A high-order well-balanced alternative finite difference WENO (A-WENO) method with the exact conservation property for systems of hyperbolic balance laws*

Ziyao Xu† and Chi-Wang Shu‡

Abstract

In this paper, we develop a high-order well-balanced alternative finite difference weighted essentially non-oscillatory (A-WENO) method with the exact conservation property and high efficiency for a class of hyperbolic balance laws whose steady states are characterized by constant equilibrium variables. In particular, the method preserves the non-hydrostatic equilibria of the shallow water equations with non-flat bottom topography and the Euler equations in gravitational fields. Our method comprises three essential gradients. First, we adopt the finite difference framework to discretize the equations, thus we approximate the value of source terms at grid points rather than their averages on cells. Then, we rewrite the source terms in flux-gradient forms at local reference equilibrium states and discretize them using the same approach as the true flux gradient to achieve the well-balanced property. Most importantly, the exact conservation property and high efficiency are achieved through the interpolation of equilibrium variables in the A-WENO framework [41, 20], which is different from the more widely used finite difference framework [42] based on the reconstruction of fluxes. Since the equilibrium variables are constants at equilibria in the equations we study, the WENO interpolation becomes trivial for the local reference equilibrium states in the flux-gradient formulation of source terms, which is the key to efficiency and preservation of conservation property. Ample numerical tests are presented to validate the good performance of our method.

Key Words: high-order, well-balanced, finite difference, non-hydrostatic equilibria, shallow water equations, Euler equations, exact conservation property

*Research supported by NSF grant DMS-2309249.

†Department of Applied and Computational Mathematics and Statistics, University of Notre Dame, Notre Dame, IN 46556, USA. E-mail: zxu25@nd.edu

‡Division of Applied Mathematics, Brown University, Providence, RI 02912, USA. E-mail: chi-wang_shu@brown.edu

1 Introduction

In this paper, we consider the one-dimensional system of hyperbolic balance laws

$$\mathbf{u}_t + \mathbf{f}(\mathbf{u})_x = \mathbf{s}(\mathbf{u}, x), \quad x \in \mathbb{R},\ t > 0, \tag{1.1}$$

where $\mathbf{u} \in \mathcal{O} \subset \mathbb{R}^m$ is the vector of conserved variables, $\mathbf{f}$ the vector of fluxes and $\mathbf{s}$ the vector of source terms. Throughout this paper, boldface font is used to denote vectors or matrices. If the source term vanishes, the system is called a hyperbolic conservation law.

Hyperbolic balance laws typically admit non-trivial steady-state (equilibrium) solutions, where the source term exactly balances the flux gradient. Numerical schemes that satisfy a discrete analog of this balance are known as well-balanced schemes. It is of great importance to design such schemes, as in many real-world problems, solutions of balance laws are small perturbations from steady states. Schemes that fail to preserve this property can cause spurious oscillations due to truncation errors, which may obscure the actual solution fluctuations from equilibrium, unless the meshes are refined unnecessarily. In contrast, well-balanced schemes produce solutions with excellent resolution near equilibria, even on coarse meshes.

Numerous studies have been conducted on well-balanced schemes for hyperbolic balance laws. The exact C-property of numerical schemes, referring to the hydrostatic (lake-at-rest) equilibria being exactly preserved for the shallow water equations, was introduced and studied by Bermudez and Vazquez in [2]. The strategy of hydrostatic reconstruction was proposed by Audusse et al. in [1], and has been adopted in many subsequent studies, e.g. [50, 30, 53, 33]. Xing and Shu [47] proposed a high-order finite difference WENO scheme for the shallow water equations, based on a reformulation of the separable source term, and generalized the approach in [48, 49]. For other studies on the well-balanced schemes for the shallow water equations with hydrostatic equilibria, we refer to the readers the non-exhaustive list [25, 34, 26, 8, 4, 37, 18] and references therein. The design of well-balanced schemes for the shallow water equations with non-hydrostatic (moving-water) equilibria is more challenging yet highly meaningful [52]. A high-order well-balanced finite volume scheme was developed by Noelle et al. in [35]. Cheng and Kurganov [12] developed a second-order well-balanced scheme based on the reconstruction of equilibrium variables. Well-balanced DG methods for non-hydrostatic equilibria were studied in [46] and [55]. The works of well-balancing via flux globalization can be found in [11, 13, 32]. Well-balanced schemes for the Euler equations with gravitational fields have also been intensively studied since the notion was first proposed by Cargo and LeRoux in [5], see [43, 45, 51, 15, 22, 29, 30, 31, 24, 10] for the hydrostatic isothermal or polytropic equilibria. For the

well-balanced schemes preserving the non-hydrostatic equilibria of Euler equations, we refer the readers to [3, 16, 14]. More information on the history and recent development of well-balanced methods can be found in the recent survey [21].

In this work, we propose a high-order well-balanced finite difference method with exact conservation property for systems of balance laws in the framework of the alternative finite difference WENO (A-WENO) method [41, 20]. The work builds upon the idea of reformulating source terms as flux gradients at local reference equilibrium states. Some existing works along this line can be found in [36, 7, 46, 9, 16].

We start with a reformulation [21] of the equation (1.1),

$$\mathbf{u}_t + \mathbf{f}(\mathbf{u})_x = \frac{\mathbf{s}(\mathbf{u}, x)}{\mathbf{s}(\mathbf{u}_{e,\mathrm{loc}}, x)} \mathbf{f}(\mathbf{u}_{e,\mathrm{loc}})_x, \quad x \in \mathcal{I},\ t > 0, \tag{1.2}$$

where $\mathbf{u}_{e,\mathrm{loc}}$ is a local equilibrium state of the balance law (1.1) satisfying $\mathbf{f}(\mathbf{u}_{e,\mathrm{loc}})_x = \mathbf{s}(\mathbf{u}_{e,\mathrm{loc}}, x)$ for $x \in \mathcal{I}$, and the quotient $\frac{1}{\mathbf{s}(\mathbf{u}_{e,\mathrm{loc}},x)}\mathbf{f}(\mathbf{u}_{e,\mathrm{loc}})_x$ is understood componentwisely and equal to one in the case of $\frac{0}{0}$. We call $\mathbf{u}_{e,\mathrm{loc}}$ a local reference equilibrium state of $\mathbf{u}$ if $\mathbf{u}_{e,\mathrm{loc}} = \mathbf{u}$ when $\mathbf{u}$ is at equilibrium.

Based on the above formulation, we establish our well-balanced semi-discrete finite difference scheme

$$\frac{d\mathbf{u}_j}{dt} + \frac{1}{\Delta x}\left(\hat{\mathbf{f}}_{j+\frac{1}{2}} - \hat{\mathbf{f}}_{j-\frac{1}{2}}\right) = \frac{\mathbf{s}(\mathbf{u}_j, x_j)}{\mathbf{s}(\mathbf{u}_{e,j}(x_j), x_j)} \frac{1}{\Delta x}\left(\hat{\mathbf{f}}^e_{j,j+\frac{1}{2}} - \hat{\mathbf{f}}^e_{j,j-\frac{1}{2}}\right), \quad j = 0, \pm 1, \ldots, \tag{1.3}$$

where $\mathbf{u}_j$ is the numerical solution to approximate $\mathbf{u}$ at the node $x_j := j\Delta x$, $\hat{\mathbf{f}}_{j+\frac{1}{2}} = \mathcal{F}(\mathbf{u}_{j-r+1}, \ldots, \mathbf{u}_{j+s})$ is the numerical flux to approximate $\mathbf{f}(\mathbf{u})$ at the node $x_{j+\frac{1}{2}} := (j+\frac{1}{2})\Delta x$, $\mathcal{F}$ is a Lipschitz continuous function, the right-hand side is an approximation of the reformulated source term $\frac{\mathbf{s}(\mathbf{u},x)}{\mathbf{s}(\mathbf{u}_{e,j},x)}\mathbf{f}(\mathbf{u}_{e,j})_x$ at x_j, $\mathbf{u}_{e,j}$ is a local reference equilibrium state in a neighborhood of x_j whose definition will be given in later sections for the shallow water equations and Euler equations. To achieve the well-balanced property, we discretize the "flux" gradient in the reformulated source term using the same approach as the true flux gradient, i.e. $\hat{\mathbf{f}}^e_{j,j+\frac{1}{2}} = \mathcal{F}(\mathbf{u}_{e,j}(x_{j-r+1}), \ldots, \mathbf{u}_{e,j}(x_{j+s}))$ and $\hat{\mathbf{f}}^e_{j,j-\frac{1}{2}} = \mathcal{F}(\mathbf{u}_{e,j}(x_{j-r}), \ldots, \mathbf{u}_{e,j}(x_{j+s-1}))$. It is not difficult to see that, if the solution is at equilibrium, i.e. $\mathbf{u}_j = \mathbf{u}_e(x_j)$ for some equilibrium states $\mathbf{u}_e$, the right-hand side of (1.3) becomes $\frac{1}{\Delta x}\left(\hat{\mathbf{f}}^e_{j,j+\frac{1}{2}} - \hat{\mathbf{f}}^e_{j,j-\frac{1}{2}}\right)$ due to $\mathbf{u}_{e,j} = \mathbf{u}_e$, which exactly balances the flux discretization on the left-hand side by the definition.

Quite a similar strategy was used in [36], where the local reference equilibrium state $\mathbf{u}_{e,j}$ was taken as the solution of the ODE problem

$$\begin{cases} \mathbf{f}\left(\mathbf{u}_{e,j}\right)_x = \mathbf{s}\left(\mathbf{u}_{e,j}, x\right), \\ \mathbf{u}_{e,j}(x_j) = \mathbf{u}_j. \end{cases} \tag{1.4}$$

If such $\mathbf{u}_{e,j}$ exists in the neighborhood of x_j we shall use, the balance law (1.1) has a simplified reformulation

$$\mathbf{u}_t + \left(\mathbf{f}(\mathbf{u}) - \mathbf{f}(\mathbf{u}_{e,j})\right)_x = \mathbf{0},$$

and consequently, the semi-discrete finite difference scheme

$$\frac{d\mathbf{u}_j}{dt} + \frac{1}{\Delta x}\left(\hat{\mathbf{f}}_{j,j+\frac{1}{2}} - \hat{\mathbf{f}}_{j,j-\frac{1}{2}}\right) = \mathbf{0}, \quad j = 0, \pm 1, \ldots, \tag{1.5}$$

is well-balanced, where $\hat{\mathbf{f}}_{j,j+\frac{1}{2}}$ and $\hat{\mathbf{f}}_{j,j-\frac{1}{2}}$ are numerical fluxes to approximate $\mathbf{f}(\mathbf{u}) - \mathbf{f}(\mathbf{u}_{e,j})$ at $x_{j+\frac{1}{2}}$ and $x_{j-\frac{1}{2}}$, respectively, based on the WENO flux reconstruction [42, 19].

However, (1.5) is generally not conservative because of $\hat{\mathbf{f}}_{j,j+\frac{1}{2}} \neq \hat{\mathbf{f}}_{j+1,j+\frac{1}{2}}$, meaning the subsystem of conservation laws included in the balance law (1.1) may be discretized in a non-conservative form. Moreover, the number of flux reconstructions in (1.5) is doubled compared with normal schemes for hyperbolic conservation laws, since the numerical flux at $x_{j+\frac{1}{2}}$ has to be evaluated for both $\mathbf{f}(\mathbf{u}) - \mathbf{f}(\mathbf{u}_{e,j})$ and $\mathbf{f}(\mathbf{u}) - \mathbf{f}(\mathbf{u}_{e,j+1})$. The same issues are also encountered in the scheme (1.3) in general. While these issues have been addressed for some particular steady states in [36], the problems are not fully resolved for arbitrary non-hydrostatic equilibria. We refer the readers to [36] for more discussions.

In this paper, we refine the generic scheme (1.3) to establish a well-balanced finite difference method with exact conservation property and higher efficiency through the interpolation of equilibrium variables in the A-WENO framework. To achieve this goal, we make a further assumption on the equilibrium states of (1.1).

