 200 elements; *dashed line* 1000 elements; *solid line* bottom profile

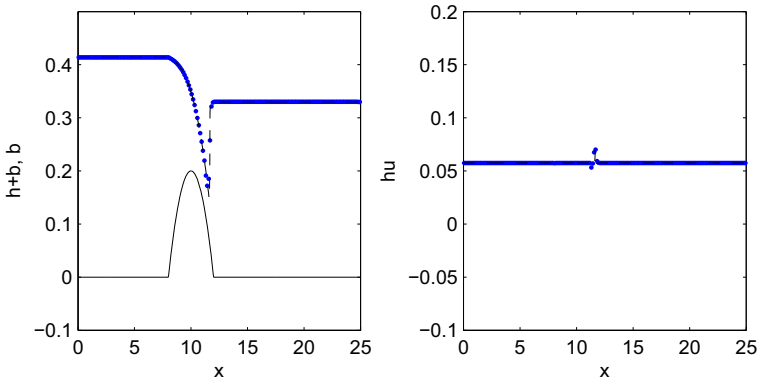


Fig. 10 Numerical results by the well-balanced CDG method for a transcritical steady flow at $t = 626.5$. *Dots* 200 elements; *dashed line* 1000 elements; *solid line* bottom profile

$$\begin{cases} h(x,0) = 0.33, \quad (hu)(x,0) = 0.0575, \quad 0 \le x \le 25, \\ h(25,t) = 0.33, \quad (hu)(0,t) = 0.0575, \quad 0 \le t \le 626.5. \end{cases}$$

The bottom topography is the same as prescribed in (49). We again use 200 and 1000 elements, and simulate this problem up to $t = 626.5$ by the well-balanced CDG method. The surface level $h + b$ and discharge hu are depicted in Fig. 10. Observations similar to those for the subcritical steady flow can be made. Overall, the solution convergence is well illustrated and, in particular, the abrupt depression followed by a shock downstream of the hump is well captured by the numerical scheme.

3.1.7 Drying and Wetting Riemann Problems

In this example, we consider a Riemann problem where a dry area emerges during the simulation. The initial conditions are given by

$$(h(x,0), u(x,0)) = \begin{cases} (h^- = 0.5,\ u^- = 0), \text{ if } x < 0, \\ (h^+ = 1, \quad u^+ = 4), \text{ if } x > 0, \end{cases} \tag{50}$$

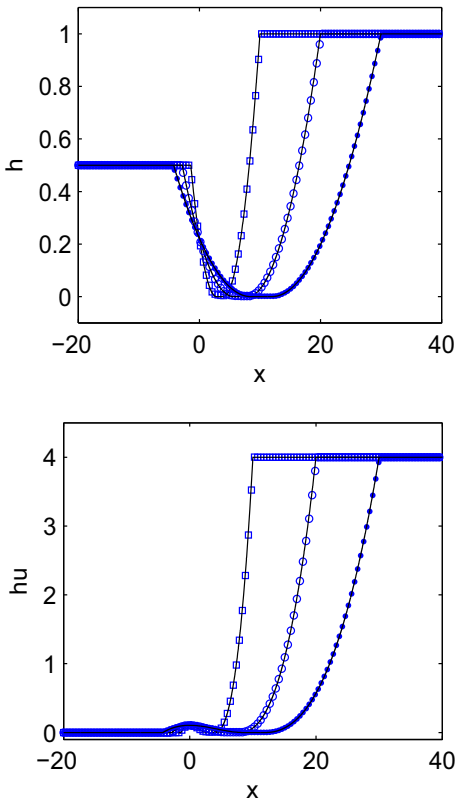


Fig. 11 Numerical results at $t = 2, 4, 6$ obtained from the positivity-preserving well-balanced CDG method for the drying Riemann problem. *Solid line* exact solution; *squares* numerical solution at $t = 2$; *circles* numerical solution at $t = 4$; *dots* numerical solution at $t = 6$

and the bottom topography is $b = 0$. This problem has been examined in [4,38] and an exact solution can be found in [3]. The initial conditions do not contain dry areas, but satisfy the drying criterion $(u^- + 2\sqrt{gh^-}) - (u^+ - 2\sqrt{gh^+}) < 0$. Two expansion waves propagate away from each other, and thus a dry area emerges when $t > 0$.

The computational domain $[-20, 40]$ is divided into 1000 cells, with outgoing boundary conditions. We ran this test using the standard and positivity-preserving CDG methods. For the standard CDG method, negative water depth was generated during the computation and this causes the numerical solution to shortly blow up. For the positivity-preserving CDG method however, the water depth remains non-negative during the entire computation.

Numerical results at times $t = 2, 4, 6$ obtained from the positivity-preserving CDG method are shown in Fig. 11. The exact solution is also plotted in this figure for comparison. To check the non-negativity of the water depth, a zoom-in near the wet/dry front is presented in Fig. 12. We can see that the numerical approximation matches the exact solution well, and the positivity-preserving CDG scheme successfully maintains the non-negativity of the water depth.

Similarly, we also tested our positivity-preserving CDG scheme in the context of the dam-breaking problem over a flat dry bed. The initial conditions are given by

$$(h, u) = \begin{cases} (h^- = 1,\ u^- = 0), \text{ if } x < 0, \\ (h^+ = 0,\ u^+ = 0), \text{ if } x > 0, \end{cases} \tag{51}$$

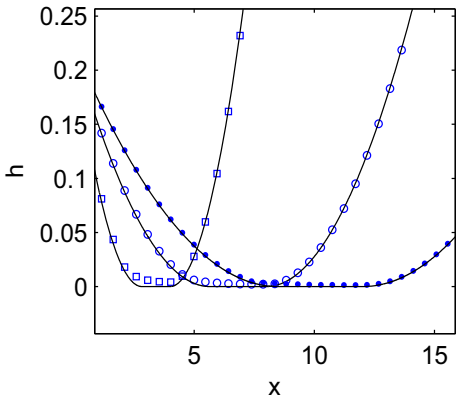
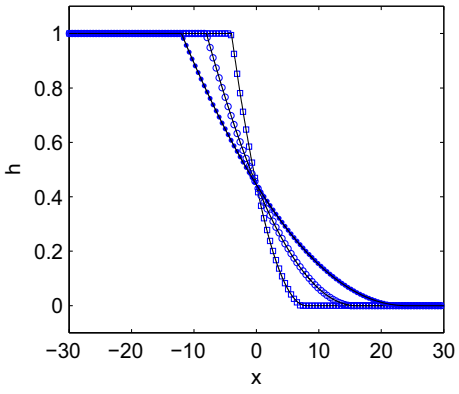


