Steady States and Well-balanced Schemes for Shallow Water Moment Equations with Topography

Julian Koellermeier,* Ernesto Pimentel-García†

November 16, 2020

Abstract

In this paper, we investigate steady states of shallow water moment equations including bottom topographies. We derive a new hyperbolic shallow water moment model based on linearized moment equations that allows for a simple assessment of the steady states. After proving hyperbolicity of the new model, the steady states are fully identified. A well-balanced scheme is adopted to the specific structure of the new model and allows to preserve the steady states in numerical simulations.

Keywords: Shallow Water Equations, hyperbolic moment equations, well-balanced, steady states

1 Introduction

Applications of shallow flows can be found in many scientific fields, e.g., in hydrodynamics [31] or granular flows [16]. An important class of problems considers changing topographies, for example related to snow avalanches [15] or sediment transport [17]. The main assumption for the widely used Shallow Water Equations (SWE) is that the horizontal velocity profile is constant along the vertical axis from the bottom to the surface. However, this assumption quickly brakes down for more complex flows that yield velocity variations. This is true in practically all applications of shallow flows and especially in presence of friction terms. But even in typical tsunami or dam break situations, the assumption of constant velocity profiles is often violated, see [21]. A new model that takes into account horizontal velocity changes over the vertical direction was developed in [23] based on an expansion of the velocity profile in polynomial basis functions modeling the deviation from a constant velocity profile. The resulting Shallow Water Moment Equations (SWME) are more accurate the more basis functions are considered. Despite the success for simple test cases, the model lacks hyperbolicity, which was studied in detail in [21]. In the same paper, a new model called Hyperbolic Shallow Water Moment Equations (HWSME) and a second variant called the β-HSWME were derived. The models are essentially based on a linearization of the original SWME model around linear velocity profiles, i.e., all contributions of coefficients of higher order basis functions are neglected. In [21] the eigenvalues of these models were analyzed and the first numerical tests confirmed that the models yield similar accuracy as the SWME models with additional guaranteed hyperbolicity.

While the numerical tests in [23, 21] included standard friction terms for a Newtonian fluid, only a flat bottom topography was considered. This is obviously a strong simplification and bottom topographies need to be taken into account as has been done for the SWE in a numerical and analytical way, see [1, 5, 29, 25] and the references therein. In the context of varying bottom

---

*Department of Computer Science, KU Leuven, 3001 Leuven, Belgium, email: julian.koellermeier@kuleuven.be

†Dpto. Análisis Matemático. Universidad de Málaga, 14071 Málaga, Spain

topographies, it is paramount to consider steady states of the models because any numerical simulation should be able to exactly preserve steady states when present. Otherwise, numerical solutions starting from steady state initial conditions would lead to numerical artifacts or numerical instabilities. It is therefore important to first study the steady states of the models and then design tailored well-balanced numerical schemes, which means that the schemes preserve those steady states by balancing the topography source term and the numerical flux in the correct way so they cancel out. Since [4], the study and design of well-balanced numerical methods have been very active fields in the last years, see for instance [2, 3, 7, 12, 8, 26, 14, 30]. In the context of path-conservative methods introduced in [27], the authors in [10] and more recently in [13] developed a strategy to obtain well-balanced high-order numerical methods for systems of balance laws. We will follow this strategy and apply it to a newly derived moment model.

In this paper, we investigate steady states of shallow water moment equations including bottom topographies and use this to derive a new first order and second order well-balanced numerical scheme for a new shallow water moment model. The analysis of the existing SWME, including the hyperbolic versions HSWME and β-HSWME, shows that steady states are difficult to access analytically and numerically, despite the simple case where the velocity profile is just a linear function of the vertical variable. Knowing about the problematic terms in the existing models, we derive a new model that is valid for small deviations from the constant velocity profile. For this model, we can neglect only the non-linear contributions of the basis coefficients while keeping the linear contributions of all coefficients. The model is thus called Shallow Water Linearized Moment Equations (SWLME). It is surprisingly simple, in the sense that it removes some coupling terms between the equations, but it keeps the overall structure even in the higher order equations. Subsequently, we prove hyperbolicity, analyze the eigenstructure, and show that the model yields more realistic propagation speeds than the previous models, while still being hyperbolic. Most importantly, the model allows for a concise characterization of its steady states with and without topography terms. The characterization of the steady states then allows to derive a potentially high-order well-balanced numerical scheme based on the possible steady states of the new model. We explicitly construct the first order and second order well-balanced scheme in this paper. The numerical schemes are tested extensively with a standard lake-at-rest test case, two subcritical stationary solutions, and a transcritical solution. In the end, we also present a test case comparing the new SWLME to the existing HSWME and β-HSWME models, to outline the good approximation properties of the new model despite its simplicity.

The rest of the paper is organized as follows: In Section 2 we review the derivation of a vertically resolved shallow flow model that is employed to derive all the shallow water moment models presented in this paper. In the following sections we derive and analyze the standard Shallow Water Equations (SWE) (Section 3), the extended Shallow Water Moment Equations (SWME) (Section 4), and the new Shallow Water Linearized Moment Equations (SWLME) (Section 5) including their hyperbolicity, steady states, and Rankine-Hugoniot conditions including bottom topography. In Section 6, we develop a first order and second order well-balanced numerical scheme for the special case of the shallow water models used in this paper. Numerical tests in Section 7 show the preservation of the steady states and allow for a comparison of the new SWLME model with respect to other existing models.

2 Vertically resolved shallow flow model

In this paper, we are concerned with free-surface water flows in one horizontal direction. Modeling of free-surface flows starts with the incompressible Navier-Stokes equations, which model the evolution of the horizontal velocity u in direction x and the vertical velocity w in direction z.

$$\partial_x u + \partial_z w = 0, \tag{2.1}$$

$$\partial_t u + \partial_x u^2 + \partial_z uw = -\frac{1}{\rho}\partial_x p + \frac{1}{\rho}\partial_z \sigma_{xz} + g, \tag{2.2}$$

where ρ is the density and g the gravitation constant. The hydrostatic pressure in relation to the vertical position z with respect to the surface $h + b$, where b represents the bottom topography

and h is the water height, is given by

$$p = (h + b - z)\rho g \tag{2.3}$$

and the stress σ_{xz} is modeled using the assumption of a Newtonian fluid with dynamic viscosity μ, i.e.,

$$\sigma_{xz} = \mu \partial_z u \tag{2.4}$$

to close the system.

To allow for a more efficient representation of the horizontal velocity variation along the vertical axis, a mapping is introduced in [23]. This mapping shifts and scales the vertical variable, which is defined between the bottom at $z = b$ and the surface at $z = h + b$ according to the following transformation

$$\zeta = \frac{z - b}{h}, \tag{2.5}$$

where the denominator is precisely the water height h. The variable ζ is thus defined within the interval $[0, 1]$. According to the derivation in [23], the following vertically-resolved system for the simulation of shallow flows is derived using the mapping from (2.5)

$$\partial_t h + \partial_x h u_m \quad = \quad 0, \tag{2.6}$$

$$\partial_t h u + \partial_x \left(h u^2 + \frac{g}{2} h^2 \right) + \partial_\zeta \left(h u \omega - \frac{1}{\rho} \sigma_{xz} \right) \quad = \quad -g h \partial_x b, \tag{2.7}$$

where u_m is the mean velocity over the vertical ζ-axis and the so-called vertical coupling term ω is given by

$$\omega = \int_0^\zeta \left(\int_0^1 \partial_x (hu) \left(\hat{\zeta} \right) d\hat{\zeta} - \partial_x (hu) \left(\widetilde{\zeta} \right) \right) d\widetilde{\zeta}. \tag{2.8}$$

The following boundary conditions in the ζ-direction are used:

$$\partial_\zeta u |_{\zeta=1} \quad = \quad 0, \tag{2.9}$$

$$\partial_\zeta u |_{\zeta=0} \quad = \quad \frac{h}{\lambda} u |_{\zeta=0}, \tag{2.10}$$

modeling a stress-free top surface and a slip condition at the bottom with slip length λ, see [23] for more details.