Assumption 1.1. *The equilibrium states* $\mathbf{u}_e$ *of* (1.1) *satisfy* $\mathbf{v}_e := V(\mathbf{u}_e(x), x) =$ *constant and* $\mathbf{u}_e(x) = U(\mathbf{v}_e, x)$*, where* $V : \mathcal{O} \times \mathbb{R} \to \mathbb{R}^m$ *and* $U : G \to \mathbb{R}^m$ *with* $G := \{(\mathbf{v}, x) : \mathbf{v} = V(\mathbf{w}, x), \mathbf{w} \in \mathcal{O}, x \in \mathbb{R}\}$ *are mappings between the conserved and equilibrium variables. Furthermore,* $\forall (\mathbf{v}^*, x^*) \in G^o$, $U(\mathbf{v}^*, x)$ *is a local equilibrium state of* (1.1) *in a neighborhood of* x^**, and for any local equilibrium states* $\mathbf{u}_{e,loc}$, $V(\mathbf{u}_{e,loc}, x) =$ *constant.*

The key to the improvement of our algorithm is the specially designed flux approximation $\mathcal{F}$, which is based on the interpolation of equilibrium variables in the A-WENO framework, such that the source term approximation in (1.3) is efficient and conservative. We shall provide the definition of $\mathcal{F}$ and discuss the properties of the resulting scheme in the next section. Before we conclude, we would like to notify that, the idea of interpolation of equilibrium variables was also used in [12] to yield a second-order moving-water

equilibria preserving scheme for the shallow water equations. A similar idea of applying interpolation to Riemann invariants in the A-WENO scheme to improve the efficiency was used in our recent work [54].

The remaining sections of this paper are structured as follows. In Section 2, we establish the high-order well-balanced A-WENO scheme for general hyperbolic balance laws and discuss its properties. The scheme is then applied to the shallow water equations with non-flat bottom typography and Euler equations in gravitational fields to preserve their non-hydrostatic equilibria in Section 3. In Section 4, we provide extensive numerical tests to demonstrate the good performance and well-balanced property of the proposed scheme. Finally, we conclude the paper with some closing remarks in Section 5.

2 High-order well-balanced alternative finite difference WENO scheme

In this section, we establish the high-order well-balanced A-WENO scheme for the hyperbolic balance law (1.1). For simplicity, we assume the grid is uniform with mesh size Δx, and denote by $I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$ the j-th cell centered at $x_j = j\Delta x$. To fix the ideas, we only present the fifth-order scheme in this paper, but the method can be generalized to arbitrary high order.

We first give a brief review of the fifth-order WENO interpolation. Consequently, the numerical flux $\mathcal{F}$ in (1.3) is defined to determine our well-balanced A-WENO scheme.

2.1 A review of the fifth-order WENO-JS interpolation

Given the function values $p_{j-2}, p_{j-1}, p_j, p_{j+1}, p_{j+2}$ at nodes x_k for $k = j-2, \ldots, j+2$, the fifth-order WENO-JS interpolation [19, 6, 40] at $x_{j+\frac{1}{2}}$ on I_j is given by

$$p^-_{j+\frac{1}{2}} = w_1 p^{-(1)}_{j+\frac{1}{2}} + w_2 p^{-(2)}_{j+\frac{1}{2}} + w_3 p^{-(3)}_{j+\frac{1}{2}}, \tag{2.1}$$

where

$$\begin{aligned} p^{-(1)}_{j+\frac{1}{2}} &= \frac{3}{8}p_{j-2} - \frac{5}{4}p_{j-1} + \frac{15}{8}p_j, \\ p^{-(2)}_{j+\frac{1}{2}} &= -\frac{1}{8}p_{j-1} + \frac{3}{4}p_j + \frac{3}{8}p_{j+1}, \\ p^{-(3)}_{j+\frac{1}{2}} &= \frac{3}{8}p_j + \frac{3}{4}p_{j+1} - \frac{1}{8}p_{j+2}, \end{aligned} \tag{2.2}$$

are linear interpolations from sub-stencils, and $w_1, w_2, w_3 \geq 0$ are nonlinear weights defined as

$$\omega_k = \frac{\tilde{\omega}_k}{\sum_{m=1}^{3} \tilde{\omega}_m}, \quad \tilde{\omega}_k = \frac{\gamma_k}{(\beta_k + \epsilon)^2}, \quad k = 1, 2, 3, \tag{2.3}$$

with ϵ a small positive number, e.g. $\epsilon = 10^{-6}$,

$$\gamma_1 = \frac{1}{16}, \quad \gamma_2 = \frac{5}{8}, \quad \gamma_3 = \frac{5}{16}, \tag{2.4}$$

and

$$\begin{aligned}
\beta_1 &= \frac{13}{12}(p_{j-2} - 2p_{j-1} + p_j)^2 + \frac{1}{4}(p_{j-2} - 4p_{j-1} + 3p_j)^2, \\
\beta_2 &= \frac{13}{12}(p_{j-1} - 2p_j + p_{j+1})^2 + \frac{1}{4}(p_{j-1} - p_{j+1})^2, \\
\beta_3 &= \frac{13}{12}(p_j - 2p_{j+1} + p_{j+2})^2 + \frac{1}{4}(3p_j - 4p_{j+1} + p_{j+2})^2.
\end{aligned} \tag{2.5}$$

The WENO interpolation is designed to achieve optimal accuracy in smooth regions, while, when the stencil includes discontinuities, the interpolation provides an approximation in an essentially non-oscillatory fashion by minimizing the weights of stencils containing such discontinuities.

For brevity, we denote the WENO interpolation procedure (2.1)-(2.5) by $p^-_{j+\frac{1}{2}} = \text{weno}\,(p_{j-2}, p_{j-1}, p_j, p_{j+1}, p_{j+2})$. The WENO interpolation at $x_{j-\frac{1}{2}}$ on I_j follows from mirror symmetry, i.e. $p^+_{j-\frac{1}{2}} = \text{weno}(p_{j+2}, p_{j+1}, p_j, p_{j-1}, p_{j-2})$. We shall abuse the notation to also let it denote the component-wise WENO interpolation for vector-valued inputs, e.g. $\mathbf{p}^-_{j+\frac{1}{2}} = \text{weno}\,(\mathbf{p}_{j-2}, \mathbf{p}_{j-1}, \mathbf{p}_j, \mathbf{p}_{j+1}, \mathbf{p}_{j+2})$. It is clear that $\mathbf{c} = \text{weno}\,(\mathbf{c}, \mathbf{c}, \mathbf{c}, \mathbf{c}, \mathbf{c})$ for arbitrary constant $\mathbf{c} \in \mathbb{R}^m$.

2.2 The definition of $\mathcal{F}$ and resulting A-WENO scheme (1.3)

We define

$$\mathcal{F}\left(\mathbf{u}_{j-2}, \ldots, \mathbf{u}_{j+3}\right) := \mathbf{h}(\mathbf{u}^-_{j+\frac{1}{2}}, \mathbf{u}^+_{j+\frac{1}{2}}) - \frac{1}{24}\Delta x^2 \left(\frac{\partial^2}{\partial x^2}\mathbf{f}\right)_{j+\frac{1}{2}} + \frac{7}{5760}\Delta x^4 \left(\frac{\partial^4}{\partial x^4}\mathbf{f}\right)_{j+\frac{1}{2}}, \tag{2.6}$$

to calculate the numerical flux $\hat{\mathbf{f}}_{j+\frac{1}{2}}$ on the left-hand side of the A-WENO scheme (1.3), where

$$\left(\frac{\partial^2}{\partial x^2}\mathbf{f}\right)_{j+\frac{1}{2}} = \frac{1}{\Delta x^2}\left(-\frac{5}{48}\mathbf{f}(\mathbf{u}_{j-2}) + \frac{13}{16}\mathbf{f}(\mathbf{u}_{j-1}) - \frac{17}{24}\mathbf{f}(\mathbf{u}_j) - \frac{17}{24}\mathbf{f}(\mathbf{u}_{j+1}) + \frac{13}{16}\mathbf{f}(\mathbf{u}_{j+2}) - \frac{5}{48}\mathbf{f}(\mathbf{u}_{j+3})\right), \tag{2.7}$$

and

$$\left(\frac{\partial^4}{\partial x^4}\mathbf{f}\right)_{j+\frac{1}{2}} = \frac{1}{\Delta x^4}\left(\frac{1}{2}\mathbf{f}(\mathbf{u}_{j-2}) - \frac{3}{2}\mathbf{f}(\mathbf{u}_{j-1}) + \mathbf{f}(\mathbf{u}_j) + \mathbf{f}(\mathbf{u}_{j+1}) - \frac{3}{2}\mathbf{f}(\mathbf{u}_{j+2}) + \frac{1}{2}\mathbf{f}(\mathbf{u}_{j+3})\right), \tag{2.8}$$

are high-order central differences to approximate $\frac{\partial^2}{\partial x^2}\mathbf{f}(\mathbf{u})$ and $\frac{\partial^4}{\partial x^4}\mathbf{f}(\mathbf{u})$ at $x_{j+\frac{1}{2}}$, respectively, $\mathbf{h}\,(\cdot,\cdot)$ is a consistent numerical flux based on exact or approximate Riemann solvers, e.g. the Lax-Friedrichs flux

$\mathbf{h}(\mathbf{w}_1, \mathbf{w}_2) = \frac{1}{2}(\mathbf{f}(\mathbf{w}_1) + \mathbf{f}(\mathbf{w}_2) - \alpha(\mathbf{w}_2 - \mathbf{w}_1))$, with $\alpha = \max_{\mathbf{u}}\{|\lambda_1(\mathbf{u})|, \ldots, |\lambda_m(\mathbf{u})|\}$, $\lambda_i(\mathbf{u})$ is the i-th eigenvalue of the Jacobian matrix $\frac{\partial \mathbf{f}}{\partial \mathbf{u}}$, and $\mathbf{u}^{\pm}_{j+\frac{1}{2}} = U\left(\mathbf{v}^{\pm}_{j+\frac{1}{2}}, x_{j+\frac{1}{2}}\right)$ are approximations of $\mathbf{u}$ at $x_{j+\frac{1}{2}}$ from the left/right side.

Most importantly, the calculation of $\mathbf{v}^{\pm}_{j+\frac{1}{2}}$ is performed as follows,

$$\begin{aligned} \mathbf{v}^{-}_{j+\frac{1}{2}} &= \text{weno}\left(V(\mathbf{u}_{j-2}, x_{j-2}), \ldots, V(\mathbf{u}_{j+2}, x_{j+2})\right), \\ \mathbf{v}^{+}_{j+\frac{1}{2}} &= \text{weno}(V(\mathbf{u}_{j+3}, x_{j+3}), \ldots, V(\mathbf{u}_{j-1}, x_{j-1})). \end{aligned} \tag{2.9}$$

For systems of hyperbolic equations, the component-wise WENO interpolation (2.9) may not be enough to ensure essentially non-oscillatory performance when there are multiple interacting shocks [38]. A more robust way to calculate $\mathbf{v}^{\pm}_{j+\frac{1}{2}}$ is to use the local characteristic decomposition [17]:

$$\begin{aligned} \mathbf{v}^{-}_{j+\frac{1}{2}} &= \mathbf{K}_{j+\frac{1}{2}}\text{weno}(\mathbf{K}^{-1}_{j+\frac{1}{2}}V(\mathbf{u}_{j-2}, x_{j-2}), \ldots, \mathbf{K}^{-1}_{j+\frac{1}{2}}V(\mathbf{u}_{j+2}, x_{j+2})), \\ \mathbf{v}^{+}_{j+\frac{1}{2}} &= \mathbf{K}_{j+\frac{1}{2}}\text{weno}(\mathbf{K}^{-1}_{j+\frac{1}{2}}V(\mathbf{u}_{j+3}, x_{j+3}), \ldots, \mathbf{K}^{-1}_{j+\frac{1}{2}}V(\mathbf{u}_{j-1}, x_{j-1})), \end{aligned} \tag{2.10}$$

where $\mathbf{K}_{j+\frac{1}{2}} = \frac{\partial V}{\partial \mathbf{u}}(\mathbf{u}_{j+\frac{1}{2}}, x_{j+\frac{1}{2}})\mathbf{R}_{j+\frac{1}{2}}$ [55], $\mathbf{R}_{j+\frac{1}{2}}$ is obtained from the eigendecomposition of Jacobian matrix $\frac{\partial \mathbf{f}}{\partial \mathbf{u}}(\mathbf{u}_{j+\frac{1}{2}}) = \mathbf{R}_{j+\frac{1}{2}}\mathbf{\Lambda}_{j+\frac{1}{2}}\mathbf{R}^{-1}_{j+\frac{1}{2}}$, $\mathbf{\Lambda}_{j+\frac{1}{2}}$ is a diagonal matrix, and $\mathbf{u}_{j+\frac{1}{2}}$ is obtained by either the arithmetic mean or Roe's average [39] of $\mathbf{u}_j$ and $\mathbf{u}_{j+1}$.