Fig. 12 Zoom-in of the water depth for the drying Riemann problem. *Solid line* exact solution; *squares* numerical solution at $t = 2$; *circles* numerical solution at $t = 4$; *dots* numerical solution at $t = 6$



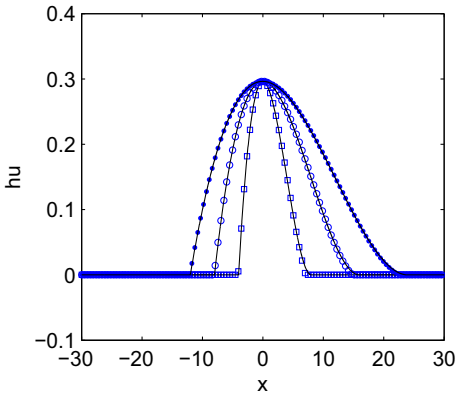


Fig. 13 Numerical results at $t = 4, 8, 12$ obtained from the positivity-preserving well-balanced CDG method for the dam-breaking problem over a flat dry bed. *Solid line* exact solution; *squares* numerical solution at $t = 4$; *circles* numerical solution at $t = 8$; *dots* numerical solution at $t = 12$

and further details (e.g. exact solution) can be found in [3,4,38]. This situation is opposite to the previous one since a wet front now moves into a dry area. Here again, Figs. 13 and 14 indicate that the numerical results compare well with the exact solution and no negative water depth is observed during the entire simulation. Standard CDG methods would produce negative values of the water depth and, combined with the nonlinear TVB minmod limiter in local characteristic fields (which requires calculating the eigenvalues $u \pm \sqrt{gh}$), they would

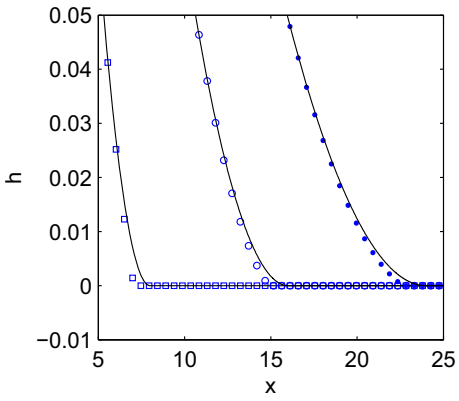


Fig. 14 Zoom-in of the water depth for the dam-breaking problem over a flat dry bed. *Solid line* exact solution; *squares* numerical solution at $t = 4$; *circles* numerical solution at $t = 8$; *dots* numerical solution at $t = 12$

lead to solution blow-up. If we instead replace h by $\max(h, 0)$, the numerical scheme yields satisfactory results yet it is no longer mass-conservative.

3.1.8 Parabolic Bowl

We consider a plane surface oscillating in the parabolic bowl given by

$$b(x) = h_0 \left(\frac{x}{a}\right)^2,$$

with constants h_0 and a to be specified below. The corresponding exact water surface for the 1D NSWEs reads

$$h(x,t) + b(x) = \max\left\{b, h_0 - \frac{B^2}{4g}\left(\cos(2\omega t) + 1\right) - \frac{Bx}{2a}\sqrt{\frac{8h_0}{g}}\cos(\omega t)\right\}, \tag{52}$$

where $\omega = \sqrt{2gh_0}/a$ and B is a given constant [38].

The computational domain is $[-2, 2]$ with outgoing boundary conditions, and is discretized with 200 elements. The exact solution (52) at $t = 0$ serves as the initial condition for the water surface level, and the initial discharge hu is set to zero. We choose the constants to be $h_0 = 0.1$, $a = 1.0$ and $B = 0.3$. Figure 15 plots the water surface level and corresponding discharge at $t = 2T, 2T + T/6, 2T + T/5, 2T + T/4, 2T + T/3, 2T + T/2$ where $T = 2\pi/\omega$ is the oscillation period. The positivity-preserving well-balanced CDG method is applied to solving this problem. Comparison between the numerical and analytical results for $h + b$ shows excellent agreement. No spurious oscillations or negative values are discernible in the numerical solution at any time.

3.1.9 Runup of Solitary Waves on a Slope

In this example, we investigate the runup of solitary waves on a mild slope of 1:19.85 following the experimental work of Synolakis [32]. The bottom slope is specified by

$$b(x) = \begin{cases} -x\tan\beta, & x \le \cot\beta, \\ -1, & x > \cot\beta, \end{cases} \tag{53}$$