The system (2.6)-(2.7) is called *vertically resolved system* in [23] as it includes the dependence on the vertical variable ζ. This system is at the core of this paper as all the models are derived directly from it.

3   Shallow Water Equations

Similar to the work in [23], we will start with the simple Shallow Water Equations (SWE), which assume a constant velocity $u(t, x, \zeta) = u_m(t, x)$ over the whole vertical axis ζ, see Figure 1a. The dependency on the vertical variable ζ is then resolved by integrating over $\zeta \in [0, 1]$ and using the constant velocity $u(t, x, \zeta) = u_m(t, x)$. It was shown in [23] that the vertically resolved system (2.6)-(2.7) then simplifies to the following set of equations called Shallow Water Equations (SWE)

$$\partial_t \begin{pmatrix} h \\ h u_m \end{pmatrix} + \partial_x \begin{pmatrix} h u_m \\ h u_m^2 + \frac{1}{2} g h^2 \end{pmatrix} = \begin{pmatrix} 0 \\ -g h \partial_x b \end{pmatrix} - \frac{\nu}{\lambda} \begin{pmatrix} 0 \\ u_m \end{pmatrix} \tag{3.1}$$

where $u_m = u_m(t, x)$ is the horizontal water velocity, $h = h(t, x)$ is the water height, g is the gravitational constant (we later set it to $g = 1$ in our simulations) the known function $b(x)$ is the bottom topography, and ν and λ are the kinematic viscosity and the slip length, respectively.

In non-conservative matrix form, the model can be written as

$$\partial_t \begin{pmatrix} h \\ h u_m \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ -u_m^2 + g h & 2 u_m \end{pmatrix} \partial_x \begin{pmatrix} h \\ h u_m \end{pmatrix} = \begin{pmatrix} 0 \\ -g h \partial_x b \end{pmatrix} - \frac{\nu}{\lambda} \begin{pmatrix} 0 \\ u_m \end{pmatrix}. \tag{3.2}$$

3
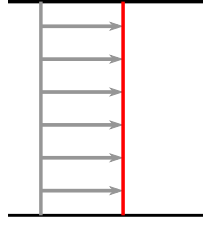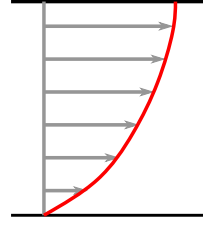
(a) Constant velocity profile (b) Varying velocity profile

Figure 1: Constant velocity ansatz of SWE model (a) and variable velocity ansatz of SWME model (b).

The eigenvalues of the left hand side transport matrix are the standard propagation speeds of the Shallow Water Equations

$$\lambda_{1,2} = u_m \pm \sqrt{gh}. \tag{3.3}$$

For flat bottom $\partial_x b = 0$ and zero friction, the steady state fulfils

$$\partial_x (hu_m) = 0, \tag{3.4}$$

$$\partial_x \left(hu_m^2 + \frac{1}{2}gh^2 \right) = 0, \tag{3.5}$$

so that the jump conditions (also called Rankine-Hugoniot conditions) from a given state $(h_0, h_0 u_{m,0})$ to a state (h, hu_m) can be derived by solving the system

$$hu_m = h_0 u_{m,0}, \tag{3.6}$$

$$hu_m^2 + \frac{1}{2}gh^2 = h_0 u_{m,0}^2 + \frac{1}{2}gh_0^2, \tag{3.7}$$

for which the solution is

$$\left(\frac{h}{h_0} \right) = -\frac{1}{2} + \frac{1}{2} \cdot \sqrt{1 + 8Fr^2}, \tag{3.8}$$

where Fr is the Froude number for the given state defined by

$$Fr = \frac{u_{m,0}}{\sqrt{gh_0}}. \tag{3.9}$$

For a smooth frictionless flow including a bottom topography, the steady state momentum equation can be modified using the mass equation (3.4) to

$$\partial_x \left(\frac{1}{2}u_m^2 + g(h + b) \right) = 0. \tag{3.10}$$

The steady state solution can thus be found using

$$hu_m = const, \tag{3.11}$$

$$\frac{1}{2}u_m^2 + g(h + b) = const. \tag{3.12}$$

The SWE are widely used in simulations of water flows. However, the main deficiency is that the horizontal velocity u is constant over the height by assumption. The model is thus not able to predict more complex flow phenomena.

4

4 Shallow Water Moment Equations

For the Shallow Water Moment Equations (SWME) derived in [23], the idea is to allow for a vertical variation of the water velocity profile. This is done by assuming the following ansatz for the velocity profile, see Figure 1b:

$$u(t,x,\zeta) = u_m(t,x) + \sum_{j=1}^{N} \alpha_j(t,x)\phi_j(\zeta), \tag{4.1}$$

where $u_m(t,x)$ is the mean horizontal velocity also used in the SWE in Section 3, ζ is the scaled vertical coordinate (2.5), α_j are coefficients, and ϕ_j are Legendre ansatz functions for $j = 1,\dots,N$ defined by

$$\phi_j(\zeta) = \frac{1}{j!}\frac{d^j}{d\zeta^j}(\zeta - \zeta^2)^j. \tag{4.2}$$

Note that the larger N, the more variation is allowed in vertical direction. Furthermore, the ansatz functions form a group of orthogonal basis functions as [23]

$$\int_0^1 \phi_m\phi_n d\zeta = \frac{1}{2n+1}\delta_{mn},$$

with Kronecker delta $\delta_{m,n}$.

The initial values for u_m and α_j for $j = 1,\dots,N$ can be computed from some initial velocity profile $u(0,x,\zeta) = u_0(x,\zeta)$ by projecting the initial velocity profile to the basis functions ϕ_j.

$$u_0(x,\zeta) = u_m(0,x) + \sum_{j=1}^{N}\alpha_j(0,x)\phi_j(\zeta) \tag{4.3}$$

$$\int_0^1 u_0(x,\zeta)\phi_i(\zeta)\,d\zeta = \int_0^1 \left(u_m(0,x) + \sum_{j=1}^{N}\alpha_j(0,x)\phi_j(\zeta)\right)\phi_i(\zeta)\,d\zeta \tag{4.4}$$

$$= u_m(0,x)\delta_{i,0} + \sum_{j=1}^{N}\alpha_j(0,x)\frac{1}{2i+1}\delta_{i,j}, \tag{4.5}$$

which leads to the initial mean and coefficients

$$u_m(0,x) = \int_0^1 u_0(x,\zeta)\,d\zeta, \tag{4.6}$$

$$\alpha_i(0,x) = (2i+1))\int_0^1 u_0(x,\zeta)\phi_i(\zeta)\,d\zeta \text{ for } i = 1,\dots,N. \tag{4.7}$$

The model for the evolution of the coefficients for arbitrary N can be derived by inserting the ansatz (4.1) into the vertically resolved system (2.6)-(2.7) and integrating over $\zeta \in [0,1]$. According to [23] this leads to

$$\partial_t h + \partial_x hu_m = 0, \tag{4.8}$$

$$\partial_t hu_m + \partial_x\left(hu_m^2 + h\sum_{j=1}^{N}\frac{\alpha_j^2}{2j+1} + \frac{g}{2}h^2\right) = -\frac{\nu}{\lambda}\left(u_m + \sum_{j=1}^{N}\alpha_j\right) - gh\partial_x b, \tag{4.9}$$

$$\partial_t h\alpha_i + \partial_x\left(2hu_m\alpha_i + h\sum_{j,k=1}^{N}A_{ijk}\alpha_j\alpha_k\right) = u_m\partial_x h\alpha_i - \sum_{j,k=1}^{N}B_{ijk}\partial_x(h\alpha_j)\alpha_k \tag{4.10}$$

$$-(2i+1)\frac{\nu}{\lambda}\left(u_m + \sum_{j=1}^{N}\left(1 + \frac{\lambda}{h}C_{ij}\right)\alpha_j\right), \tag{4.11}$$