Based on the definition of $\mathcal{F}$, we can determine the source term approximation in (1.3):

$$\hat{\mathbf{f}}^{e}_{j,j\pm\frac{1}{2}} = \mathbf{f}(U(\mathbf{v}_{e,j}, x_{j\pm\frac{1}{2}})) - \frac{1}{24}\Delta x^2\left(\frac{\partial^2}{\partial x^2}\mathbf{f}^{e}_{j}\right)_{j\pm\frac{1}{2}} + \frac{7}{5760}\Delta x^4\left(\frac{\partial^4}{\partial x^4}\mathbf{f}^{e}_{j}\right)_{j\pm\frac{1}{2}}, \tag{2.11}$$

where $\mathbf{v}_{e,j} := V(\mathbf{u}_{e,j}, x)$ is a constant, $\left(\frac{\partial^2}{\partial x^2}\mathbf{f}^{e}_{j}\right)_{j\pm\frac{1}{2}}$ and $\left(\frac{\partial^4}{\partial x^4}\mathbf{f}^{e}_{j}\right)_{j\pm\frac{1}{2}}$ are defined by simply replacing $\mathbf{u}_{j+k}$ with $\mathbf{u}_{e,j}(x_{j+k})$ for $k = -3, \ldots, 3$, in the definitions of $\left(\frac{\partial^2}{\partial x^2}\mathbf{f}\right)_{j\pm\frac{1}{2}}$ and $\left(\frac{\partial^4}{\partial x^4}\mathbf{f}\right)_{j\pm\frac{1}{2}}$, respectively. The first term on the right-hand side of (2.11) is $\mathbf{f}(U(\mathbf{v}_{e,j}, x_{j\pm\frac{1}{2}}))$ because $\mathbf{h}(\cdot,\cdot)$ is consistent and the WENO interpolation procedure (2.9) or (2.10) returns the constant $\mathbf{v}_{e,j}$ if $\mathbf{u}_k$ are replaced with $\mathbf{u}_{e,j}(x_k)$ therein.

The equations (1.3), (2.6), (2.7), (2.8), (2.10) (or (2.9)), and (2.11) form our A-WENO scheme for (1.1). From the arguments in the introduction, it is clear that the scheme is well-balanced. In addition, the source term discretization (2.11) does not involve any WENO interpolation, local characteristic decomposition or Riemann solver-based flux calculation, thus the scheme is efficient.

More importantly, conservation laws contained in (1.1) are discretized in a conservative form. To see this, we assume the i-th component of (1.1) is a conservation law, i.e. $u_{it} + f_i(\mathbf{u})_x = 0$, where u_i and f_i denotes the i-th component of the vector of conserved variables and fluxes, respectively. At steady states, we have

$f_i(\mathbf{u}) =$ constant, indicating $f_i(\mathbf{u})$ is one of the components of the vector of equilibrium variables, say the k-th component. Thus, $f_i(U(\mathbf{v}_{e,j}, x_{j+\frac{1}{2}})) = \mathbf{v}_{e,j}(k) = f_i(U(\mathbf{v}_{e,j}, x_{j-\frac{1}{2}}))$ in (2.11). For the same reason, the i-th component in the high-order terms in (2.11) are all zero. Therefore the i-th component of the source term approximation in (1.3) is always zero, ensuring the subsystem of conservation laws contained in the balance law (1.1) is discretized in a conservative form.

3 Applications

In this section, we consider the shallow water equations with non-flat bottom topography and Euler equations in gravitational fields as applications of the well-balanced A-WENO scheme established in Section 2. The definitions of mappings V, U and choice of local reference equilibrium states $\mathbf{u}_{e,j}$ for these equations shall be given.

3.1 Shallow water equations

The shallow water equations with non-flat bottom topography take the form of

$$\begin{cases} h_t + m_x = 0 \\ m_t + \left(\frac{m^2}{h} + \frac{1}{2}gh^2\right)_x = -ghb_x, \end{cases} \tag{3.1}$$

i.e. $\mathbf{u} = (h, m)^T$, $\mathbf{f}(\mathbf{u}) = \left(m, \frac{m^2}{h} + \frac{1}{2}gh^2\right)^T$ and $\mathbf{s}(\mathbf{u}, x) = (0, -ghb_x(x))^T$ in (1.1), where h is the water depth, m is the discharge, g is the gravitational constant, and b is the prescribed bottom topography. We assume b is continuous and $h > 0$ (no dry area).

The moving-water equilibrium of the shallow water equations (3.1) is characterized by

$$m = \text{constant} \quad \text{and} \quad Q := \frac{1}{2}\frac{m^2}{h^2} + g(h + b) = \text{constant}, \tag{3.2}$$

where m and Q are referred to as the equilibrium variables.

Given the conservative variables h and m, it is straightforward to calculate Q by definition in (3.2). However, recovering h from the equilibrium variables m and Q is more complex. For this, we introduce the Froude number as $\text{Fr} := \frac{|m|}{\sqrt{gh^3}}$, and the flow regime indicator as $\sigma := \text{sign}\,(\text{Fr} - 1)$. A state is classified as subcritical, critical, or supercritical if $\sigma = -1, 0$, or 1. The water depth h is determined implicitly by the following equation:

$$Q - gb = \frac{1}{2}\frac{m^2}{h^2} + gh. \tag{3.3}$$

The right-hand side of (3.3) defines a function $\psi(h) = \frac{1}{2}\frac{m^2}{h^2} + gh$ for $h > 0$, whose graph in the case $m \neq 0$ is shown in Figure 1. Note that $\psi(h) \geq \frac{3}{2}(g|m|)^{\frac{2}{3}}, \forall h > 0$, from the graph. If $m = 0$, (3.3) has a unique solution $h = \frac{Q-gb}{g}$. Otherwise, the equation has no positive solution, one positive solution, or two positive solutions in the subcritical and supercritical flow regimes, respectively, if $Q - gb <, =,$ or $> \frac{3}{2}(g|m|)^{\frac{2}{3}}$. For further details, we refer to [35].

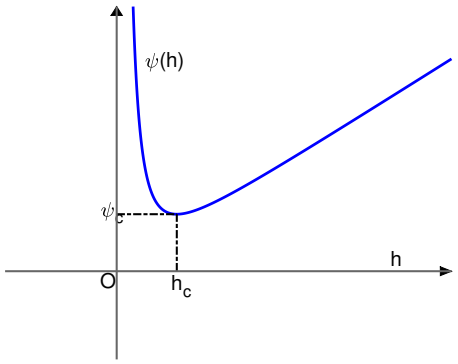


Figure 1: A graph of $\psi(h)$ in the case $m \neq 0$, where $h_c = \left(\frac{m^2}{g}\right)^{\frac{1}{3}}, \psi_c = \frac{3}{2}(g|m|)^{\frac{2}{3}}$ is the critical point. The states are supercritical for $0 < h < h_c$, critical for $h = h_c$, and subcritical for $h > h_c$.

There is an explicit formula of the solution of (3.3), including the case $m = 0$, given by the formula of trigonometric solution of cubic equation [27, 23]:

$$h(m, Q, b; \sigma) = \frac{Q - gb}{3g}\left(1 + \cos(\frac{\theta}{3}) - \sqrt{3}\sigma\sin(\frac{\theta}{3})\right), \quad \text{where} \quad \theta = \arccos\left(\frac{27}{4}\frac{g^2m^2}{(Q-gb)^3} - 1\right), \tag{3.4}$$

where σ is the flow regime indicator.

We denote $\mathbf{v} = (m, Q)^T$ and define the mappings V and U in Assumption 1.1 as $V(\mathbf{u}, x) = (m, \frac{1}{2}\frac{m^2}{h^2} + g(h + b(x)))^T$ and $U(\mathbf{v}, x) = (h(\mathbf{v}, b(x); \sigma), m)^T$, where σ indicates the flow regime. One can check that Assumption 1.1 is satisfied by the mappings. The local reference equilibrium state $\mathbf{u}_{e,j}$ in scheme (1.3) is taken as

$$\mathbf{u}_{e,j}(x) = U(\mathbf{v}_{e,j}, x), \tag{3.5}$$

where

$$\mathbf{v}_{e,j} = \begin{cases} V(\mathbf{u}_\alpha, x_\alpha), & \text{if } \alpha \in \{j-3, \dots, j+3\}, \\ \mathbf{v}_\alpha^+, & \text{if } \alpha \in \{j \pm \frac{1}{2}\} \text{ and } Q_\alpha^+ - \frac{3}{2}(g|m_\alpha^+|)^{\frac{2}{3}} > Q_\alpha^- - \frac{3}{2}(g|m_\alpha^-|)^{\frac{2}{3}} \\ \mathbf{v}_\alpha^-, & \text{otherwise,} \end{cases}$$

with $\alpha = \arg\max_{\beta \in \{j-3,\dots,j+3,j\pm\frac{1}{2}\}} b(x_\beta)$. Such a choice ensures $\mathbf{u}_{e,j}$ is always well-defined in the stencil we use.

3.2 Euler equations

The Euler equations of polytropic gas in gravitational fields take the form of

$$\begin{cases} \rho_t + m_x = 0 \\ m_t + \left(\rho v^2 + p\right)_x = -\rho\phi_x \\ E_t + \left((E+p)v\right)_x = -m\phi_x, \end{cases} \tag{3.6}$$

i.e. $\mathbf{u} = (\rho, m, E)^T$, $\mathbf{f}(\mathbf{u}) = \left(m, (\rho v^2 + p), (E+p)v\right)^T$ and $\mathbf{s}(\mathbf{u}, x) = (0, -\rho\phi_x(x), -m\phi_x(x))^T$ in (1.1), where ρ is the density of fluid, $m = \rho v$ is the momentum, v is the velocity, $E = \frac{1}{2}\rho v^2 + \frac{p}{\gamma - 1}$ is the total energy, p is the pressure, $\gamma > 1$ is the ratio of specific heats, and $\phi(x)$ is the prescribed gravitational potential. We assume ϕ is continuous and $\rho, p > 0$ (no vacuum).

The isentropic equilibrium of the Euler equations (3.6) is characterized by

$$s := \frac{p}{\rho^\gamma} = \text{constant}, \quad m = \text{constant}, \quad \text{and} \quad Q := \frac{1}{2}\frac{m^2}{\rho^2} + \frac{\gamma}{\gamma - 1} s\rho^{\gamma - 1} + \phi = \text{constant}, \tag{3.7}$$

where s, m and Q are referred to as the equilibrium variables.

To recover ρ and E from the equilibrium variables, we introduce the Mach number as $\text{Ma} := \frac{|v|}{c}$, where $c = \sqrt{\frac{\gamma p}{\rho}}$ is the sound speed, and the flow regime indicator as $\sigma := \text{sign}(\text{Ma} - 1)$. A state is classified as subcritical, critical, or supercritical if $\sigma = -1, 0$ or 1. The density ρ is determined implicitly by the following equation:

$$Q - \phi = \frac{1}{2}\frac{m^2}{\rho^2} + \frac{\gamma}{\gamma - 1} s\rho^{\gamma - 1}. \tag{3.8}$$

The right-hand side of (3.8) defines a function $\psi(\rho) = \frac{1}{2}\frac{m^2}{\rho^2} + \frac{\gamma}{\gamma - 1} s\rho^{\gamma - 1}$ for $\rho > 0$, whose graph in the case $m \neq 0$ is shown in Figure 2. Note that $\psi(\rho) \geq \left(\frac{1}{2} + \frac{1}{\gamma - 1}\right)(\gamma s)^{\frac{2}{\gamma+1}} |m|^{\frac{2(\gamma-1)}{\gamma+1}}$, for $\rho > 0$, from the graph. If $m = 0$, (3.8) has a unique solution $\rho = \left(\frac{\gamma - 1}{s\gamma}(Q - \phi)\right)^{\frac{1}{\gamma - 1}}$. Otherwise, the equation has no positive solution,

one positive solution, or two positive solutions corresponding to the subcritical and supercritical regimes, respectively, if $Q-\phi <, =,$ or $> \left(\frac{1}{2}+\frac{1}{\gamma-1}\right)(\gamma s)^{\frac{2}{\gamma+1}}|m|^{\frac{2(\gamma-1)}{\gamma+1}}$. For further details, we refer to [16].