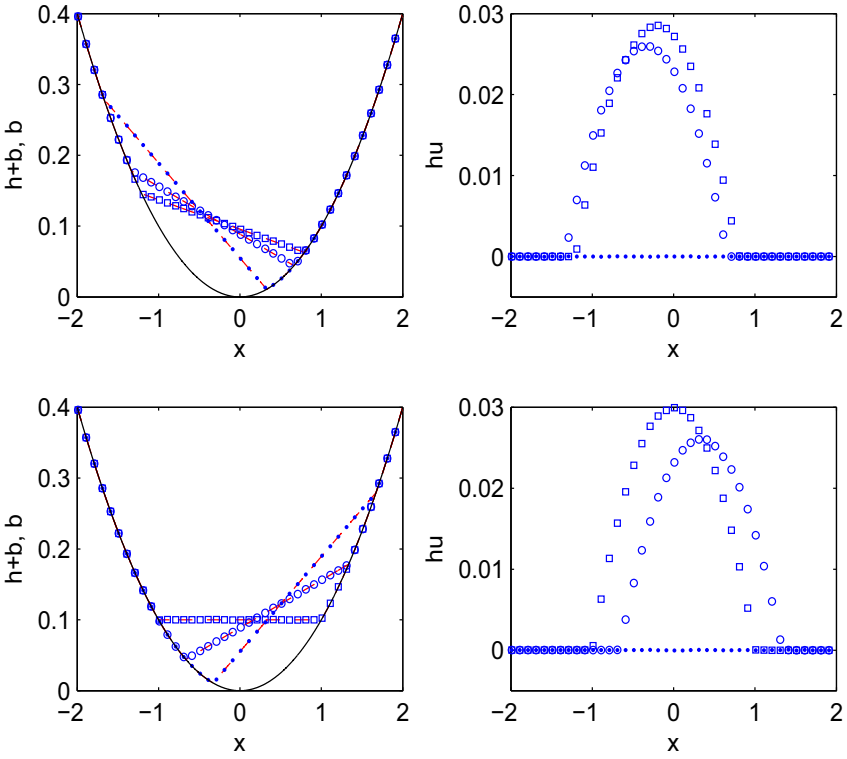


Fig. 15 Numerical results for a plane surface oscillating in a parabolic bowl at $t = 2T$ (*dots*), $2T + T/6$ (*circles*), $2T + T/5$ (*squares*) in the *top panel*, and at $t = 2T + T/4$ (*squares*), $2T + T/3$ (*circles*), $2T + T/2$ (*dots*) in the *lower panel*. *Dashed line* exact water surface; *solid line* domain geometry

with $\beta = 2.884^\circ$ and the initial solitary wave is given by

$$h(x,t) + b(x) = (h_2 - h_1)\text{sech}^2\left(\frac{x - Dt}{2}\sqrt{\frac{3(h_2 - h_1)}{h_2 h_1^2}}\right), \tag{54}$$

where h_1 is a reference water depth, h_2 corresponds to the solitary wave height and $D = \sqrt{gh_2}$ is the wave speed.

We first examine the runup of a solitary wave of small amplitude $a_0 = h_2 - h_1 = 0.0185$ with $h_1 = 1$ and $h_2 = 1.0185$ in (54). The wave crest is initially located at $x = 38.5$ and travels to the left up the sloping beach. The computational domain is $[-15, 85]$, divided into 1000 elements and with outgoing boundary conditions.

A sequence of wave profiles is presented in Fig. 16. To illustrate the performance of the NSWEs and of the numerical method, we also plot experimental data from [32] for comparison. We see that the wave profiles computed from the NSWEs match the experimental data well during both runup and rundown phases.

Since the sloping bottom is a linear function of space, its approximation satisfies condition (26) in the L^2 sense, thus we did not observe negative values of the water depth during the entire simulation. To further test the reliability of the new approximations to the bottom

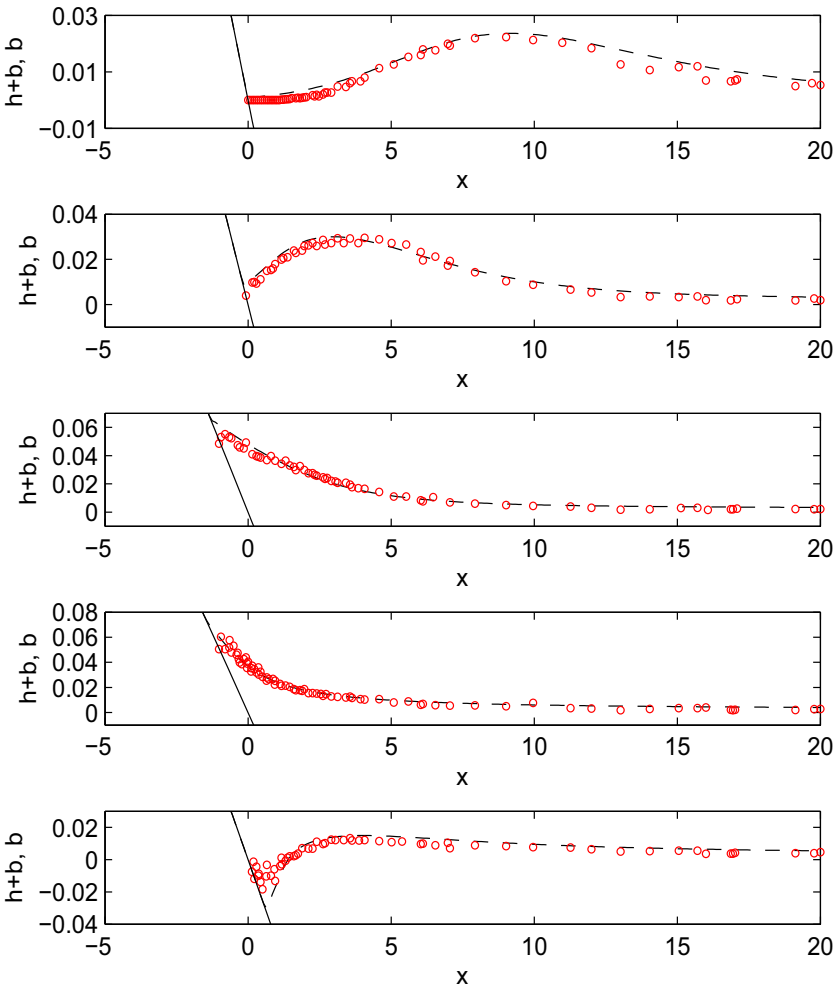


Fig. 16 Runup of a solitary wave with amplitude $a_0 = 0.0185$ at $t = 30, 40, 50, 60, 70$ (from *top* to *bottom*). *Red circles* experimental data [32], *black dashed line* shallow water equations, *black solid line* bottom geometry (Color figure online)

obtained from (28), we investigate the runup of a solitary wave with larger amplitude $a_0 = 0.08$ ($h_1 = 1$ and $h_2 = 1.08$) on a more complicated topography which is a perturbation to the previous slope, as defined by

$$b(x) = \begin{cases} -x \tan\beta + \sin(x)/(x+31), & x \le \cot\beta, \\ -1, & x > \cot\beta. \end{cases} \tag{55}$$

The solitary wave is also initially located at $x = 38.5$ and again travels to the left. We specify the same domain, discretization and boundary conditions as previously. We show a sequence of computed wave profiles in Fig. 17, which confirms that the positivity-preserving well-balanced CDG scheme is effective at maintaining the non-negativity of the water depth, under general wave and topographical conditions.