5

for $i = 1, \ldots, N$, the unknown variables $(h, u, \alpha_1, \ldots, \alpha_N)$ and

$$A_{ijk} = (2i+1) \int_0^1 \phi_i \phi_j \phi_k \, d\zeta, \tag{4.12}$$

$$B_{ijk} = (2i+1) \int_0^1 \partial_\zeta \phi_i \left(\int_0^\zeta \phi_j \, d\hat{\zeta} \right) \phi_k \, d\zeta, \tag{4.13}$$

$$C_{ij} = \int_0^1 \partial_\zeta \phi_i \partial_\zeta \phi_j \, d\zeta. \tag{4.14}$$

The model can be written in closed form with the precomputed terms A_{ijk}, B_{ijk}, C_{ij} for large N. We then write it as

$$\partial_t W + \frac{\partial F}{\partial W} \partial_x W = Q \partial_x W + P, \tag{4.15}$$

with variables $W = (h, hu_m, h\alpha_1, \ldots, h\alpha_N)^T \in \mathbb{R}^{N+2}$, the flux Jacobian (also called conservative matrix) $\frac{\partial F}{\partial W}$ given by

$$\frac{\partial F}{\partial W} = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ gh - u_m^2 - \sum_{i=1}^{N} \frac{\alpha_i}{2i+1} & 2u_m & \frac{2\alpha_1}{2\cdot1+1} & \ldots & \frac{2\alpha_N}{2N+1} \\ -2u_m\alpha_1 - \sum_{j,k=1}^{N} A_{1jk}\alpha_j\alpha_k & 2\alpha_1 & 2u_m\delta_{11} + 2\sum_{k=1}^{N} A_{11k}\alpha_k & \ldots & 2u_m\delta_{1N} + 2\sum_{k=1}^{N} A_{1Nk}\alpha_k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -2u_m\alpha_N - \sum_{j,k=1}^{N} A_{Njk}\alpha_j\alpha_k & 2\alpha_N & 2u_m\delta_{NN} + 2\sum_{k=1}^{N} A_{N1k}\alpha_k & \ldots & 2u_m\delta_{NN} + 2\sum_{k=1}^{N} A_{NNk}\alpha_k \end{pmatrix},$$

and right-hand side non-conservative matrix Q

$$Q = \begin{pmatrix} 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & u_m\delta_{11} + \sum_{k=1}^{N} B_{11k}\alpha_k & \ldots & u_m\delta_{1N} + \sum_{k=1}^{N} B_{1Nk}\alpha_k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & u_m\delta_{N1} + \sum_{k=1}^{N} B_{N1k}\alpha_k & \ldots & u_m\delta_{NN} + \sum_{k=1}^{N} B_{NNk}\alpha_k \end{pmatrix},$$

with Kronecker delta δ_{ij}. The friction term P on the right-hand side is defined in [23] as first entry $P_0 = 0$ and

$$P_i = -(2i+1) \frac{\nu}{\lambda} \left(u_m + \sum_{j=1}^{N} \left(1 + \frac{\lambda}{h} C_{ij} \right) \alpha_j \right), i = 1, \ldots, N+1. \tag{4.16}$$

The friction term can be given explicitly as

$$P_i = -\frac{\nu}{\lambda}(2i+1) \left(u_m + \sum_{j=1}^{N} \alpha_j \right) - \frac{\nu}{h} 4(2i+1) \sum_{j=1}^{N} a_{i,j}\alpha_j, i = 1, \ldots, N+1, \tag{4.17}$$

where the constants $a_{i,j}$ are computed by

$$a_{i,j} = \begin{cases} 0 & \text{if } i+j = \text{even}, \\ \frac{\min(i-1,j)(\min(i-1,j)+1)}{2} & \text{if } i+j = \text{odd}. \end{cases} \tag{4.18}$$

Note that the right-hand side friction term can become quite stiff for large N, even though the friction coefficients λ, ν can be of order 1. This should be accounted for by appropriate numerical methods, e.g. Projective Integration [18, 24]. For most of this work, we will neglect the friction terms but consider non-zero topography changes $\partial_x b$.

All the models covered in this paper use the form (4.15) for different simplifications of the conservative and non-conservative matrix.

We first consider the case $N = 1$, also called the first order system. This model is described in [23] and [21]. The velocity profile is then given depending on the mean velocity u_m and the coefficient $\alpha = \alpha_1$ as

$$u(t, x, z) == u_m(t, x) + \left(1 - 2\frac{z - b}{h}\right)\alpha(t, x). \tag{4.19}$$

Note that the two values for the velocity at the top ($z = b + h$) and at the bottom ($z = b$) are given by

$$u(z = b + h) = u_m - \alpha, \tag{4.20}$$
$$u(z = b) = u_m + \alpha. \tag{4.21}$$

It seems reasonable, to require $u(z)$ to have the same sign over the whole velocity profile. Otherwise, the flow can no longer be approximated by means of a *shallow* model assumption, as a vortex could form. Thus we require in this paper

$$|\alpha(t, x)| \leq |u(t, x)|. \tag{4.22}$$

Compare Figure 2.

(a) no change of sign   (b) change of sign

Figure 2: Velocity profile without change of sign (a) and with change of sign (b). We only consider velocity profiles without change of sign in this paper.

Choosing this linear velocity change with the vertical variable, the first order shallow water moment model reads [23]

$$\partial_t \begin{pmatrix} h \\ hu_m \\ h\alpha_1 \end{pmatrix} + \partial_x \begin{pmatrix} hu_m \\ hu_m^2 + \frac{1}{2}gh^2 + \frac{1}{3}h\alpha_1^2 \\ 2hu_m\alpha_1 \end{pmatrix} = Q\partial_x \begin{pmatrix} h \\ hu_m \\ h\alpha_1 \end{pmatrix} - \begin{pmatrix} 0 \\ gh\partial_x b \\ 0 \end{pmatrix} - \frac{\nu}{\lambda}P, \tag{4.23}$$

with

$$Q = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & u_m \end{pmatrix}, P = \begin{pmatrix} 0 \\ u_m + \alpha_1 \\ 3\left(u_m + \alpha_1 + 4\frac{\lambda}{h}\alpha_1\right) \end{pmatrix} \text{ and Jacobian } \frac{\partial F}{\partial V} = \begin{pmatrix} 0 & 1 & 0 \\ -u_m^2 - \frac{\alpha_1^2}{3} + gh & 2u_m & \frac{2\alpha_1}{3} \\ -2u_m\alpha_1 & 2\alpha_1 & 2u_m \end{pmatrix},$$

leading to the system matrix

$$A = \frac{\partial F}{\partial V} - Q = \begin{pmatrix} 0 & 1 & 0 \\ -u_m^2 - \frac{\alpha_1^2}{3} + gh & 2u_m & \frac{2\alpha_1}{3} \\ -2u_m\alpha_1 & 2\alpha_1 & u_m \end{pmatrix}. \tag{4.24}$$

7

The first order system has the distinct real eigenvalues

$$\lambda_{1,2} = u_m \pm \sqrt{gh + \alpha_1^2} \text{ and } \lambda_3 = u_m. \tag{4.25}$$

For positive water height $h > 0$, the first order shallow water moment model is hyperbolic.

So far, there has been no analysis of the first order system except for the eigenvalues in [21, 23]. In this paper, we investigate the steady state of the model.