In practice, we compute ρ by Newton's method, with the initial guess $\rho_{\text{init}} = \left(\frac{\gamma-1}{\gamma s}(Q-\phi)\right)^{\frac{1}{\gamma-1}}$ and $\rho_{\text{init}} = \frac{|m|}{\sqrt{2(Q-\phi)}}$ in the subcritical and supercritical regimes, respectively. Such choices of the initial guess satisfy $\rho_{\text{init}} > \rho$ if the solution is in the subcritical regime while $\rho_{\text{init}} < \rho$ if the solution is in the supercritical regime, ensuring the Newton iteration always converges to the exact solution. Once ρ is obtained, E can be determined by $E = \frac{1}{2}\frac{m^2}{\rho} + \frac{p}{\gamma-1}$, where $p = s\rho^\gamma$.

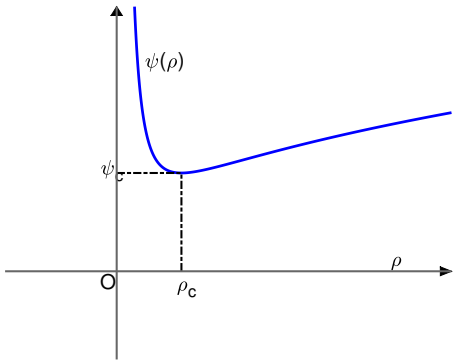


Figure 2: A graph of $\psi(\rho)$ in the case $m \neq 0$, where $\rho_c = \left(\frac{m^2}{\gamma s}\right)^{\frac{1}{\gamma+1}}$, $\psi_c = \left(\frac{1}{2}+\frac{1}{\gamma-1}\right)(\gamma s)^{\frac{2}{\gamma+1}}|m|^{\frac{2(\gamma-1)}{\gamma+1}}$ is the critical point. The states are supercritical for $0 < \rho < \rho_c$, critical for $\rho = \rho_c$, and subcritical for $\rho > \rho_c$.

We denote $\mathbf{v} = (s, m, Q)^T$ and define the mappings V and U in Assumption 1.1 as $V(\mathbf{u}, x) = ((\gamma-1)(E-\frac{1}{2}\frac{m^2}{\rho})\rho^{-\gamma}, m, \frac{1}{2}\frac{m^2}{\rho^2}+\gamma(E-\frac{1}{2}\frac{m^2}{\rho})\rho^{-1}+\phi(x))^T$ and $U(\mathbf{v}, x) = (\rho(\mathbf{v}, \phi(x); \sigma), m, E(\mathbf{v}, \phi(x); \sigma))^T$, where σ indicates the flow regime. One can check that Assumption 1.1 is satisfied by the mappings. The local reference equilibrium state $\mathbf{u}_{e,j}$ in scheme (1.3) is taken as

$$\mathbf{u}_{e,j}(x) = U(\mathbf{v}_{e,j}, x), \tag{3.9}$$

where

$$\mathbf{v}_{e,j} = \begin{cases} V(\mathbf{u}_\alpha, x_\alpha), & \text{if } \alpha \in \{j-3, \ldots, j+3\}, \\ \mathbf{v}_\alpha^+, & \text{if } \alpha \in \{j \pm \frac{1}{2}\} \text{ and } Q_\alpha^+ - \left(\frac{1}{2}+\frac{1}{\gamma-1}\right)(\gamma s_\alpha^+)^{\frac{2}{\gamma+1}}|m_\alpha^+|^{\frac{2(\gamma-1)}{\gamma+1}} > Q_\alpha^- - \left(\frac{1}{2}+\frac{1}{\gamma-1}\right)(\gamma s_\alpha^-)^{\frac{2}{\gamma+1}}|m_\alpha^-|^{\frac{2(\gamma-1)}{\gamma+1}}, \\ \mathbf{v}_\alpha^-, & \text{otherwise}, \end{cases}$$

with $\alpha = \arg\max_{\beta\in\{j-3,\dots,j+3,j\pm\frac{1}{2}\}} \phi(x_\beta)$. The above choice ensures $\mathbf{u}_{e,j}$ is always well-defined in the stencil we use.

Remark 3.1. *To this end, we would like to comment that, the mapping $U(\mathbf{v}, x)$ is non-smooth at critical points, i.e. $\left|\frac{\partial U}{\partial \mathbf{v}}\right| = \infty$ when $Fr = 1$ or $Ma = 1$. At such non-smooth points, even $\mathbf{v}^{\pm}_{j+\frac{1}{2}}$ obtained from interpolation in* (2.9) *or* (2.10) *are constant up to round-off error at steady states, the computation $\mathbf{u}^{\pm}_{j+\frac{1}{2}} = U\left(\mathbf{v}^{\pm}_{j+\frac{1}{2}}, x_{j+\frac{1}{2}}\right)$ may amplify errors to truncation error level, hence our well-balanced approach may fail there. Therefore, we need to avoid $x_{j+\frac{1}{2}}$ landing at critical points to ensure the well-balanced property. In practice, we return the critical value h_c (or ρ_c) in the computation of $U(\mathbf{v}, x)$ when $|Q - Q_c| < 100\epsilon_0$, where ϵ_0 is the machine epsilon, to enhance the robustness of the well-balanced property.*

4 Numerical tests

In this section, we test the performance of our well-balanced A-WENO scheme established in previous sections for the shallow water equations with non-flat bottom topography and Euler equations in gravitational fields, with the emphases on the high order accuracy, well-balanced property and capability of capturing perturbations upon equilibrium states. There is no need to test the conservation property as the source term approximation corresponding to the subsystem of conservation laws is simply zero. In all the tests, time discretization is carried out by the classical fourth order Runge–Kutta method. The CFL number is taken as 0.4 if not otherwise stated.

4.1 Shallow water equations

These examples are classical tests for the accuracy and well-balanced property of algorithms for the shallow water equations. One can find results of other numerical schemes in the articles cited and references therein for comparison. Unless otherwise stated, the gravitational constant is taken as $g = 9.812$ and boundary conditions are taken the same as the initial conditions there.

Example 4.1. ***Accuracy test***

In this example [35, 26], we test the accuracy of the well-balanced A-WENO scheme for a smooth solution. The bottom topography and initial conditions are given by

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

on the computational domain $\Omega = [0, 1]$ *with periodic boundary condition.*

Since the explicit formulation of the exact solution is not available, we use the ninth-order non-well-balanced finite difference WENO scheme[42, 19] with $N = 6400$ *cells and CFL*$= 0.1$ *to give the reference solution. The* L^1 *errors and orders of convergence of* h *and* m *at the terminal time* $t = 0.1$ *(before shock formation) is presented in Table 1. From the table, we can clearly observe the high-order accuracy.*

		h		m	
N	CFL	L^1 error	Order	L^1 error	Order
50	0.6	1.82E-03	-	1.78E-02	-
100	0.4	2.77E-04	2.72	2.32E-03	2.94
200	0.3	1.69E-05	4.04	1.42E-04	4.03
400	0.2	6.61E-07	4.67	5.69E-06	4.64
800	0.1	2.33E-08	4.82	2.03E-07	4.81

Table 1: L^1 errors and orders of convergence in the accuracy test for Example 4.1

Example 4.2. ***Well-balanced test***

In this example [44, 35, 12], we test the well-balanced property of the scheme for the shallow water equations with moving-water equilibria.

We consider the bottom topography

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

on the computational domain $\Omega = [0, 25]$ *with three sets of steady-state initial conditions:*

(a) Supercritical flow

$$m(x,0) = 24, \quad Q(x,0) = \frac{24^2}{2 \times 2^2} + 9.812 \times 2, \quad \textit{with} \quad Fr(x,0) > 1.$$

(b) Subcritical flow

$$m(x,0) = 4.42, \quad Q(x,0) = \frac{4.42^2}{2 \times 2^2} + 9.812 \times 2, \quad \textit{with} \quad Fr(x,0) < 1.$$

(c) Transcritical flow

$$m(x,0) = 1.53, \quad Q(x,0) = \frac{3}{2}\left(9.812 \times 1.53\right)^{\frac{2}{3}} + 9.812 \times 0.2,$$

$$\textit{with} \quad \begin{cases} Fr(x,0) < 1, & x < 10, \\ Fr(x,0) = 1, & x = 10, \\ Fr(x,0) > 1, & x > 10. \end{cases}$$

We compute the solutions up to $t = 20$ with $N = 200$ cells, and list their errors in Table 2. From the table, we can observe that the equilibrium solutions are preserved within machine epsilon, which verifies the well-balanced property of our scheme.

		h		m	
Case	Precision	L^1 error	L^∞ error	L^1 error	L^∞ error
(a)	Single	5.06E-04	4.03E-05	6.72E-03	5.25E-04
	Double	5.58E-14	1.07E-14	5.42E-13	1.21E-13
	Quadruple	6.64E-32	1.31E-32	5.97E-31	1.42E-31
(b)	Single	1.02E-03	6.16E-05	1.55E-03	1.33E-04
	Double	7.74E-15	1.33E-15	4.54E-14	8.88E-15
	Quadruple	2.05E-32	2.89E-33	4.52E-32	9.24E-33
(c)	Single	2.34E-04	3.44E-05	6.17E-04	5.84E-05
	Double	1.08E-14	3.15E-14	1.54E-14	3.55E-15
	Quadruple	6.12E-33	1.44E-33	2.31E-32	3.27E-33

Table 2: L^1 and L^∞ errors for different precision in the well-balanced test for Example 4.2

Example 4.3. ***A small perturbation of moving-water equilibria***

In this example, we consider a small perturbation on the moving-water equilibria of Example 4.2, and compute the solutions using both the well-balanced scheme and a non-well-balanced scheme with a straightforward approximation of the source term for comparison.

We add $\delta = 0.001$ on the initial water depth of case (a), (b) and (c) of Example 4.2 in the region $x \in [5.75, 6.25]$, and compute the solutions up to $t = 1, t = 1.5$ and $t = 1.5$, respectively. We draw the

discrepancy of the solutions with respect to their corresponding base steady states in Figure 3, 4, and 5 for the case (a), (b) and (c), respectively.

As we can observe from the figures, the well-balanced scheme provides excellent resolution on all grids, while the non-well-balanced scheme only resolves the spurious oscillations when the meshes are extremely refined.

Example 4.4. ***A small perturbation of lake-at-rest***

In this example [28],we consider a small perturbation of lake-at-rest and compute the solutions using both the well-balanced scheme and a non-well-balanced scheme with a straightforward approximation of the source term for comparison.