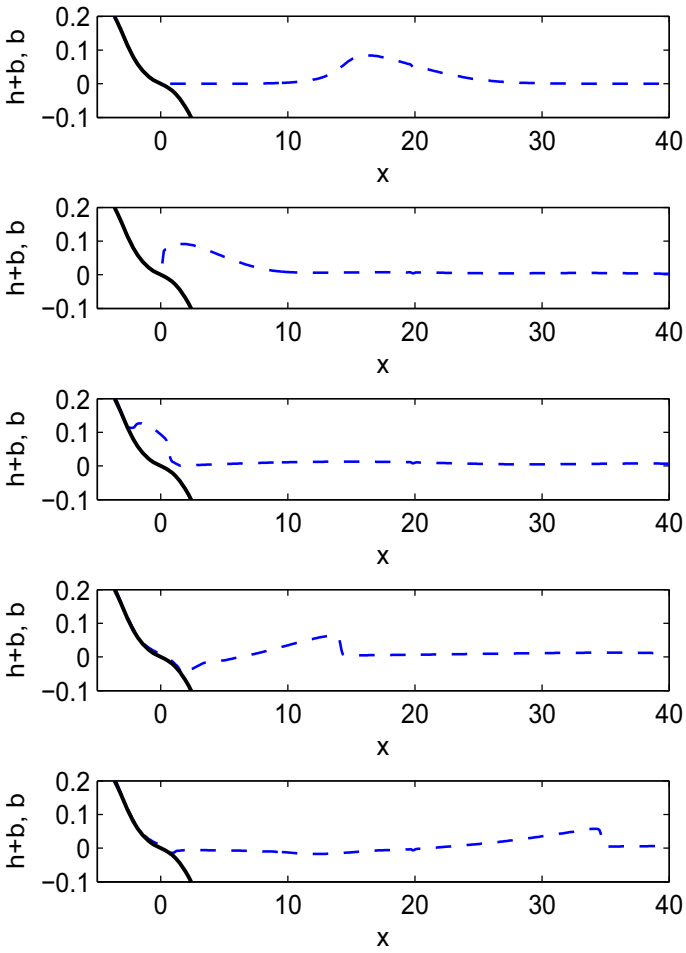


Fig. 17 Runup of a solitary wave with amplitude $a_0 = 0.0185$ at $t = 20, 40, 60, 80, 100$ (from *top* to *bottom*). *Dashed line* computed water surface, *solid line* bottom geometry

3.2 Examples for 2D NSWEs

3.2.1 Accuracy Tests

The first test aims to verify the high-order accuracy of the 2D well-balanced CDG method. We take the smooth functions

$$\begin{aligned} h(x, y, 0) &= 10 + e^{\sin(2\pi x)} \cos(2\pi y), \\ hu(x, y, 0) &= \sin(\cos(2\pi x)) \sin(2\pi y), \\ hv(x, y, 0) &= \cos(2\pi x) \cos(\sin(2\pi y)), \end{aligned}$$

as initial conditions, and the bottom profile is described by

$$b(x, y) = \sin(2\pi x) + \cos(2\pi y).$$

Table 5 L^2 errors and orders of accuracy for (h, hu, hv) at $t = 0.1$ by the standard CDG method

Mesh	h		hu		hv	
	L^2 error	Order	L^2 error	Order	L^2 error	Order
P^1						
10×10	2.58E−02	–	4.88E−02	–	6.98E−02	–
20×20	6.92E−03	2.04	1.15E−02	2.08	1.81E−02	1.95
40×40	1.55E−03	2.02	2.70E−03	2.10	4.51E−03	2.01
80×80	3.87E−04	2.01	6.53E−04	2.05	1.11E−03	2.02
160×160	9.42E−05	2.04	1.58E−04	2.05	2.71E−04	2.04
P^2						
10×10	2.94E−03	–	5.89E−03	–	6.47E−03	–
20×20	3.78E−04	2.96	7.23E−04	3.03	8.52E−04	2.92
40×40	4.66E−05	3.02	8.88E−05	3.02	1.06E−04	3.00
80×80	5.44E−06	3.10	1.05E−05	3.08	1.25E−05	3.09
160×160	5.75E−07	3.24	1.13E−06	3.23	1.34E−05	3.22

Table 6 L^2 errors and orders of accuracy for (h, hu, hv) at $t = 0.1$ by the well-balanced CDG method

Mesh	h		hu		hv	
	L^2 error	Order	L^2 error	Order	L^2 error	Order
P^1						
10×10	2.68E−02	–	4.45E−02	–	6.19E−02	–
20×20	6.54E−03	2.03	1.02E−02	2.11	1.59E−02	1.96
40×40	1.60E−03	2.02	2.36E−03	2.11	3.94E−03	2.01
80×80	3.95E−04	2.01	5.68E−04	2.05	9.73E−04	2.02
160×160	9.58E−05	2.04	1.36E−04	2.05	2.35E−04	2.04
P^2						
10×10	2.93E−03	–	5.88E−03	–	6.44E−03	–
20×20	3.76E−04	2.96	7.22E−04	3.03	8.50E−04	2.92
40×40	4.64E−05	3.02	8.87E−05	3.02	1.06E−04	3.00
80×80	5.41E−06	3.10	1.05E−05	3.08	1.25E−05	3.09
160×160	5.75E−07	3.23	1.12E−06	3.23	1.34E−05	3.22

The computational domain is $[0, 1] \times [0, 1]$ with periodic boundary conditions. The final time is $t = 0.1$ at which the solution is still smooth. Since the exact solution is unknown for this problem, we first compute the numerical solution on a fine mesh with 320×320 elements, and then use it as the reference solution to evaluate errors and orders of accuracy for coarser resolutions. We present L^2 errors and orders of accuracy for h, hu and hv obtained by the standard CDG method (30) and its well-balanced version (36) in Tables 5 and 6, respectively. We can observe that both methods are $(k + 1)$st-order accurate for P^k with $k = \{1, 2\}$ and therefore the added correction terms in (36) do not affect the accuracy of the standard CDG method.