For flat bottom $\partial_x b = 0$ and zero friction, the steady state fulfills

$$\partial_x (hu_m) = 0, \tag{4.26}$$

$$\partial_x \left(hu_m^2 + \frac{1}{2}gh^2 + \frac{1}{3}h\alpha^2 \right) = 0, \tag{4.27}$$

$$\partial_x (2hu_m\alpha) = u_m \partial_x (h\alpha), \tag{4.28}$$

From the first and last equation, we obtain after some modification

$$hu_m = const, \tag{4.29}$$

$$u_m = 0 \text{ or } \frac{\alpha}{h} = const. \tag{4.30}$$

Using those relations in the remaining second equation, we can derive the Rankine-Hugoniot conditions from a given state $(h_0, h_0 u_{m,0}, h_0 \alpha_0)$ to a state $(h, hu_m, h\alpha)$ and obtain (after some modifications)

$$(h - h_0) \left[-\frac{u_{m,0}^2}{gh_0} + \frac{1}{2} \left(\left(\frac{h}{h_0}\right)^2 + \left(\frac{h}{h_0}\right) \right) + \frac{1}{3}\frac{\alpha_0^2}{gh_0} \left(\left(\frac{h}{h_0}\right)^3 + \left(\frac{h}{h_0}\right)^2 + \left(\frac{h}{h_0}\right) \right) \right] = 0. \tag{4.31}$$

We now use the following dimensionless flow numbers:

$$Fr = \frac{u_{m,0}}{\sqrt{gh_0}}, \tag{4.32}$$

$$M\alpha = \frac{\alpha_0}{u_{m,0}}, \tag{4.33}$$

and write $y = \frac{h}{h_0}$ to arrive at the two solutions

$$h = h_0 \quad \vee \quad -Fr^2 + \frac{1}{2}\left(y^2 + y\right) + \frac{1}{3}M\alpha^2 Fr^2\left(y^3 + y^2 + y\right) = 0. \tag{4.34}$$

That means that the jump conditions for the SWME with $N = 1$ lead to a third order polynomial with two parameters which are the flow numbers Fr and $M\alpha$, a consistent extension from the standard case of the Shallow Water Equations. The new parameter $M\alpha$ measures how far away the flow is from the standard shallow water model. For $M\alpha = 0$, the shallow water equations are recovered with a constant velocity profile, whereas for $|M\alpha| = 1$, the flow velocity is changing the most along the z-axis. For values $|M\alpha| > 1$, the assumption (4.22) is no longer fulfilled and the model assumption of a shallow flow is not valid any more.

Note that the third order polynomial in (4.34) always has at least one real zero.

For a smooth frictionless flow including a bottom topography, the steady state momentum equation can be modified using the mass equation to

$$\partial_x \left(\frac{1}{2}u_m^2 + g(h + b) + \frac{1}{2}\alpha^2 \right) = 0. \tag{4.35}$$

The non-trivial steady state solution can thus be found using

$$hu_m = const, \tag{4.36}$$

$$\frac{1}{2}u_m^2 + g(h + b) + \frac{1}{2}\alpha^2 = const, \tag{4.37}$$

$$\frac{\alpha}{h} = const. \tag{4.38}$$

In Section 6, we will use this form of the non-trivial steady state solution to preserve steady
states within the numerical scheme.

Unfortunately, it is not possible to extend the study of steady states of the SWME form $N = 1$ to $N > 1$. The first problem is that the SWME loose hyperbolicity for $N > 1$ as analyzed in detail in [21]. Hyperbolicity is a mathematical requirement for first order partial differential equations to be robust against small perturbations of the initial data, a key property of the real-world physical processes [32]. The model is only hyperbolic for certain states depending on the values of the coefficients α_i. As one example, consider the case $N = 2$. This so-called second order moment model is given by

$$\partial_t \begin{pmatrix} h \\ hu_m \\ h\alpha_1 \\ h\alpha_2 \end{pmatrix} + \partial_x \begin{pmatrix} hu_m \\ hu_m^2 + g\frac{h^2}{2} + \frac{1}{3}h\alpha_1^2 + \frac{1}{5}h\alpha_2^2 \\ 2hu_m\alpha_1 + \frac{4}{5}h\alpha_1\alpha_2 \\ 2hu_m\alpha_2 + \frac{2}{3}h\alpha_1^2 + \frac{2}{7}h\alpha_2^2 \end{pmatrix} = Q\partial_x \begin{pmatrix} h \\ hu_m \\ h\alpha_1 \\ h\alpha_2 \end{pmatrix} - \begin{pmatrix} 0 \\ gh\partial_x b \\ 0 \\ 0 \end{pmatrix} - \frac{\nu}{\lambda}P \qquad (4.39)$$

with

$$Q = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & u_m - \frac{\alpha_2}{5} & \frac{\alpha_1}{5} \\ 0 & 0 & \alpha_1 & u_m + \frac{\alpha_2}{7} \end{pmatrix} \text{ and } P = \begin{pmatrix} 0 \\ u_m + \alpha_1 + \alpha_2 \\ 3\left(u_m + \alpha_1 + \alpha_2 + 4\frac{\lambda}{h}\alpha_1\right) \\ 5\left(u_m + \alpha_1 + \alpha_2 + 12\frac{\lambda}{h}\alpha_2\right) \end{pmatrix}.$$

Where the two coefficients are now α_1, α_2.

This leads to the Jacobian

$$\frac{\partial F}{\partial V} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -\frac{\alpha_1^2}{3} - u_m^2 + gh - \frac{\alpha_2^2}{5} & 2u_m & \frac{2\alpha_1}{3} & \frac{2\alpha_2}{5} \\ -2\alpha_1 u_m - \frac{4}{5}\alpha_1\alpha_2 & 2\alpha_1 & 2u_m + \frac{4\alpha_2}{5} & \frac{4\alpha_1}{5} \\ -\frac{2}{3}\alpha_1^2 - 2\alpha_2 u_m - \frac{2}{7}\alpha_2^2 & 2\alpha_2 & \frac{4\alpha_1}{3} & 2u_m + \frac{4\alpha_2}{7} \end{pmatrix}$$

and the full system matrix reads

$$A = \frac{\partial F}{\partial V} - Q = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -\frac{\alpha_1^2}{3} - u_m^2 + gh - \frac{\alpha_2^2}{5} & 2u_m & \frac{2\alpha_1}{3} & \frac{2\alpha_2}{5} \\ -2\alpha_1 u_m - \frac{4}{5}\alpha_1\alpha_2 & 2\alpha_1 & u_m + \alpha_2 & \frac{3\alpha_1}{5} \\ -\frac{2}{3}\alpha_1^2 - 2u_m\alpha_2 - \frac{2}{7}\alpha_2^2 & 2\alpha_2 & -\frac{\alpha_1}{3} & u_m + \frac{3\alpha_2}{7} \end{pmatrix}. \qquad (4.40)$$

However, the system is not hyperbolic and the non-hyperbolic regions are clearly shown in
Figure 3. In particular, the eigenvalues depend on α_1 and α_2. It was shown in [21] that the
non-hyperbolic regions can be reached in standard simulations which makes the SWME models
with $N > 1$ prone to stability problems.

There are several hyperbolic regularization of the SWME with arbitrary N that restore hyper-
bolicity and yield more stable solutions while achieving similar accuracy as the original model. For
more details, see [21]. However, it is very difficult to investigate the steady states for these models
as the number of non-conservative terms is large. At the same time, those models do not depend
on the higher order coefficients α_i any more, which leads to a drastic simplification.
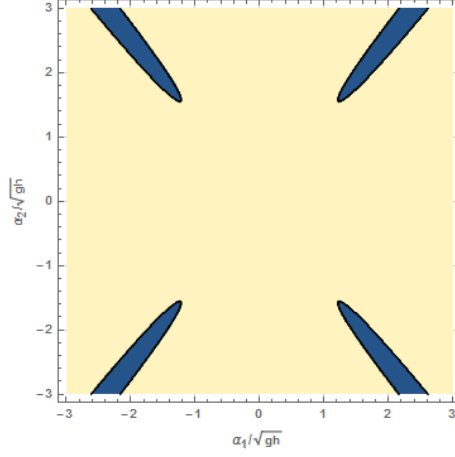
Figure 3: Non-hyperbolic region of $N = 2$ model in blue, from [21].