The bottom topography is given by

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

on the computational domain $\Omega = [0, 2]$. *The initial conditions are*

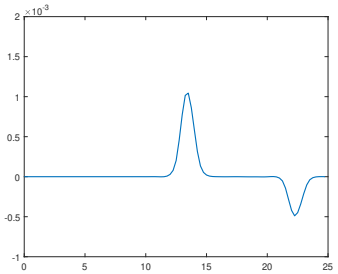
$$h(x,0) = \begin{cases} 1 - b(x) + \delta, & 1.1 \le x \le 1.2, \\ 1 - b(x), & \textit{otherwise}. \end{cases}, \quad m(x,0) = 0.$$

The lake-at-rest ($\delta = 0$) *is preserved within machine epsilon in our test and the results are omitted to save space. We take* $\delta = 0.001$ *to test the capability of capturing small perturbations of our well-balanced scheme. The surface level* $h + b$ *and discharge* m *at* $t = 0.2$ *are shown in Figure 6 and 7, respectively, for different methods on various grids. As shown by the figures, our scheme successfully captures the perturbations on coarse meshes, whereas the non-well-balanced scheme requires much finer grids to attain a comparable resolution.*

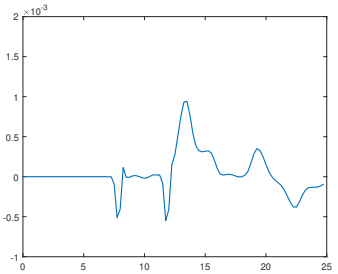
Example 4.5. ***Dam breaking over a rectangular bump***

In this example [50], we simulate the dam breaking over a rectangular bump. This problem has a rapidly varying flow over a discontinuous bed, which is described by the bottom topography

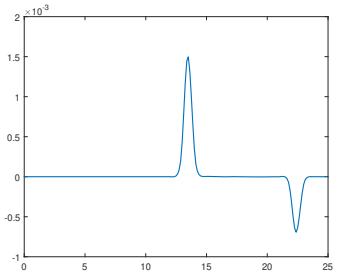
$$b(x) = \begin{cases} 8, & |x - 750| \le 187.5, \\ 0, & \textit{otherwise}, \end{cases}$$



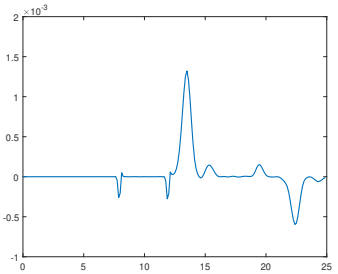
(a) Well-balanced scheme, $N = 100$



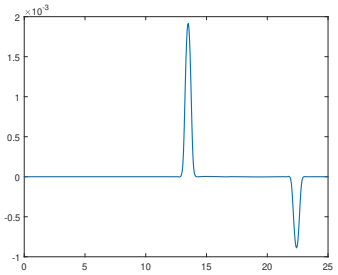
(b) Non-well-balanced scheme, $N = 100$



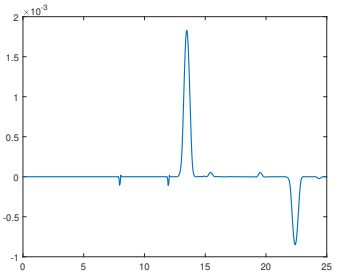
(c) Well-balanced scheme, $N = 200$



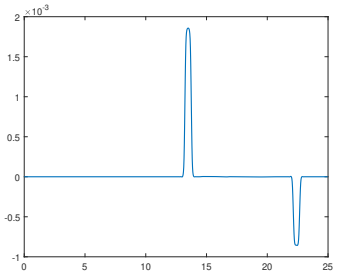
(d) Non-well-balanced scheme, $N = 200$



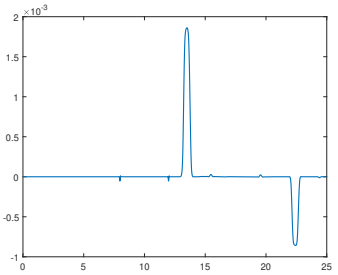
(e) Well-balanced scheme, $N = 500$



(f) Non-well-balanced scheme, $N = 500$

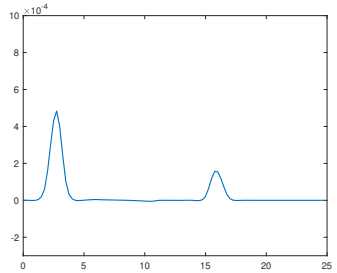


(g) Well-balanced scheme, $N = 1000$

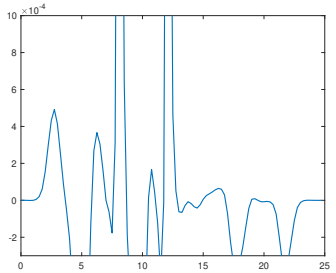


(h) Non-well-balanced scheme, $N = 1000$

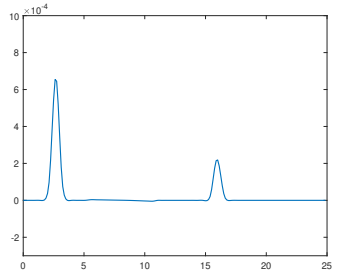
Figure 3: Discrepancy of the solutions of case (a) of Example 4.3 with respect to the corresponding base steady state at $t = 1$.



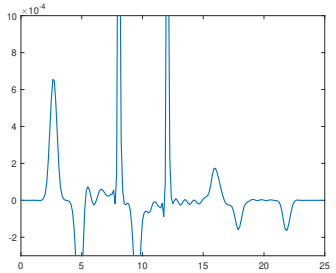
(a) Well-balanced scheme, $N = 100$



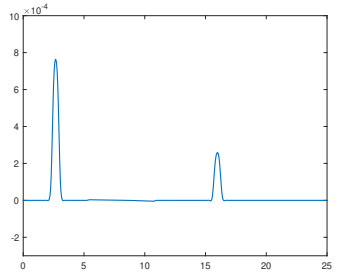
(b) Non-well-balanced scheme, $N = 100$



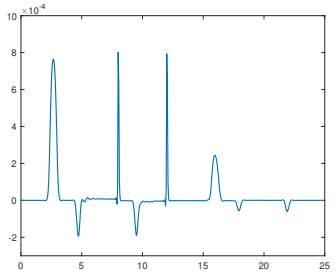
(c) Well-balanced scheme, $N = 200$



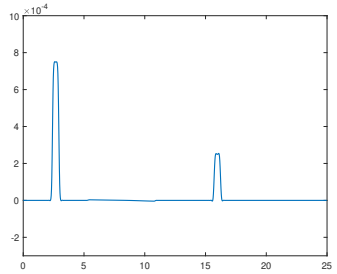
(d) Non-well-balanced scheme, $N = 200$



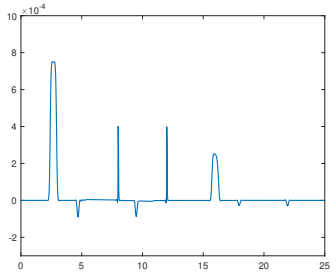
(e) Well-balanced scheme, $N = 500$



(f) Non-well-balanced scheme, $N = 500$

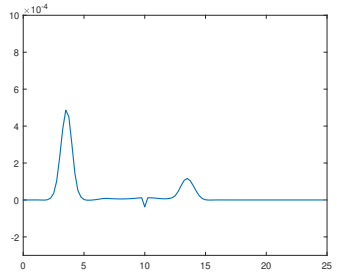


(g) Well-balanced scheme, $N = 1000$

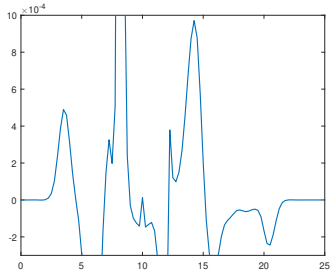


(h) Non-well-balanced scheme, $N = 1000$

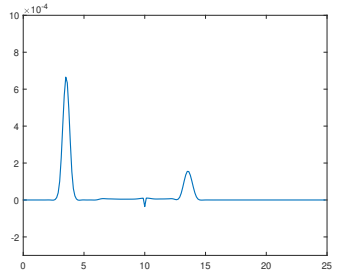
Figure 4: Discrepancy of the solutions of case (b) of Example 4.3 with respect to the corresponding base steady state at $t = 1.5$.



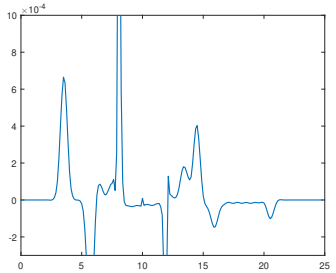
(a) Well-balanced scheme, $N = 100$



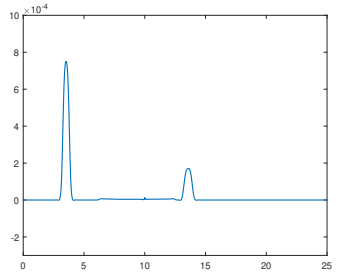
(b) Non-well-balanced scheme, $N = 100$



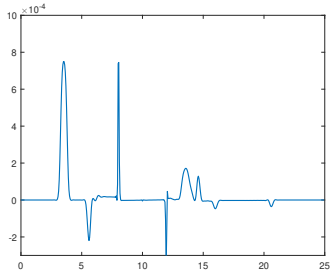
(c) Well-balanced scheme, $N = 200$



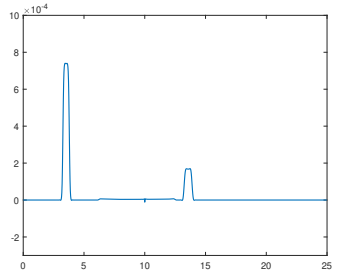
(d) Non-well-balanced scheme, $N = 200$



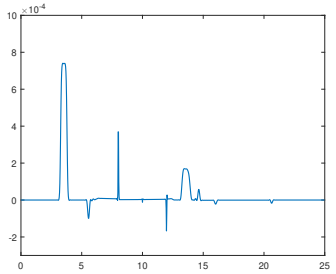
(e) Well-balanced scheme, $N = 500$



(f) Non-well-balanced scheme, $N = 500$

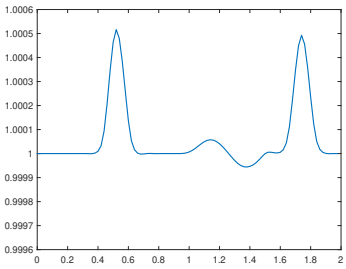


(g) Well-balanced scheme, $N = 1000$

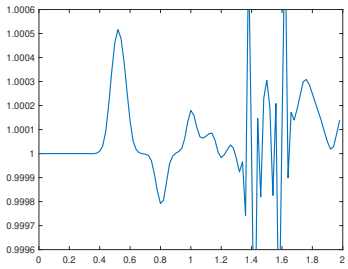


(h) Non-well-balanced scheme, $N = 1000$

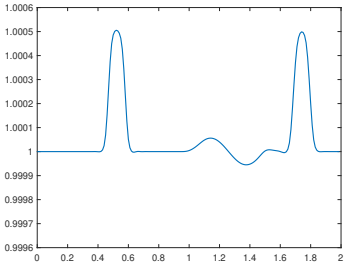
Figure 5: Discrepancy of the solutions of case (c) of Example 4.3 with respect to the corresponding base steady state at $t = 1.5$. A small spike is located at $x = 10$ due to the non-smooth transform between conserved and equilibrium variables at critical point.