3.2.2 Stationary Solution

In this test, we validate the well-balanced feature of the 2D method (36) as applied to a continuous and discontinuous variable bottom. The initial conditions are

$$hu(x, y, 0) = 0, \qquad hv(x, y, 0) = 0, \qquad h(x, y, 0) + b(x, y) = 1, \tag{56}$$

and the continuous bottom profile is defined by

$$b(x, y) = \begin{cases} 0.2, & r \leq 0.3, \\ 0.5 - r, & 0.3 \leq r \leq 0.5, \\ 0, & \text{otherwise}, \end{cases} \tag{57}$$

with $r = \sqrt{x^2 + y^2}$, while the discontinuous bottom profile is given by

$$b(x, y) = \begin{cases} 0.5, & -0.5 \leq x, y \leq 0.5, \\ 0, & \text{otherwise}. \end{cases} \tag{58}$$

We choose $[-1, 1] \times [-1, 1]$ as the computational domain, divided into 50×50 elements, and use outgoing boundary conditions. We compute the solution up to $t = 1$ by the standard and well-balanced CDG methods.

Contours of the computed water level $h + b$, discharges hu and hv are plotted in Fig. 18 for the continuous variable bottom and in Fig. 19 for the discontinuous variable bottom. We see that the well-balanced CDG method maintains the stationary solution (56) in both cases, while the standard CDG method fails to do so. To demonstrate that the well-balanced scheme indeed preserves the still-water stationary solution exactly (i.e. up to machine precision), we

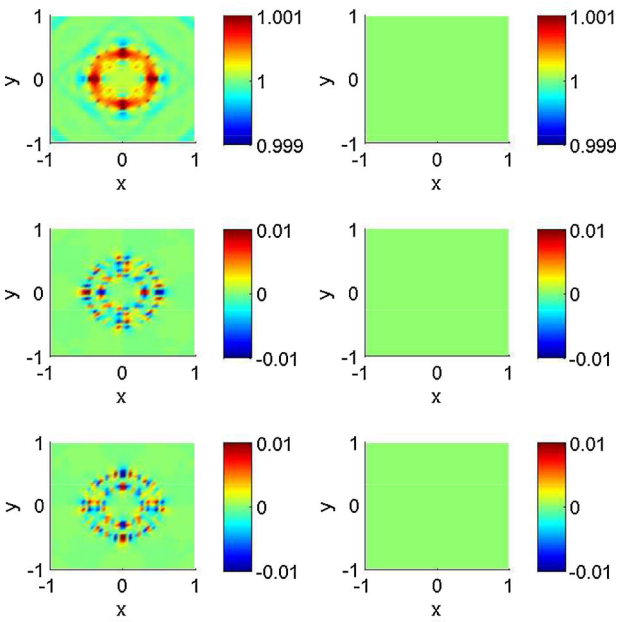


Fig. 18 Numerical results for the stationary solution over a 2D continuous variable bottom at $t = 1$. *Left* standard CDG method; *right* well-balanced CDG method. *Upper panel* water level $h + b$; *middle panel* discharge hu; *lower panel* discharge hv

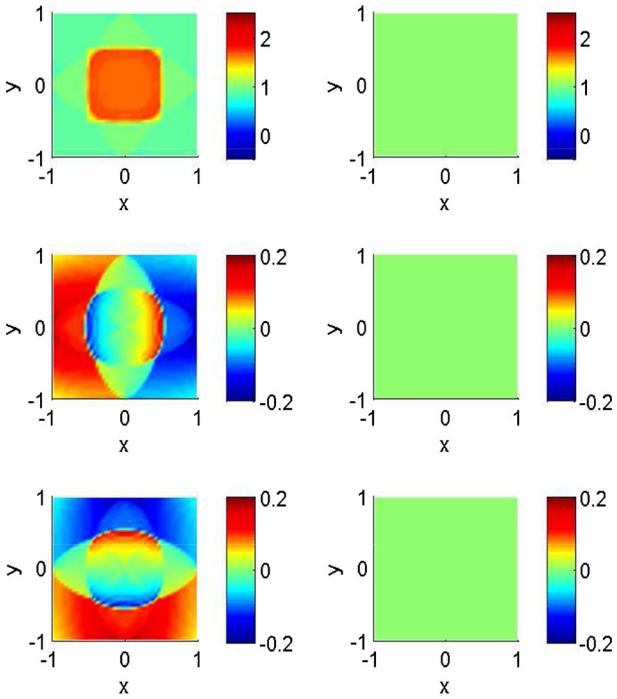


Fig. 19 Numerical results for the stationary solution over a 2D discontinuous variable bottom at $t = 1$. *Left* standard CDG method; *right* well-balanced CDG method. *Upper panel* water level $h + b$; *middle panel* discharge hu; *lower panel* discharge hv

Table 7 L^2 and L^∞ errors on $(h + b, hu, hv)$ for the stationary solution at $t = 1$

Precision	$h + b$		hu		hu	
	L^2 error	L^∞ error	L^2 error	L^∞ error	L^2 error	L^∞ error
Continuous bottom						
Single	1.19E−07	1.19E−07	9.25E−09	3.01E−08	9.78E−09	3.50E−08
Double	8.85E−16	1.67E−15	1.79E−15	7.69E−16	2.24E−16	8.42E−16
Discontinuous bottom						
Single	1.19E−07	2.38E−07	1.40E−08	8.30E−08	1.42E−07	8.43E−07
Double	8.95E−16	2.00E−15	2.67E−16	1.38E−15	2.49E−16	1.03E−15

performed the computation in both single and double precision. The corresponding L^2 and L^∞ errors on the water surface $h + b$, discharges hu and hv are listed in Table 7 for both topographies. Their values have orders of magnitude consistent with machine single and double precision, thus verifying the well-balanced property.