5   Shallow Water Linearized Moment Equations

In the previous section we have seen that the general SWME lacks hyperbolicity and a proper analysis of steady states is difficult due to the non-conservative terms. Note that even the hyperbolic HSWME model and the related β-HSWME model in [21] pose the same problems for computing the steady states.

In this paper, we propose a new hyperbolic model for the simulation of shallow flows, which is called Shallow Water Linearized Moment Equations (SWLME). Its derivation is based on the insights from the SWME $N = 1$ model. We saw that the steady states are easy to obtain as long as there are not that many non-conservative terms in the model and as long as the higher-order equations for the variables $h\alpha_i$ are not too complicated.

The difficult expressions in the higher-order equations are obtained by the non-linear terms $\partial_x \left(hu^2 \right)$ and $\partial_\zeta \left(hu\omega \right)$ in the vertically-resolved system (2.6)-(2.7), which require the computation of the following terms after insertion of the ansatz (4.1)

$$\int_0^1 \phi_i u^2 \, d\zeta \quad \text{and} \quad \int_0^1 \phi_i \partial_\zeta \left(u\omega \right) \, d\zeta.$$

Following an exact derivation, the first term evaluates to

$$\int_0^1 \phi_i u^2 \, d\zeta \;\; = \;\; \int_0^1 \phi_i \left(u_m + \sum_{j=1}^N \alpha_j \phi_j \right)^2 d\zeta \tag{5.1}$$

$$= \;\; u_m^2 \int_0^1 \phi_i \, d\zeta + \sum_{j=1}^N 2 u_m \alpha_j \int_0^1 \phi_i \phi_j \, d\zeta + \sum_{j,k=1}^N 2 \alpha_j \alpha_k \int_0^1 \phi_i \phi_j \phi_k \, d\zeta \tag{5.2}$$

$$= \;\; 0 + \frac{2}{2i+1} u_m \alpha_i + \frac{1}{2i+1} \sum_{j,k}^N A_{ijk} \alpha_j \alpha_k. \tag{5.3}$$

Assuming small deviations from a constant profile, i.e., $\alpha_i = \mathcal{O}\left(\epsilon \right)$ allows for neglecting the last term containing the coefficient coupling $\alpha_j \alpha_k = \mathcal{O}\left(\epsilon^2 \right)$. This results in

$$\int_0^1 \phi_i u^2 \, d\zeta \approx \frac{2}{2i+1} u_m \alpha_i.$$

The second term exactly evaluates to

$$\int_0^1 \phi_i \partial_\zeta \left(u\omega \right) \, d\zeta \;\; = \;\; -\frac{1}{2i+1} u_m \partial_x (h\alpha_i) + \sum_{j,k}^N B_{ijk} \alpha_j \partial_x (h\alpha_k). \tag{5.4}$$

10

Again assuming small coefficients $\alpha_i = \mathcal{O}(\epsilon)$ that only change moderately, the last term containing the coefficient coupling $\alpha_j \partial_x(h\alpha_k)$ is neglected. This results in

$$\int_0^1 \phi_i \partial_\zeta (u\omega) \approx -\frac{1}{2i+1} u_m \partial_x(h\alpha_i).$$

164   This leads to two changes in the equation system:

165   1. The left-hand side transport term does no longer include the non-linear couplings between
166   different α_i.

167   2. The right-hand side non-conservative term does no longer contain coupling terms between
168   different α_i.

169   Due to the linearization, the new model is called Shallow Water Linearized Moment Equations
170   (SWLME).

171   *Remark* 1. The linearization procedure outlined for the SWLME is related to the hyperbolic
172   regularization procedure that leads to the so-called Hyperbolic Moment Equations (HME) for
173   rarefied gases in [6], which are linearized around the equilibrium point in conservative variables.
174   Another similar linearization was performed in the derivation of the so-called Simplified Hyperbolic
175   Moment Equations (SHME) for rarefied gases in [22], which neglects the non-linearity in the ansatz
176   to derive a hyperbolic but much simpler moment model.

To see the effect of the changes in practice, consider the simple case $N = 2$ that will later be extended for larger N. The model reads

$$\partial_t \begin{pmatrix} h \\ hu_m \\ h\alpha_1 \\ h\alpha_2 \end{pmatrix} + \partial_x \begin{pmatrix} hu \\ hu_m^2 + g\frac{h^2}{2} + \frac{1}{3}h\alpha_1^2 + \frac{1}{5}h\alpha_2^2 \\ \textcolor{red}{2hu_m\alpha_1} \\ \textcolor{red}{2hu_m\alpha_2} \end{pmatrix} = Q\partial_x \begin{pmatrix} h \\ hu \\ h\alpha_1 \\ h\alpha_2 \end{pmatrix} - \frac{\nu}{\lambda}P$$

with

$$\textcolor{red}{Q} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \textcolor{red}{u_m} & \textcolor{red}{0} \\ 0 & 0 & \textcolor{red}{0} & \textcolor{red}{u_m} \end{pmatrix} \text{ and } P = \begin{pmatrix} 0 \\ u_m + \alpha_1 + \alpha_2 \\ 3(u_m + \alpha_1 + \alpha_2 + 4\frac{\lambda}{h}\alpha_1) \\ 5(u_m + \alpha_1 + \alpha_2 + 12\frac{\lambda}{h}\alpha_2) \end{pmatrix}.$$

The changed entries are given in red, illustrating the derivation above. While the model looks simpler than the SWME model (4.39), in comparison with the HSWME from [21], the differences are smaller as the HSWME model also neglects the high-order linear terms. Most importantly, the momentum equation, which is the second equation of the model, is exactly recovered by the SWLME and the system matrix A still depends on the second coefficient α_2, which is both not the case for the HSWME model. The system matrix is given by

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -\frac{\alpha_1^2}{3} - u_m^2 + gh - \frac{\alpha_2^2}{5} & 2u_m & \frac{2\alpha_1}{3} & \frac{2\alpha_2}{5} \\ \textcolor{red}{-2u_m\alpha_1} & 2\alpha_1 & \textcolor{red}{u_m} & \textcolor{red}{0} \\ \textcolor{red}{-2u_m\alpha_2} & 2\alpha_2 & \textcolor{red}{0} & \textcolor{red}{u_m} \end{pmatrix}. \tag{5.5}$$

177   Albeit being a simpler model, the model captures most of the original model, including the conser-
178   vation of mass and momentum and the dependence of the momentum terms hu on the higher order
179   equations. The second column of the system matrix is not changed at all, leading to the correct
180   momentum influence on the higher order equations. Only the coupling between the higher-order
181   equations, induced by the non-linear parts (e.g. $h\alpha_1\alpha_2$ and $h\alpha_2^2$ and the additional non-conservative
182   terms) is reduced. However, there is still a non-linear velocity and momentum coupling between
183   all higher-order equations.

This procedure can be generalized to an explicit system for arbitrary N following the same strategy. The model equations read:

$$\partial_t \begin{pmatrix} h \\ hu_m \\ h\alpha_1 \\ \vdots \\ h\alpha_N \end{pmatrix} + \partial_x \begin{pmatrix} hu_m \\ hu_m^2 + g\frac{h^2}{2} + \frac{1}{3}h\alpha_1^2 + \ldots + \frac{1}{2N+1}h\alpha_N^2 \\ 2hu_m\alpha_1 \\ \vdots \\ 2hu_m\alpha_N \end{pmatrix} = Q\partial_x \begin{pmatrix} h \\ hu_m \\ h\alpha_1 \\ \vdots \\ h\alpha_N \end{pmatrix} + P. \qquad (5.6)$$

The non-conservative term is simplified to

$$Q = \operatorname{diag}\left(0, 0, u_m, \ldots, u_m\right).$$

The system matrix of the new SWLME then reads

$$A_N = \begin{pmatrix} 0 & 1 & 0 & \vdots & 0 \\ gh - u_m^2 - \frac{\alpha_1^2}{3} - \ldots - \frac{\alpha_N^2}{2N+1} & 2u_m & \frac{2\alpha_1}{3} & \cdots & \frac{2\alpha_N}{2N+1} \\ -2u_m\alpha_1 & 2\alpha_1 & u_m & & \\ \vdots & \vdots & & \ddots & \\ -2u_m\alpha_N & 2\alpha_N & & & u_m \end{pmatrix} \in \mathbb{R}^{(N+2)\times(N+2)}. \qquad (5.7)$$

For the model with general $N > 2$, the same observations as for the $N = 2$ model hold, including the conservation of mass and momentum as well as the exact second column of the system matrix. The coupling between the higher-order equations is reduced, but still present.