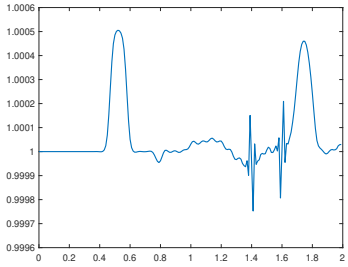
(a) Well-balanced scheme, $N = 100$



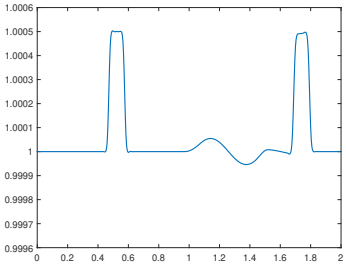
(b) Non-well-balanced scheme, $N = 100$



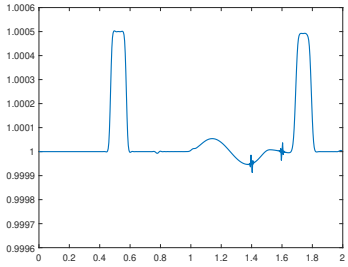
(c) Well-balanced scheme, $N = 200$



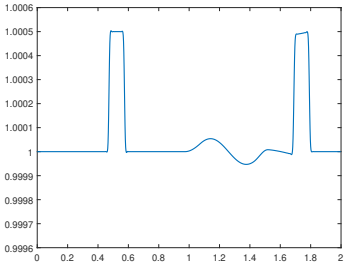
(d) Non-well-balanced scheme, $N = 200$



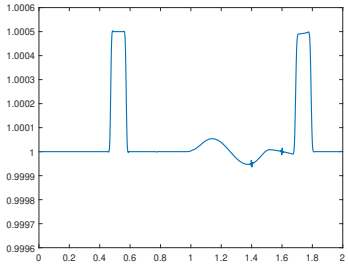
(e) Well-balanced scheme, $N = 500$



(f) Non-well-balanced scheme, $N = 500$

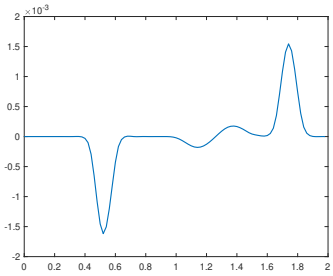


(g) Well-balanced scheme, $N = 1000$

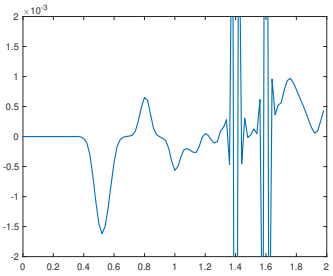


(h) Non-well-balanced scheme, $N = 1000$

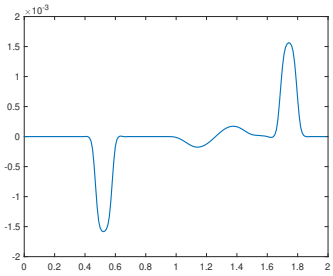
Figure 6: Surface level $h + b$ of Example 4.4 at $t = 0.2$.



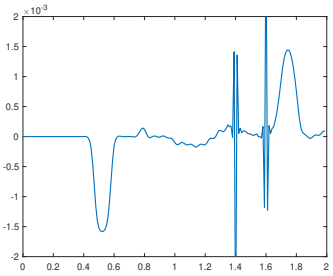
(a) Well-balanced scheme, $N = 100$



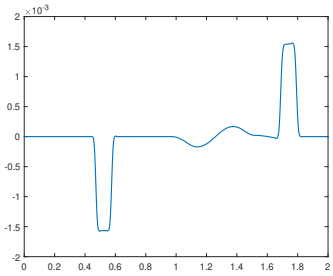
(b) Non-well-balanced scheme, $N = 100$



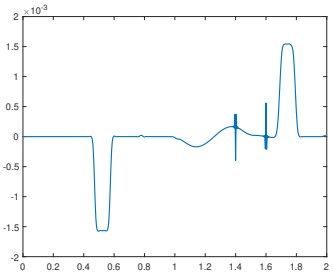
(c) Well-balanced scheme, $N = 200$



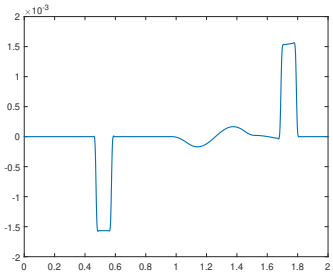
(d) Non-well-balanced scheme, $N = 200$



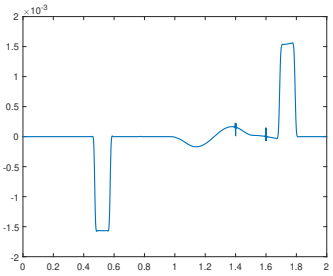
(e) Well-balanced scheme, $N = 500$



(f) Non-well-balanced scheme, $N = 500$



(g) Well-balanced scheme, $N = 1000$



(h) Non-well-balanced scheme, $N = 1000$

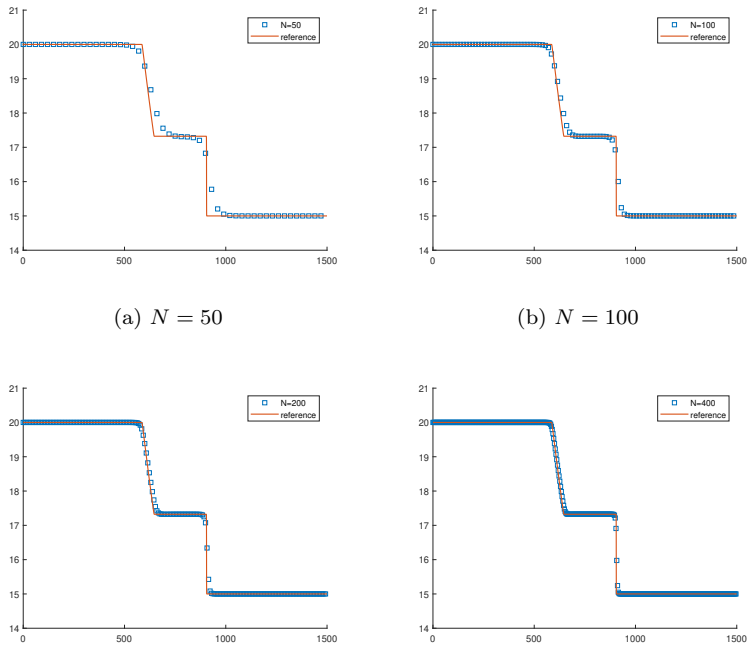
Figure 7: Discharge m of Example 4.4 at $t = 0.2$.

and initial conditions

$$h(x,0) = \begin{cases} 20 - b(x), & x \le 750, \\ 15 - b(x), & otherwise, \end{cases}, \quad m(x,0) = 0,$$

on the domain $\Omega = [0, 1500]$.

We compute the solution on a sequence of grids $N = 50, 100, 200, 400$*, and draw the surface level* $h + b$ *in Figure 8 and 9, at the time* $t = 15$ *and* 60*, respectively. The results of* 6400 *cells are taken as the reference solutions. The figures show good resolution and essentially non-oscillatory property of our well-balanced scheme for solutions away from equilibrium states.*



(a) $N = 50$ (b) $N = 100$

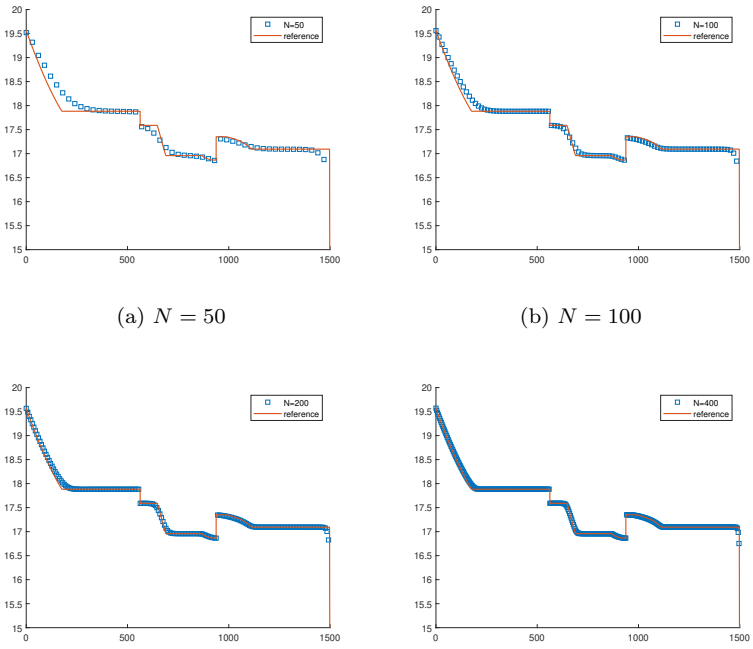
(c) $N = 200$ (d) $N = 400$

Figure 8: Surface level $h + b$ of the dam breaking problem in Example 4.5 at $t = 15$.

4.2 Euler equations

We take several classic examples [51, 16] to demonstrate the good performance of our scheme for the Euler equations in gravitational fields.

Example 4.6. ***Accuracy test***



(a) $N = 50$ (b) $N = 100$

(c) $N = 200$ (d) $N = 400$

Figure 9: Surface level $h + b$ of the dam breaking problem in Example 4.5 at $t = 60$.

In this example, we test the accuracy of the well-balanced A-WENO scheme for the Euler equations in gravitational fields with a smooth solution.

One can check that an exact solution for the linear gravitational filed $\phi(x) = x$ *takes the form of*

$$\rho(x,t) = 1 + \frac{1}{5}\sin(\pi(x-u_0t)), \quad u(x,t) = u_0, \quad p(x,t) = p_0 + u_0t - x + \frac{1}{5\pi}\cos(\pi(x-u_0t)),$$

where u_0, p_0 *are constants. In the computation, we take* $u_0 = 1, p_0 = 4.5$ *and* $\gamma = \frac{5}{3}$.

The L^1 *errors and orders of convergence of* ρ, m, E *at* $t = 0.1$ *are collected in Table 3, which high-order accuracy.*

		ρ		m		E	
N	CFL	L^1 error	Order	L^1 error	Order	L^1 error	Order
20	0.6	5.48E-04	-	7.18E-04	-	3.31E-03	-
40	0.4	1.38E-05	5.31	1.89E-05	5.24	6.68E-05	5.63
80	0.3	4.91E-07	4.81	6.16E-07	4.94	2.13E-06	4.97
160	0.2	1.71E-08	4.84	2.07E-08	4.89	7.05E-08	4.92
320	0.1	6.13E-10	4.80	7.21E-10	4.84	2.23E-09	4.99

Table 3: L^1 errors and orders of convergence in the accuracy test for Example 4.6

Example 4.7. ***Well-balanced test***

In this example, we test the well-balanced property of our scheme for isentropic flow in steady-state.

We consider the linear gravitational filed $\phi(x) = x$ *on the domain* $\Omega = [0, 2]$*, with the initial conditions*

$$s(x,0) = 1, \quad m(x,0) = -M\gamma^{\frac{1}{2}}, \quad Q(x,0) = \frac{1}{2}M^2\gamma + \frac{\gamma}{\gamma-1},$$

where $\gamma = \frac{5}{3}$ *is the ratio of specific heat,* M *is the Mach number at* $x = 0$ *taking different values in the following three cases:*

(a) Hydrostatic flow $M = 0$.

(b) Subcritical flow $M = 0.01$.

(c) Supercritical flow $M = 2.5$.

We compute the solutions up tp $t = 20$ *with* $N = 200$ *cells, and gather the errors in Table 4. We can observe that the steady flow is preserved within machine epsilon during simulation, which confirms the*

well-balanced property of the scheme.

		ρ		m		E	
Case	Precision	L^1 error	L^∞ error	L^1 error	L^∞ error	L^1 error	L^∞ error
	Single	8.95E-07	1.85E-06	4.90E-07	7.86E-07	7.04E-07	1.37E-06
(a)	Double	1.70E-15	3.50E-15	4.88E-16	8.74E-16	1.09E-15	3.11E-15
	Quadruple	1.05E-33	2.02E-33	4.66E-34	9.31E-34	8.41E-34	1.73E-33
	Single	6.93E-07	1.25E-06	3.14E-07	7.15E-07	7.25E-07	1.79E-06
(b)	Double	1.35E-15	2.55E-15	5.63E-16	1.11E-15	1.09E-15	3.11E-15
	Quadruple	1.19E-33	2.89E-33	4.75E-34	9.69E-34	8.97E-34	2.31E-33
	Single	3.28E-06	5.96E-06	7.21E-06	1.62E-05	1.33E-05	3.19E-05
(c)	Double	5.56E-15	8.88E-15	1.30E-14	2.31E-14	2.37E-14	4.35E-14
	Quadruple	4.10E-33	7.51E-33	9.51E-33	1.77E-32	1.86E-32	3.54E-32

Table 4: L^1 and L^∞ errors for different precision in the well-balanced test for Example 4.7

Example 4.8. ***A small perturbation of steady states***

In this example, we test the capability of the well-balanced scheme for capturing small perturbations near equilibria.