3.2.3 Perturbation of a Stationary Solution

We next consider the problem of 2D perturbations to a stationary state [20]. The computational domain is $[0, 2] \times [0, 1]$ and the bottom topography is an isolated elliptical hump defined by

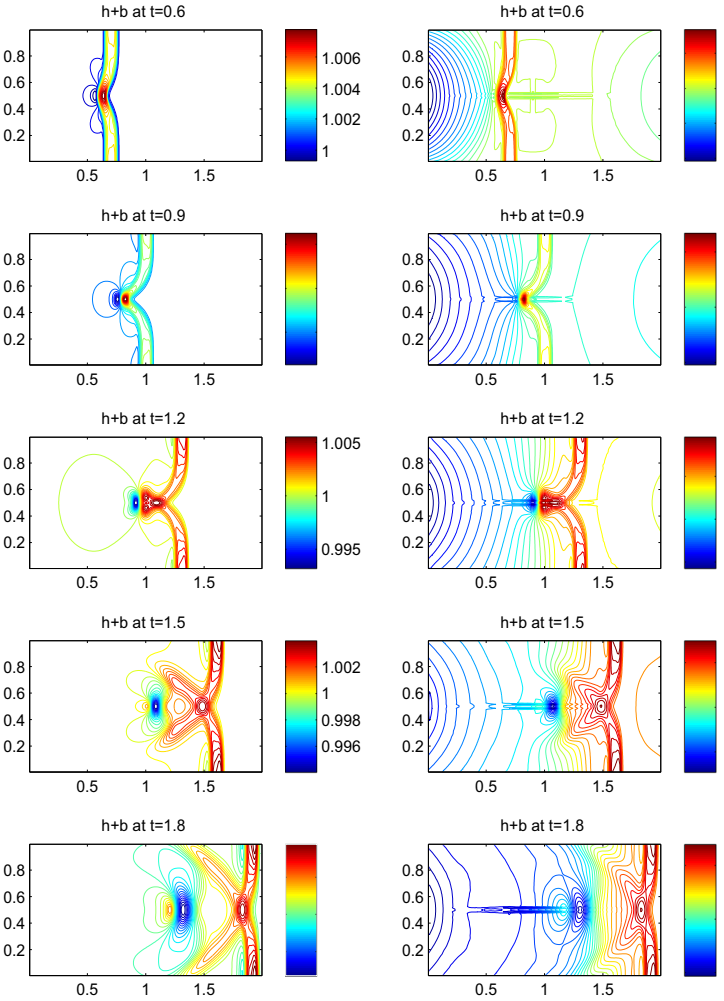


Fig. 20 Contours of the surface level $h+b$ for the perturbation of a stationary solution at $t = 0.6, 0.9, 1.2, 1.5$ and 1.8 (*from top to bottom*). *Left* well-balanced CDG method, *right* standard CDG method

$$b(x, y) = 0.8\, e^{-5(x-0.9)^2-50(y-0.5)^2}.$$

The initial conditions are given by

$$h(x, y, 0) = \begin{cases} 1 - b(x, y) + \varepsilon, & 0.05 \le x \le 0.15, \\ 1 - b(x, y), & \text{otherwise}, \end{cases}$$
$$u(x, y, 0) = v(x, y, 0) = 0,$$

where ε is a non-zero perturbation parameter which is selected to be 0.01 in our computation.

The well-balanced CDG method is used to solve this problem on a 200×100 mesh. The left column in Fig. 20 displays the water surface $h+b$ at $t = 0.6, 0.9, 1.2, 1.5$ and 1.8. Clearly, we can see that the numerical solution is able to capture complex small features of the flow as reported in [18,35]. To demonstrate the importance of the well-balanced property, we also

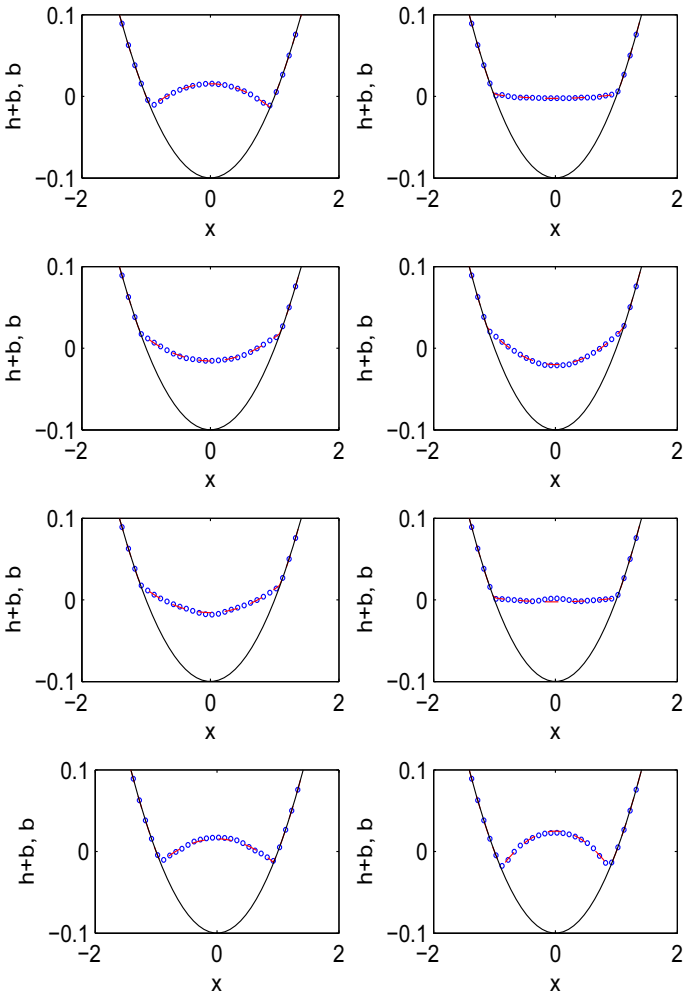


Fig. 21 Vertical cross-section of the water surface level $h + b$ at $y = 0$ in a parabolic bowl at $t = T/8, T/4, 3T/8, T/2, 5T/8, 3T/4, 7T/8$ and T (from *left* to *right* and from *top* to *bottom*). *Circles* numerical solution; *dashed line* exact water surface; *solid line* domain geometry

compare with numerical results by the standard CDG method for the same 200×100 mesh, as presented in the right column of Fig. 20. The corresponding solution tends to produce parasitic waves which are of about the same amplitude as the principal waves generated by the small bottom perturbation.