An analysis of the system matrix reveals the following theorem.

Theorem 1. *The SWLME system matrix $A_N \in \mathbb{R}^{(N+2)\times(N+2)}$ (5.7) has the following characteristic polynomial*

$$\chi_{A_N}(\lambda) = (u_m - \lambda)\left[(\lambda - u_m)^2 - gh - \sum_{i=1}^{N} \frac{3\alpha_i^2}{2i+1}\right]$$

and the eigenvalues are given by

$$\lambda_{1,2} = u_m \pm \sqrt{gh + \sum_{i=1}^{N} \frac{3\alpha_i^2}{2i+1}} \quad and \quad \lambda_{i+2} = u, \ \ for \ \ i = 1, \ldots, N. \qquad (5.8)$$

The system is thus hyperbolic.

Proof. The proof closely follows the proof of the characteristic polynomial of the HSWME system matrix in [21]. However, we can compute the characteristic polynomial and all eigenvalues explicitly here.

We write $\widetilde{\lambda} = \lambda - u_m$, so that we can compute the characteristic polynomial using

$$\begin{aligned} \chi_{A_N}(\lambda) &= \det\left(A_N - \lambda I\right) \\ &= \det\left(A_N - \left(\widetilde{\lambda} + u_m\right)I\right). \end{aligned}$$

When writing A_N, the following notation is used for conciseness:

$$\begin{aligned} d_0 &= gh + \sum_{i=1}^{N} \frac{\alpha_i^2}{2i+1}, \\ d_i &= -2u\alpha_i, \ \text{for} \ i = 1, \ldots, N \\ c_i &= 2\alpha_i, \ \text{for} \ i = 1, \ldots, N \\ b_i &= \frac{2\alpha_i}{2i+1}, \ \text{for} \ i = 1, \ldots, N. \end{aligned}$$

12

Computing the determinant $\left|A_N - \left(\widetilde{\lambda} + u_m\right) I\right|$ by developing with respect to the first row

yields

$$\left|A_N - \left(\widetilde{\lambda} + u_m\right) I\right| = \begin{vmatrix} -\widetilde{\lambda} - u_m & 1 & & & \\ d_0 & u_m - \widetilde{\lambda} & b_1 & \ldots & b_N \\ d_1 & c_1 & -\widetilde{\lambda} & & \\ \vdots & \vdots & & \ddots & \\ d_N & c_N & & & -\widetilde{\lambda} \end{vmatrix}$$

$$= \left(-\widetilde{\lambda} - u_m\right) \cdot \underbrace{\begin{vmatrix} u_m - \widetilde{\lambda} & b_1 & \ldots & b_N \\ c_1 & -\widetilde{\lambda} & & \\ \vdots & & \ddots & \\ c_N & & & -\widetilde{\lambda} \end{vmatrix}}_{=C_{N+1} \in \mathbb{R}^{(N+1)\times(N+1)}} -1 \cdot \underbrace{\begin{vmatrix} d_0 & b_1 & \ldots & b_N \\ d_1 & -\widetilde{\lambda} & & \\ \vdots & & \ddots & \\ d_N & & & -\widetilde{\lambda} \end{vmatrix}}_{=D_{N+1} \in \mathbb{R}^{(N+1)\times(N+1)}}$$

The determinants of C_{N+1}, D_{N+1} are computed by developing with respect to the last row as

$$|C_{N+1}| = \begin{vmatrix} u_m - \widetilde{\lambda} & b_1 & \ldots & b_N \\ c_1 & -\widetilde{\lambda} & & \\ \vdots & & \ddots & \\ c_N & & & -\widetilde{\lambda} \end{vmatrix} = (-1)^{N+2} c_N \underbrace{\begin{vmatrix} b_1 & \ldots & b_{N-1} & b_N \\ -\widetilde{\lambda} & & & \\ & \ddots & & \\ & & & -\widetilde{\lambda} \end{vmatrix}}_{=B_N \in \mathbb{R}^{N\times N}} + (-1)^{2N+2} \left(-\widetilde{\lambda}\right) |C_N|$$

and

$$|D_{N+1}| = \begin{vmatrix} d_0 & b_1 & \ldots & b_N \\ d_1 & -\widetilde{\lambda} & & \\ \vdots & & \ddots & \\ d_N & & & -\widetilde{\lambda} \end{vmatrix} = (-1)^{N+2} d_N \underbrace{\begin{vmatrix} b_1 & \ldots & b_{N-1} & b_N \\ -\widetilde{\lambda} & & & \\ & \ddots & & \\ & & & -\widetilde{\lambda} \end{vmatrix}}_{=B_N \in \mathbb{R}^{N\times N}} + (-1)^{2N+2} \left(-\widetilde{\lambda}\right) |D_N|.$$

The determinant of B_N is easily computed as

$$|B_N| = (-1)^{N+1} b_N \left(-\widetilde{\lambda}\right)^{N-1}.$$

With the help of this, we get

$$|C_{N+1}| = -c_N b_N \left(-\widetilde{\lambda}\right)^{N-1} + \left(-\widetilde{\lambda}\right) |C_N| = \ldots = \left(-\widetilde{\lambda}\right)^{N-1} \left(-\sum_{i=1}^{N} c_i b_i\right) + \left(-\widetilde{\lambda}\right)^N \underbrace{\left(u_m - \widetilde{\lambda}\right)}_{=|C_1|}$$

and analogously

$$|D_{N+1}| = -d_N b_N \left(-\widetilde{\lambda}\right)^{N-1} + \left(-\widetilde{\lambda}\right) |D_N| = \ldots = \left(-\widetilde{\lambda}\right)^{N-1} \left(-\sum_{i=1}^{N} d_i b_i\right) + \left(-\widetilde{\lambda}\right)^N \underbrace{d_0}_{=|D_1|}.$$

Note that $-\sum_{i=1}^{N} c_i b_i = -\sum_{i=1}^{N} \frac{4\alpha_i^2}{2i+1}$, $-\sum_{i=1}^{N} d_i b_i = \sum_{i=1}^{N} \frac{4\alpha_i^2}{2i+1} u_m$, $d_0 = gh + \sum_{i=1}^{N} \frac{\alpha_i^2}{2i+1}$.

196 Next, insertion of these terms into the characteristic polynomial of the system matrix A_N yields

$$
\begin{aligned}
\left| A_N - \left(\widetilde{\lambda} + u_m \right) I \right| &= \left(-\widetilde{\lambda} - u_m \right) \cdot |C_{N+1}| - 1 \cdot |D_{N+1}| \\
&= \left(-\widetilde{\lambda} - u_m \right) \cdot \left[\left(-\widetilde{\lambda} \right)^{N-1} \left(-\sum_{i=1}^{N} c_i b_i \right) + \left(-\widetilde{\lambda} \right)^{N} \left(u_m - \widetilde{\lambda} \right) \right] \\
&\quad - 1 \cdot \left[\left(-\widetilde{\lambda} \right)^{N-1} \left(-\sum_{i=1}^{N} d_i b_i \right) + \left(-\widetilde{\lambda} \right)^{N} d_0 \right] \\
&= \left(-\widetilde{\lambda} \right)^{N} \left[\widetilde{\lambda}^2 - gh - \sum_{i=1}^{N} \frac{3\alpha_i^2}{2i+1} \right] \\
&= \left(u_m - \lambda \right)^{N} \left[(\lambda - u_m)^2 - gh - \sum_{i=1}^{N} \frac{3\alpha_i^2}{2i+1} \right],
\end{aligned}
$$

197 which proves the first part of the theorem.