We take the initial conditions of density, momentum and pressure of the flow as

$$\rho(x,0) = \rho_e(x,0), \quad m(x,0) = m_e(x,0), \quad p(x,0) = p_e(x,0) + A\exp(-100(x-\bar{x})^2),$$

where $\rho_e(x,0), m_e(x,0)$ *and* $p_e(x,0)$ *are initial conditions in case (a)-(c) of Example 4.7,* $\gamma = \frac{5}{3}$*,* $A = 10^{-6}$*, and*

$$\bar{x} = \begin{cases} 1.0, & M = 0, \\ 1.1, & M = 0.01, \\ 1.5, & M = 2.5. \end{cases}$$

We compute the solutions up to $t = 0.45, 0.45$ *and* 0.25 *for the perturbations of case (a), (b) and (c), respectively, using both the well-balanced scheme and a non-well-balanced scheme. We draw the discrepancy of the density, velocity and pressure with respect to their corresponding base steady states in Figure 10, with*

the reference discrepancy given by the well-balanced scheme on $N = 6400$ cells. As the figures show, the perturbation is resolved much better by the well-balanced scheme on coarse meshes.

Example 4.9. ***Discontinuous wave propogation***

In this example, we test the capability of our the scheme for capturing shocks or large gradients. We increase the magnitude of the perturbation to $A = 1$ and keep all other settings the same as Example 4.8, such that discontinuity develops before the terminal time.

The numerical solutions of our well-balanced scheme at the terminal time with $N = 200$ cells are shown in Figure 13, from which we observe the shocks and solution with large gradients are resolved very well. The reference solutions are obtained by ninth order finite difference WENO scheme with 6400 *cells.*

5 Concluding remarks

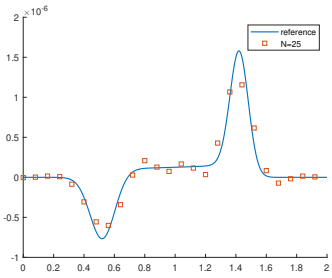
In this paper, we have established a high-order well-balanced finite difference scheme with exact conservation property for hyperbolic balance laws. Based on a flux-gradient reformulation of source terms, we achieve the well-balanced property by discretizing it using the same numerical flux as the true flux gradient. However, the resulting numerical scheme is generally not conservative and computationally costly. To fix this issue, we adopt the A-WENO framework and interpolate the equilibrium variables, which are constant at steady states in the equations we consider. As verified by ample numerical tests, the method effectively preserves the non-hydrostatic equilibria of the shallow water equations with non-flat bottom topography and Euler equations in gravitational filed.

References

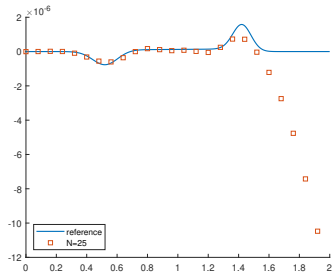
[1] E. Audusse, F. Bouchut, M.-O. Bristeau, R. Klein, and B. Perthame, A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows, SIAM Journal on Scientific Computing, 25, 2004, 2050-2065.

[2] A. Bermudez and M. E. Vazquez, Upwind methods for hyperbolic conservation laws with source terms, Computers & Fluids, 23, 1994, 1049-1071.

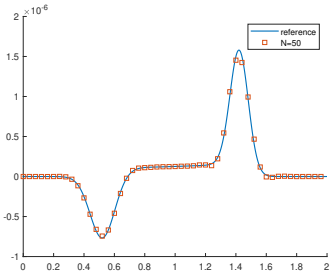
[3] F. Bouchut and T.M. De Luna, A subsonic-well-balanced reconstruction scheme for shallow water flows, SIAM Journal on Numerical Analysis, 48, 2010, 1733-1758.



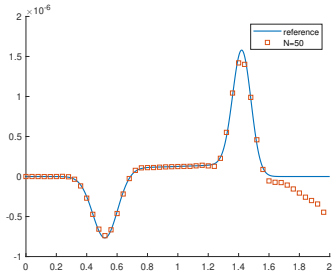
(a) Velocity u, well-balanced scheme, $N = 25$



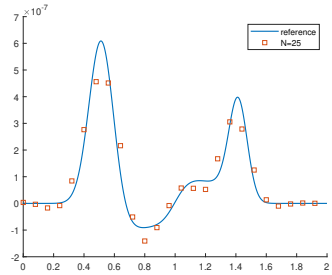
(b) Velocity u, non-well-balanced scheme, $N = 25$



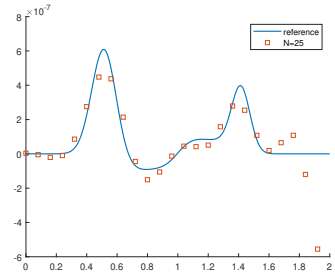
(c) Velocity u, well-balanced scheme, $N = 50$



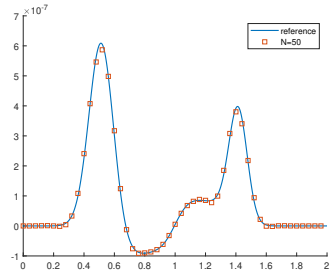
(d) Velocity u, non-well-balanced scheme, $N = 50$



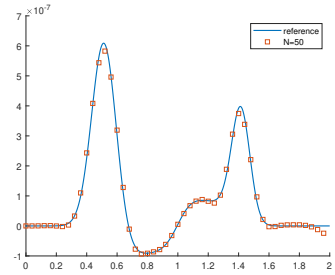
(e) Pressure p, well-balanced scheme, $N = 25$



(f) Pressure p, non-well-balanced scheme, $N = 25$

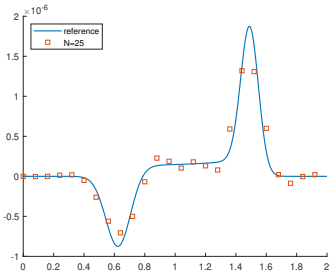


(g) Pressure p, well-balanced scheme, $N = 50$

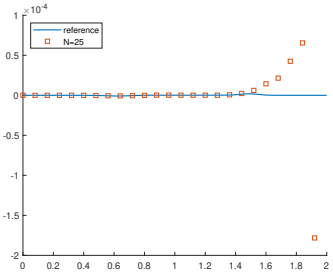


(h) Pressure p, non-well-balanced scheme, $N = 50$

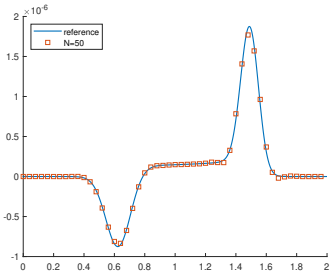
Figure 10: Discrepancy of velocity u and pressure p of case (a) of Example 4.8 with respect to the corresponding base steady state at $t = 0.45$.



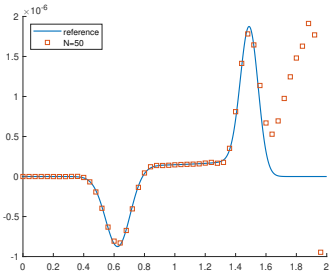
(a) Velocity u, well-balanced scheme, $N = 25$



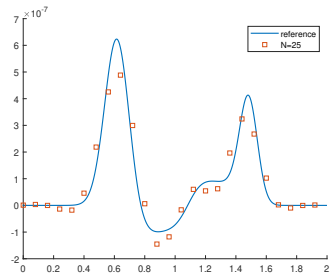
(b) Velocity u, non-well-balanced scheme, $N = 25$



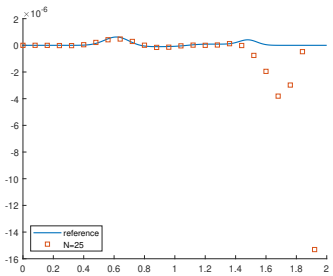
(c) Velocity u, well-balanced scheme, $N = 50$



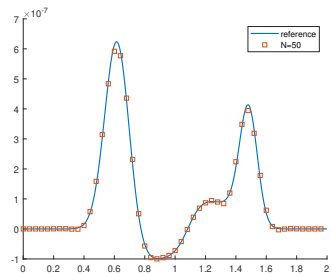
(d) Velocity u, non-well-balanced scheme, $N = 50$



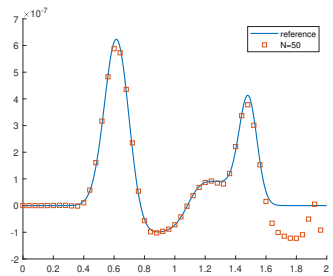
(e) Pressure p, well-balanced scheme, $N = 25$



(f) Pressure p, non-well-balanced scheme, $N = 25$

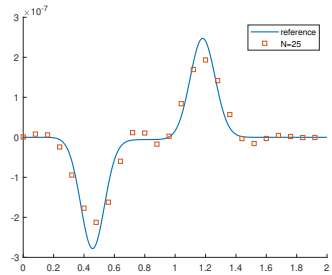


(g) Pressure p, well-balanced scheme, $N = 50$

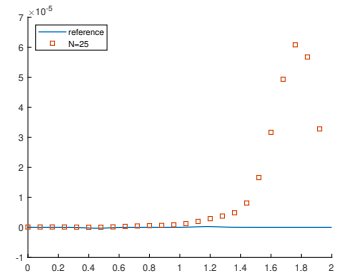


(h) Pressure p, non-well-balanced scheme, $N = 50$

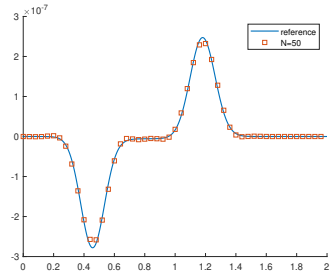
Figure 11: Discrepancy of velocity u and pressure p of case (b) of Example 4.8 with respect to the corresponding base steady state at $t = 0.45$.



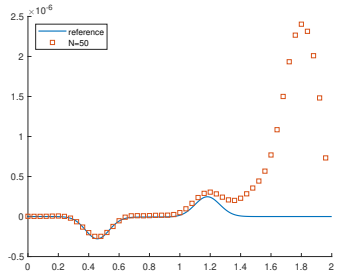
(a) Velocity u, well-balanced scheme, $N = 25$



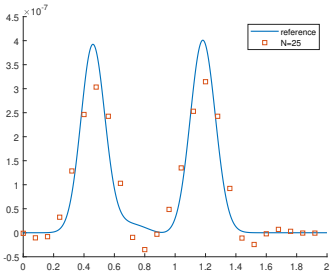
(b) Velocity u, non-well-balanced scheme, $N = 25$



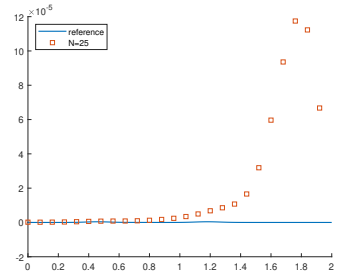
(c) Velocity u, well-balanced scheme, $N = 50$



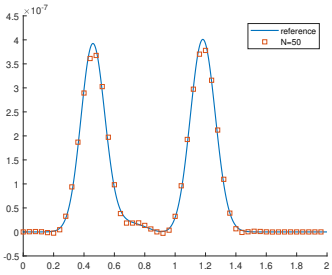
(d) Velocity u, non-well-balanced scheme, $N = 50$



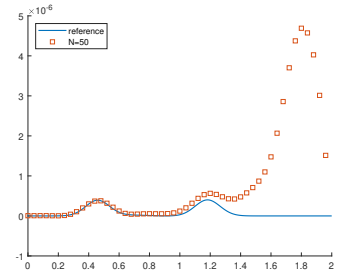
(e) Pressure p, well-balanced scheme, $N = 25$



(f) Pressure p, non-well-balanced scheme, $N = 25$

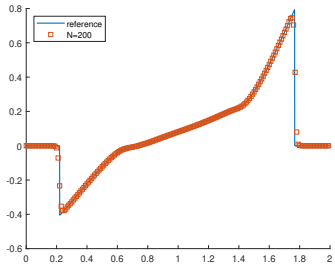


(g) Pressure p, well-balanced scheme, $N = 50$

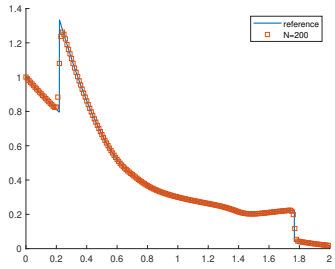


(h) Pressure p, non-well-balanced scheme, $N = 50$

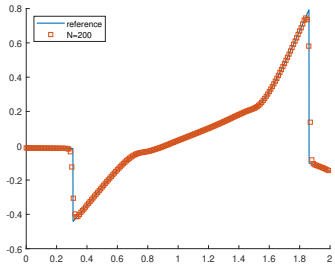
Figure 12: Discrepancy of velocity u and pressure p of case (c) of Example 4.8 with respect to the corresponding base steady state at $t = 0.25$.