3.2.4 Parabolic Bowl

Finally, we examine a water surface oscillating in a 2D parabolic bowl embedded in a square domain $[-2, 2] \times [-2, 2]$. The parabolic geometry is defined by

$$b(x, y) = h_0 \left(\frac{x^2 + y^2}{a^2} - 1 \right),$$

where h_0 and a are constants to be specified later. The exact solution for the 2D NSWEs is given by

$$\eta(x, y, t) = \max\left\{b, h_0\left[\frac{\sqrt{1-A^2}}{1-A\cos(\omega t)} - 1 - \frac{x^2+y^2}{a^2}\left(\frac{1-A^2}{(1-A\cos(\omega t))^2} - 1\right)\right]\right\},$$
$$u(x, y, t) = \frac{1}{1-A\cos(\omega t)}\left(\frac{1}{2}\omega x \sin(\omega t)\right), \quad \text{if} \quad h > 0,$$
$$v(x, y, t) = \frac{1}{1-A\cos(\omega t)}\left(\frac{1}{2}\omega y \sin(\omega t)\right), \quad \text{if} \quad h > 0, \tag{59}$$

where $\eta = h + b$, $\omega = \sqrt{8gh_0}/a$, $A = (a^2 - r_0^2)/(a^2 + r_0^2)$, and r_0 is a given constant representing the distance from the edge to the center of the water surface at the initial time $t = 0$ [29].

For this experiment, the computational domain $[-2, 2] \times [-2, 2]$ is divided into 200×200 elements. The exact solution (59) at $t = 0$ serves as the initial condition, and outgoing boundary conditions are specified. We choose $h_0 = 0.1$, $a = 1.0$ and $r_0 = 0.8$. Similar to the 1D case, since the near-boundary regions are always dry, we apply the positivity-preserving well-balanced CDG method to this problem. We plot the water surface level at $t = T/8, T/4, 3T/8, T/2, 5T/8, 3T/4, 7T/8$ and T in Fig. 21, where $T = 2\pi/\omega$ is the oscillation period. Excellent agreement between the numerical and analytical results is found and, in particular, no negative values of the water depth are discernible. However, these would appear and the computation would break down if the positivity-preserving limiter were not used.

4 Conclusions

In this paper, we have proposed a family of high-order numerical methods for 1D and 2D NSWEs, including well-balanced CDG methods, positivity-preserving CDG methods and their combined versions. Our numerical tests on various 1D and 2D problems demonstrate the properties of high-order accuracy, positivity-preservation and well-balance of the proposed schemes. Designing CDG-based methods with such features for simulations of 2D NSWEs on unstructured meshes is envisioned for future work. It is also planned to extend the proposed methods to solving 2D Green–Naghdi models for water waves over bottom topography. These equations are more complicated than the NSWEs and their numerical simulation poses major technical challenges [25,40].

Acknowledgements ML is partially supported by the Fundamental Research Funds for the Central Universities in China (Project No. 106112015CDJXY100008) and the NSFC (Grant No. 11501062). PG is partially supported by the NSF (Grant No. DMS-1615480) and the Simons Foundation (Grant No. 246170). FL is partially supported by the NSF (Grant No. DMS-1318409). LX is partially supported by the NSFC (Grant No. 11371385), the start-up fund of Youth 1000 plan of China and that of Youth 100 plan of Chongqing University.

References

1. Bermudez, A., Vazquez, M.E.: Upwind methods for hyperbolic conservation laws with source terms. Comput. Fluids **23**, 1049–1071 (1994)
2. Berthon, C., Marche, F.: A positive preserving high order VFRoe scheme for shallow water equations: a class of relaxation schemes. SIAM J. Sci. Comput. **30**, 2587–2612 (2008)