Setting the characteristic polynomial to zero results in the following propagation speeds of the system:

$$
\lambda_{1,2} = u_m \pm \sqrt{ gh + \sum_{i=1}^{N} \frac{3\alpha_i^2}{2i+1} }, \quad \text{and} \quad \lambda_{i+2} = u_m, \ \text{for} \ i = 1, \ldots, N.
$$

198 The propagation speeds prove that the system is hyperbolic for positive water height. $\qquad\square$

199 From the form of the eigenvalues, the new model for $N \geq 2$ can be seen as a consistent extension
200 of the hyperbolic $N = 1$ model from Section 4, compare also the eigenvalues in equation (4.25).

We remark that such an analysis is not possible for the original SWME model for arbitrary N as the eigenvalues have a very complicated structure and cannot be given in analytical form. For the new hyperbolic model, the eigenvalues $\lambda_{1,2}$ still depend on all flow variables. However, the analysis can be carried out analytically. For the hyperbolic HSWME and βHSWME models in [21], the eigenvalues depend solely on α_1, which is a drastic simplification. For those models, the eigenvector analysis is still very involved and theoretical results are only possible for small values of $(M\alpha)_1 \ll 1$. In this case, the model has the same wave properties as the SWLME system. From a straightforward computation, the eigenvectors v_i for $i = 1, \ldots, N+2$ of the SWLME system can be derived as

$$
v_{1,2} = \begin{pmatrix} \frac{1}{2\alpha_n} \\ \frac{1}{2\alpha_n} \left(u_m + \sqrt{ gh \pm \sum_{i=1}^{N} \frac{3\alpha_i^2}{2i+1} } \right) \\ \frac{\alpha_1}{\alpha_N} \\ \vdots \\ \frac{\alpha_N}{\alpha_N} \end{pmatrix}
\tag{5.9}
$$

$$
v_{i+2} = \begin{pmatrix} \dfrac{6\alpha_{n+1-1}}{(2(n+1-i)+1) - 3gh + \sum_{i=1}^{N} \frac{3\alpha_i^2}{2i+1}} \\ \dfrac{6\alpha_{n+1-1} u}{(2(n+1-i)+1) - 3gh + \sum_{i=1}^{N} \frac{3\alpha_i^2}{2i+1}} \\ \delta_{n+3-i,3} \\ \vdots \\ \delta_{n+3-i,N} \end{pmatrix}, \ \text{for} \ i = 1, \ldots, N,
\tag{5.10}
$$

201 for Kronecker delta $\delta_{i,j}$.
202 It can be checked that the first two eigenvalues $\lambda_{1,2}$ are genuinely non-linear, while all other
203 eigenvalues λ_{i+2} for $i = 1, \ldots, N$ are linearly degenerate. Note that the analysis of eigenvalues and

14

eigenvectors is not possible for the SWME system, due to the lack of hyperbolicity. The linearization within the N higher moment equations during the derivation procedure consistently leads to the resulting N linearly degenerate eigenvalues. However, the first two eigenvalues, corresponding to the unchanged conservation of mass and momentum, remain genuinely non-linear. The full characterization of the eigenstructure of the SWLME allows for the use of efficient numerical methods, for example using the relation between Riemann solvers and PVM methods [9].

Rankine-Hugoniot conditions can be derived analogously to the SWME case with $N = 1$ as follows. For flat bottom $\partial_x b = 0$ and zero friction, the steady state fulfills

$$\partial_x (hu_m) = 0 \tag{5.11}$$

$$\partial_x \left(hu_m^2 + \frac{1}{2}gh^2 + \frac{1}{3}h\alpha_1^2 + \ldots + \frac{1}{2N+1}h\alpha_N^2 \right) = 0 \tag{5.12}$$

$$\partial_x (2hu_m\alpha_1) = u_m\partial_x (h\alpha_1) \tag{5.13}$$

$$\vdots \tag{5.14}$$

$$\partial_x (2hu_m\alpha_N) = u_m\partial_x (h\alpha_N) \tag{5.15}$$

First looking at all equations except the second, we obtain after some modification

$$hu_m = const, \tag{5.16}$$

$$u_m = 0 \text{ or } \frac{\alpha_i}{h} = const, \text{ for } i = 1, \ldots, N. \tag{5.17}$$

Using those relations in the remaining second equation, we can derive the Rankine-Hugoniot conditions from a given state $(h_0, h_0 u_{m,0}, h_0\alpha_{1,0}, \ldots, h_0\alpha_{N,0})$ to a state $(h, hu_m, h\alpha_1, \ldots, h\alpha_N)$ and obtain (after some modifications)

$$(h - h_0) \left[-\frac{u_{m,0}^2}{gh_0} + \frac{1}{2} \left(\left(\frac{h}{h_0} \right)^2 + \left(\frac{h}{h_0} \right) \right) + \sum_{i=1}^{N} \frac{1}{2i+1} \frac{\alpha_{i,0}^2}{gh_0} \left(\left(\frac{h}{h_0} \right)^3 + \left(\frac{h}{h_0} \right)^2 + \left(\frac{h}{h_0} \right) \right) \right] = 0. \tag{5.18}$$

We extend the previous dimensionless flow numbers by using one number for each variable:

$$Fr = \frac{u_{m,0}}{\sqrt{gh_0}}, \tag{5.19}$$

$$(M\alpha)_i = \frac{\alpha_{i,0}}{u_{m,0}}, \quad \text{for } i = 1, \ldots, N, \tag{5.20}$$

writing $y = \frac{h}{h_0}$, we arrive at the two solutions

$$h = h_0 \ \lor \ -Fr^2 + \frac{1}{2} \left(y^2 + y \right) + \sum_{i=1}^{N} \frac{1}{2i+1} (M\alpha)_i^2 Fr^2 \left(y^3 + y^2 + y \right) = 0. \tag{5.21}$$

From the previous equation, we see a new dimensionless number $M\alpha^2 := \sum_{i=1}^{N} \frac{1}{2i+1} (M\alpha)_i^2$ appearing. The new number $M\alpha$ measures the total deviation from equilibrium. This leads to a consistent extension of the SWME $N = 1$ test case above. We see that the Rankine-Hugoniot conditions allow for similar solutions as before, this time with Fr and $M\alpha$ as dimensionless flow numbers. We note that the equations always have at least one solution for non-zero Fr and $M\alpha$.

Analogously, we extend the conditions for smooth and frictionless steady states including a bottom topography. We will later use this to derive a well-balancing scheme. We can derive

$$\partial_x \left(\frac{1}{2}u_m^2 + g(h + b) + \frac{3}{2} \sum_{i=1}^{N} \frac{1}{2i+1}\alpha_i^2 \right) = 0. \tag{5.22}$$

The non-trivial steady state solution can thus be found using

$$hu_m \;=\; const, \tag{5.23}$$

$$\frac{1}{2}u_m^2 + g(h+b) + \frac{3}{2}\sum_{i=1}^{N}\frac{1}{2i+1}\alpha_i^2 \;=\; const, \tag{5.24}$$

$$\frac{\alpha_i}{h} \;=\; const, \text{ for } i=1,\ldots,N. \tag{5.25}$$

This expression can be used in the following numerical methods section to obtain a proper well-balancing scheme for the new model. First, we will rewrite the model in the proper form with a conservative and non-conservative part to use it in the numerical schemes thereafter.