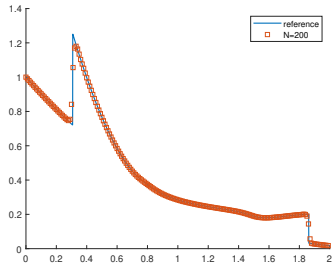
(a) Velocity u of case (a) at $t = 0.45$



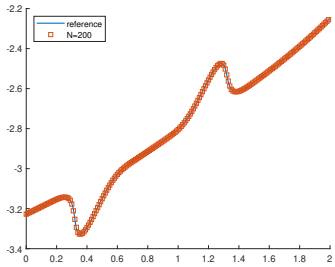
(b) Pressure p of case (a) at $t = 0.45$



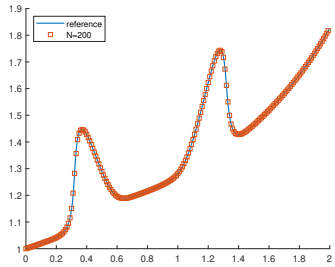
(c) Velocity u of case (b) at $t = 0.45$



(d) Pressure p of case (b) at $t = 0.45$



(e) Velocity u of case (c) at $t = 0.25$



(f) Pressure p of case (c) at $t = 0.25$

Figure 13: Numerical solutions of Example 4.9 with shocks or large gradients

[4] S. Bryson, Y. Epshteyn, A. Kurganov, and G. Petrova, Well-balanced positivity preserving central-upwind scheme on triangular grids for the Saint-Venant system, ESAIM: Mathematical Modelling and Numerical Analysis, 45, 2011, 423-446.

[5] P. Cargo and A.Y. LeRoux, Un schéma équilibre adapté au modèle d'atmosphère avec termes de gravité, Comptes rendus de l'Académie des sciences. Série 1, Mathématique, 1994, 318, 73-76.

[6] E. Carlini, R. Ferretti and G. Russo, weighted essentially nonoscillatory, large time-step scheme for Hamilton–Jacobi equations, SIAM Journal on Scientific Computing, 27, 2005, 1071-1091.

[7] M.J. Castro, J.M. Gallardo, J. López-GarcÍa and C. Parés, Well-balanced high order extensions of Godunov method for semilinear balance laws, SIAM Journal on Numerical Analysis, 46, 2008, 1012–1039.

[8] M. Castro, J. Gallardo and C. Parés, High order finite volume schemes based on reconstruction of states for solving hyperbolic systems with nonconservative products. Applications to shallow-water systems, Mathematics of computation, 75, 2006, 1103-1134.

[9] M. Castro and C. Parés, Well-balanced high-order finite volume methods for systems of balance laws, Journal of Scientific Computing, 82, 2020, 48.

[10] P. Chandrashekar and M. Zenk, Well-balanced nodal discontinuous Galerkin method for Euler equations with gravity, Journal of Scientific Computing, 71, 2017, 1062-1093.

[11] Y. Cheng, A. Chertock, M. Herty, A. Kurganov, and T. Wu, A new approach for designing moving-water equilibria preserving schemes for the shallow water equations, Journal of Scientific Computing, 80, 2019, 538-554.

[12] Y. Cheng and A. Kurganov, Moving-water equilibria preserving central-upwind schemes for the shallow water equations, Communications in Mathematical Sciences, 14, 2016, 1643-1663.

[13] A. Chertock, A. Kurganov, X. Liu, Y. Liu, and T. Wu, Well-balancing via flux globalization: Applications to shallow water equations with wet/dry fronts, Journal of Scientific Computing, 90, 2022, 1-21.

[14] I. Gómez-Bueno, M. Castro and C. Parés, High-order well-balanced methods for systems of balance laws: a control-based approach, Applied Mathematics and Computation, 394, 2021, 125820.

[15] L. Grosheintz-Laval and R. Käppeli, High-order well-balanced finite volume schemes for the Euler equations with gravitation, Journal of Computational Physics, 378, 2019, 324-343.

[16] L. Grosheintz-Laval and R. Käppeli, Well-balanced finite volume schemes for nearly steady adiabatic flows, Journal of Computational Physics, 423, 2020, 109805.

[17] A. Harten, B. Engquist, S. Osher, and S. R. Chakravarthy, Uniformly high order accurate essentially non-oscillatory schemes, III, Journal of Computational Physics, 71, 1987, 213-303.

[18] G. Huang, Y. Xing, and T. Xiong, High order well-balanced asymptotic preserving finite difference WENO schemes for the shallow water equations in all Froude numbers, Journal of Computational Physics, 463, 2022, 111255.

[19] G.-S. Jiang and C.-W. Shu, Efficient implementation of weighted ENO schemes, Journal of Computational Physics, 126, 1996, 202-228.

[20] Y. Jiang, C.-W. Shu, and M. Zhang, An alternative formulation of finite difference weighted ENO schemes with Lax–Wendroff time discretization for conservation laws, SIAM Journal on Scientific Computing, 35, 2013, A1137-A1160.

[21] R. Käppeli, Well-balanced methods for computational astrophysics, Living Reviews in Computational Astrophysics, 8, 2022, 2.

[22] R. Käppeli and S. Mishra, Well-balanced schemes for the Euler equations with gravitation, Journal of Computational Physics, 259, 2014, 199-219.

[23] C. Klingenberg, A. Kurganov, and M. Zenk, Moving-water equilibria preserving HLL-type schemes for the shallow water equations, Commun. Math. Res., 36, 2020, 247-271.

[24] C. Klingenberg, G. Puppo, and M. Semplice, Arbitrary order finite volume well-balanced schemes for the Euler equations with gravity, SIAM Journal on Scientific Computing, 41, 2019, A695-A721.

[25] A. Kurganov and D. Levy, Central-upwind schemes for the Saint-Venant system, ESAIM: Mathematical Modelling and Numerical Analysis, 36, 2002, 397-425.

[26] A. Kurganov and G. Petrova, A second-order well-balanced positivity preserving central-upwind scheme for the Saint-Venant system, Communications in Mathematical Sciences, 5, 2007, 133-160.

[27] W.D. Lambert, A generalized trigonometric solution of the cubic equation, The American Mathematical Monthly, 13, 1906, 73-76.

[28] R.J. LeVeque, Balancing source terms and flux gradients in high-resolution Godunov methods: the quasi-steady wave-propagation algorithm, Journal of Computational Physics, 146, 1998, 346-365.

[29] G. Li and Y. Xing, Well-balanced discontinuous Galerkin methods for the Euler equations under gravitational fields, Journal of Scientific Computing, 67, 2016, 493-513.

[30] G. Li and Y. Xing, Well-balanced discontinuous Galerkin methods with hydrostatic reconstruction for the Euler equations with gravitation, Journal of Computational Physics, 352, 2018, 445-462.

[31] G. Li and Y. Xing, Well-balanced finite difference weighted essentially non-oscillatory schemes for the Euler equations with static gravitational fields, Computers & Mathematics with Applications, 75, 2018, 2071-2085.

[32] X. Liu, X. Chen, S. Jin, A. Kurganov, T. Wu, and H. Yu, Moving-water equilibria preserving partial relaxation scheme for the Saint-Venant system, SIAM Journal on Scientific Computing, 42, 2020, A2206-A2229.

[33] F. Marche, P. Bonneton, P. Fabrie and N. Seguin, Evaluation of well-balanced bore-capturing schemes for 2D wetting and drying processes, International Journal for Numerical Methods in Fluids, 53, 2007, 867-894.

[34] S. Noelle, N. Pankratz, G. Puppo, and J.R. Natvig, Well-balanced finite volume schemes of arbitrary order of accuracy for shallow water flows, Journal of Computational Physics, 213, 2006, 474-499.

[35] S. Noelle, Y. Xing, and C.-W. Shu, High-order well-balanced finite volume WENO schemes for shallow water equation with moving water, Journal of Computational Physics, 226, 2007, 29-58.

[36] C. Parés and C. Parés-Pulido, Well-balanced high-order finite difference methods for systems of balance laws, Journal of Computational Physics, 425, 2021, 109880.

[37] S. Qian, G. Li, F. Shao, and Y. Xing, Positivity-preserving well-balanced discontinuous Galerkin methods for the shallow water flows in open channels, Advances in Water Resources, 115, 2018, 172-184.

[38] J. Qiu and C.-W. Shu, On the construction, comparison, and local characteristic decomposition for high-order central WENO schemes, Journal of Computational Physics, 183, 2002, 187-209.

[39] P. Roe, Approximate Riemann solvers, parameter vectors, and difference schemes, Journal of Computational Physics, 27, 1978, 1-31.

[40] K. Sebastian and C.-W. Shu, Multidomain WENO finite difference method with interpolation at subdomain interfaces, Journal of Scientific Computing, 19, 2003, 405-438.

[41] C.-W. Shu and S. Osher, Efficient implementation of essentially non-oscillatory shock-capturing schemes, Journal of Computational Physics, 77, 1988, 439-471.

[42] C.-W. Shu and S. Osher, Efficient implementation of essentially non-oscillatory shock-capturing schemes, II, Journal of Computational Physics, 83, 1989, 32-78.

[43] R. Touma, U. Koley, and C. Klingenberg, Well-balanced unstaggered central schemes for the Euler equations with gravitation, SIAM Journal on Scientific Computing, 38, 2016, B773-B807.

[44] M.E. Vázquez-Cendón, Improved treatment of source terms in upwind schemes for the shallow water equations in channels with irregular geometry, Journal of computational physics, 148, 1999, 497-526.

[45] K. Wu and Y. Xing, Uniformly high-order structure-preserving discontinuous Galerkin methods for Euler equations with gravitation: Positivity and well-balancedness, SIAM Journal on Scientific Computing, 43, 2021, A472-A510.

[46] Y. Xing, Exactly well-balanced discontinuous Galerkin methods for the shallow water equations with moving water equilibrium, Journal of Computational Physics, 257, 2014, 536-553.

[47] Y. Xing and C.-W. Shu, High order finite difference WENO schemes with the exact conservation property for the shallow water equations, Journal of Computational Physics, 208, 2005, 206-227.

[48] Y. Xing and C.-W. Shu, High-order well-balanced finite difference WENO schemes for a class of hyperbolic systems with source terms, Journal of Scientific Computing, 27, 2006, 477-494.

[49] Y. Xing and C.-W. Shu, High order well-balanced finite volume WENO schemes and discontinuous Galerkin methods for a class of hyperbolic systems with source terms, Journal of Computational Physics, 214, 2006, 567-598.

[50] Y. Xing and C.-W. Shu, A new approach of high order well-balanced finite volume WENO schemes and discontinuous Galerkin methods for a class of hyperbolic systems with source terms, Communications in Computational Physics, 1, 2006, 100-134.

[51] Y. Xing and C.-W. Shu, High order well-balanced WENO scheme for the gas dynamics equations under gravitational fields, Journal of Scientific Computing, 54, 2013, 645-662.

[52] Y. Xing, C.-W. Shu, and S. Noelle, On the advantage of well-balanced schemes for moving-water equilibria of the shallow water equations, Journal of Scientific Computing, 48, 2011, 339-349.

[53] Y. Xing, X. Zhang, and C.-W. Shu, Positivity-preserving high order well-balanced discontinuous Galerkin methods for the shallow water equations, Advances in Water Resources, 33, 2010, 1476-1493.

[54] Z. Xu and C.-W. Shu, Local characteristic decomposition free high order finite difference WENO schemes for hyperbolic systems endowed with a coordinate system of Riemann invariants, submit for review

[55] J. Zhang, Y. Xia, and Y. Xu, Moving water equilibria preserving discontinuous Galerkin method for the shallow water equations, Journal of Scientific Computing, 95, 2023, 48.