3. Bokhove, O.: Flooding and drying in discontinuous Galerkin finite-element discretizations of shallow-water equations. Part 1: one dimension. J. Sci. Comput. **22**, 47–82 (2005)
4. Bunya, S., Kubatko, E.J., Westerink, J.J., Dawson, C.: A wetting and drying treatment for the Runge–Kutta discontinuous Galerkin solution to the shallow water equations. Comput. Methods Appl. Mech. Eng. **198**, 1548–1562 (2009)
5. Brufau, P., Vazquez-Cendon, M.E., Garcia-Navarro, P.: A numerical model for the flooding and drying of irregular domains. Int. J. Numer. Methods Fluids **39**, 247–275 (2002)
6. Bryson, S., Epshteyn, Y., Kurganov, A., Petrova, G.: Well-balanced positivity preserving central-upwind scheme on triangular grids for the Saint-venant system. Math. Model. Numer. Anal. **45**, 423–446 (2011)
7. Castro, M.J., Gonzalez-Vida, J.M., Pares, C.: Numerical treatment of wet/dry fronts in shallow flows with a modified Roe scheme. Math. Models Methods Appl. Sci. **16**, 897–931 (2006)
8. Cheng, Y., Li, F., Qiu, J., Xu, L.: Positivity-preserving DG and central DG methods for ideal MHD equations. J. Comput. Phys. **238**, 255–280 (2013)
9. Cockburn, B., Shu, C.-W.: The Runge–Kutta discontinuous Galerkin method for conservation laws V multidimensional systems. J. Comput. Phys. **141**, 199–224 (1998)
10. Dong, H., Li, M.: A reconstructed central discontinuous Galerkin-finite element method for the fully nonlinear weakly dispersive Green-Naghdi model. Appl. Numer. Math. **110**, 110–127 (2016)
11. Ern, A., Piperno, S., Djadel, K.: A well-balanced Runge–Kutta discontinuous Galerkin method for the shallow-water equations with flooding and drying. Int. J. Numer. Methods Fluids **58**, 1–25 (2008)
12. Gallardo, J.M., Pares, C., Castro, M.: On a well-balanced high-order finite volume scheme for shallow water equations with topography and dry areas. J. Comput. Phys. **227**, 574–601 (2007)
13. George, D.L.: Augmented Riemann solvers for the shallow water equations over variable topography with steady states inundation. J. Comput. Phys. **227**, 3089–3113 (2008)
14. Gottlieb, S., Shu, C.-W., Tadmor, E.: Strong stability preserving high order time discretization methods. SIAM Rev. **43**, 89–112 (2001)
15. Greenberg, J.M., LeRoux, A.Y.: A well-balanced scheme for the numerical processing of source terms in hyperbolic equations. SIAM J. Numer. Anal. **33**, 1–16 (1996)
16. Guyenne, P., Nicholls, D.P.: A high-order spectral method for nonlinear water waves over moving bottom topography. SIAM J. Sci. Comput. **30**, 81–101 (2007)
17. Hubbard, M.E., Garcia-Navarro, P.: Flux difference splitting and the balancing of source terms and flux gradients. J. Comput. Phys. **165**, 89–125 (2000)
18. Kurganov, A., Levy, D.: Central-upwind schemes for the Saint-Venant system. ESAIM Math. Mod. Numer. Anal. **36**, 397–425 (2002)
19. Kurganov, A., Petrova, G.: A second-order well-balanced positivity preserving scheme for the Saint-Venant system. Commun. Math. Sci. **5**, 133–160 (2007)
20. LeVeque, R.J.: Balancing source terms and flux gradients on high-resolution Godunov methods: the quasi-steady wave-propagation algorithm. J. Comput. Phys. **146**, 346–365 (1998)
21. Levy, D., Puppo, G., Russo, G.: Central WENO schemes for hyperbolic systems of conservation laws. ESAIM Math. Model. Numer. Anal. **33**, 547–571 (1999)
22. Li, F., Yakovlev, S.: A central discontinuous Galerkin method for Hamilton–Jacobi equations. J. Sci. Comput. **45**, 404–428 (2010)
23. Li, F., Xu, L.: Arbitrary order exactly divergence-free central discontinuous Galerkin methods for ideal MHD equations. J. Comput. Phys. **231**, 2655–2675 (2012)
24. Li, F., Xu, L., Yakovlev, S.: Central discontinuous Galerkin methods for ideal MHD equations with exactly divergence-free magnetic field. J. Comput. Phys. **230**, 4828–4847 (2011)
25. Li, M., Guyenne, P., Li, F., Xu, L.: High order well-balanced CDG-FE methods for shallow water waves by a Green–Naghdi model. J. Comput. Phys. **257**, 169–192 (2014)
26. Liang, Q., Marche, F.: Numerical resolution of well-balanced shallow water equations with complex source terms. Adv. Water Resour. **32**, 873–884 (2009)
27. Liu, Y., Shu, C.-W., Tadmor, E., Zhang, M.: Central discontinuous Galerkin methods on overlapping cells with a nonoscillatary hierarchical reconstruction. SIAM J. Numer. Anal. **45**, 2442–2467 (2007)
28. Liu, Y., Shu, C.-W., Tadmor, E., Zhang, M.: L^2 stability analysis of the central discontinuous Galerkin method and a comparison between the central and regular discontinuous Galerkin methods. ESAIM Math. Model. Numer. Anal. **42**, 593–607 (2008)
29. Marche, F., Bonneton, P., Fabrie, P., Seguin, N.: Evaluation of well-balanced bore-capturing schemes for 2D wetting and drying processes. Int. J. Numer. Methods Fluids **53**, 867–894 (2007)
30. Rebollo, T.C., Delgado, A.D., Nieto, E.D.F.: A family of stable numerical solvers for the shallow water equations with source terms. Comput. Methods Appl. Mech. Eng. **192**, 203–225 (2003)
31. Rena, M.A., Li, F.: Operator bounds and time step conditions for DG and central DG methods. J. Sci. Comput. **62**, 532–554 (2015)

32. Synolakis, C.E.: The runup of solitary waves. J. Fluid Mech. **185**, 523–545 (1987)
33. Vukovic, S., Sopta, L.: ENO and WENO schemes with the exact conservation property for one-dimensional shallow water equations. J. Comput. Phys. **179**, 593–621 (2002)
34. Xing, Y., Shu, C.-W.: High order finite difference WENO schemes with the exact conservation property for the shallow water equations. J. Comput. Phys. **208**, 206–227 (2005)
35. Xing, Y., Shu, C.-W.: High order well-balanced finite difference WENO schemes for a class of hyperbolic systems with source terms. J. Sci. Comput. **27**, 477–494 (2006)
36. Xing, Y., Shu, C.-W.: High order well-balanced finite volume WENO schemes and discontinuous Galerkin Methods for a class of hyperbolic systems with source terms. J. Comput. Phys. **214**, 567–598 (2006)
37. Xing, Y., Shu, C.-W.: A new approach of high-order well-balanced finite volume WENO schemes and discontinuous Galerkin methods for a class of hyperbolic systems with source terms. Commun. Comput. Phys. **1**, 100–134 (2006)
38. Xing, Y., Zhang, X., Shu, C.-W.: Positivity-preserving high order well-balanced discontinuous Galerkin methods for the shallow water equations. Adv. Water Resour. **33**, 1476–1493 (2010)
39. Xu, K.: A well-balanced gas-kinetic scheme for the shallow-water equations with source terms. J. Comput. Phys. **178**, 533–562 (2002)
40. Xu, L., Guyenne, P.: Numerical simulation of three-dimensional nonlinear water waves. J. Comput. Phys. **228**, 8446–8466 (2009)
41. Zhou, G., Causon, D.M., Mingham, C.G., Ingram, D.M.: The surface gradient method for the treatment of source terms in the shallow-water equations. J. Comput. Phys. **168**, 1–25 (2001)