The system (5.6) with topography but without friction terms is therefore written in the form

$$U_t + \partial_x F(U) + B(U)\partial_x U = S(U)\partial_x b. \tag{5.26}$$

By straightforward calculation, we obtain

$$U = \begin{pmatrix} h \\ hu_m \\ h\alpha_1 \\ \vdots \\ h\alpha_N \end{pmatrix}, \quad F(U) = \begin{pmatrix} hu_m \\ hu_m^2 + g\frac{h^2}{2} + \frac{1}{3}h\alpha_1^2 + \ldots + \frac{1}{2N+1}h\alpha_N^2 \\ 2hu_m\alpha_1 \\ \vdots \\ 2hu_m\alpha_N \end{pmatrix}, \tag{5.27}$$

$$B(U) = diag(0,0,-u_m,...,-u_m). \tag{5.28}$$

We can also write this system in the form

$$\partial_t W + \mathcal{A}(W)\partial_x W = 0, \tag{5.29}$$

with

$$W = \begin{pmatrix} h \\ hu_m \\ h\alpha_1 \\ \vdots \\ h\alpha_N \\ b \end{pmatrix}, \quad \mathcal{A}(W) = \begin{pmatrix} A(W) & -S(W) \\ 0 & 0 \end{pmatrix},$$

where $A(W)$ has the form (5.7) and $S(W) = \begin{pmatrix} 0 \\ -gh \\ 0 \\ \vdots \\ 0 \end{pmatrix}$.

For comparison we note that also the existing HSWME and β-HSWME models from [21] can be written in the same form, see the appendix A.

6 Numerical methods

In this section, we recall the general high-order well-balanced method from [13] and construct the first order as well as the second order scheme for applications of the SWLME derived in the previous section. At the end of the section we will outline the specific spatial discretization scheme used for the numerical tests in the next section.

6.1 A general high-order well-balanced procedure

The previously derived shallow water models can all be written as non-conservative systems of the form

$$\partial_t U + \partial_x F(U) + B(U)\partial_x U = S(U)\partial_x b. \tag{6.1}$$

It is well known that these systems are equivalent to

$$\partial_t W + \mathcal{A}(W)\partial_x W = 0, \tag{6.2}$$

where

$$W = \begin{pmatrix} U \\ b \end{pmatrix}, \quad \mathcal{A}(W) = \begin{pmatrix} \frac{\partial F}{\partial U}(U) + B(U) & -S(U) \\ 0 & 0 \end{pmatrix}.$$

The goal of this section is to develop a family of numerical methods that are well-balanced for the frictionless SWLME introduced before, i.e., that preserve the stationary solutions verifying (5.23), (5.24) and (5.25). In this section we will follow [13] adding the non-conservative products. The interested reader is referred to this reference for details and proofs.

We consider semi-discrete finite-volume methods of the form

$$\frac{dW_i}{dt} = -\frac{1}{\Delta x}\left(D^-_{i+\frac{1}{2}} + D^+_{i-\frac{1}{2}} + \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \mathcal{A}(\mathbb{P}_i(x))\frac{\partial}{\partial x}\mathbb{P}_i(x)dx \right), \tag{6.3}$$

where

- $W_i(t) \cong \int_{x_{i+\frac{1}{2}}}^{x_{i-\frac{1}{2}}} W(t,x)\, dx$ is the respective cell average value,

- $\mathbb{P}_i(x)$ is a high-order well-balanced operator in the sense defined in [13].

- $D^{\pm}_{i+\frac{1}{2}} = \mathbb{D}^{\pm}\left(W^-_{i+\frac{1}{2}}, W^+_{i+\frac{1}{2}} \right)$, is the respective fluctuation with reconstructed states

$$W^-_{i+\frac{1}{2}} = \mathbb{P}_i(x_{i+\frac{1}{2}}), \quad W^+_{i+\frac{1}{2}} = \mathbb{P}_{i+1}(x_{i+\frac{1}{2}}),$$

and $\mathbb{D}(W_l, W_r)$ verifies:

$$\mathbb{D}^-(W_l, W_r) + \mathbb{D}^+(W_l, W_r) = \int_0^1 \mathcal{A}(\Psi)\frac{\partial \Psi}{\partial s}\, ds, \tag{6.4}$$

where Ψ is a family of paths joining W_l with W_r.

In order to design the high-order well-balanced operator we follow the strategy introduced in [10]. The following steps need to be performed in order to compute $\mathbb{P}_i$ at the cell $[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$ for a given family of cell values $\{W_i\}$:

1. Obtaining the steady solution $W_i^*(x)$ such that:

$$\frac{1}{\Delta x}\int_{x_{i-\frac{1}{2}}}^{x_{i-\frac{1}{2}}} W_i^*(x)dx = W_i, \tag{6.5}$$

if possible. In other cases consider $W_i^* \equiv W_i^n$.

2. Computing the fluctuations $\{V_j\}_{j \in S_i}$ within the stencil S_i:

$$V_j = W_j - \frac{1}{\Delta r}\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} W_i^*(x)dx, \quad j \in S_i. \tag{6.6}$$

3. Applying the reconstruction operator with the necessary order to the fluctuations $\{V_j\}_{j \in S_i}$:

$$Q_i(x) = Q_i(x; \{V_j\}_{j \in S_i}).$$

17

4. Defining the well-balanced operator:

$$\mathbb{P}_i(x) = W_i^*(x) + Q_i(x).$$

$\mathbb{P}_i$ is well-balanced for every steady solution provided that the reconstruction operator Q_i is exact for the null function. Moreover, it is conservative, i.e.,

$$\frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \mathbb{P}_i(x) dr = W_i, \quad \text{for all } i,$$

provided that Q_i is conservative, and it is high-order accurate provided that the steady solutions are smooth (see [13] for details).

6.2 First order well-balanced scheme

We apply the steps of the previous subsection to the system (5.29) in a first order setup before considering the second order scheme in the next section. As the bottom topography b is known, we will focus on the other variables of the system.

The cell averages of the initial condition will be computed using the mid-point rule, that is

$$W_i^0 = W_0(x_i), \quad \text{for all } i,$$

where $W_0(x)$ is the initial condition.

In the case of the SWLME system, the steady state solutions verify:

$$
\begin{aligned}
hu_m &= C_1 \equiv const, \\
\frac{1}{2} u_m^2 + g(h + b) + \frac{3}{2} \sum_{i=1}^{N} \frac{1}{2i+1} \alpha_i^2 &= C_2 \equiv const, \\
\frac{\alpha_1}{h} &= C_3 \equiv const, \\
\frac{\alpha_2}{h} &= C_4 \equiv const, \\
&\vdots \\
\frac{\alpha_N}{h} &= C_{N+2} \equiv const.
\end{aligned}
$$

Using the mid-point rule in (6.5) the first step is to obtain, if possible, the stationary solution W_i^* such that:

$$W_i^*(x_i) = W_i. \tag{6.7}$$

With this information the constants $C_1, C_2, C_3,...,C_{N+2}$ can be computed as

$$
\begin{cases}
C_1 = h_i u_{m,i}, \\
C_2 = \frac{1}{2} u_{m,i}^2 + g(h_i + b(x_i)) + \frac{3}{2} \sum_{j=1}^{N} \frac{1}{2j+1} \alpha_{j,i}^2, \\
C_3 = \frac{\alpha_{1,i}}{h_i}, \\
C_4 = \frac{\alpha_{2,i}}{h_i}, \\
\quad \vdots \\
C_{N+2} = \frac{\alpha_{N,i}}{h_i}.
\end{cases}
\tag{6.8}
$$

Using the relations (6.8), the stationary solution can be evaluated in a point $x = a$. The evaluation of the steady state solution requires finding roots of the function

$$f(h) = Dh^4 + 2h^3 g + 2h^2(gb(a) - C_2) + C_1^2, \tag{6.9}$$

where the parameter D is given by

$$D = C_3^2 + \frac{3}{5} C_4^2 + \cdots + \frac{3}{2N+1} C_{N+2}^2.$$

18

The derivative of the function f is given by

$$f'(h) = 4Dh^3 + 6h^2g + 4h(gb(a) - C_2).$$

The positive root h_c of $f'(h)$ is

$$h_c = \frac{-3g + \sqrt{9g^2 - 16D(b(a)g - C_2)}}{4D}, \tag{6.10}$$

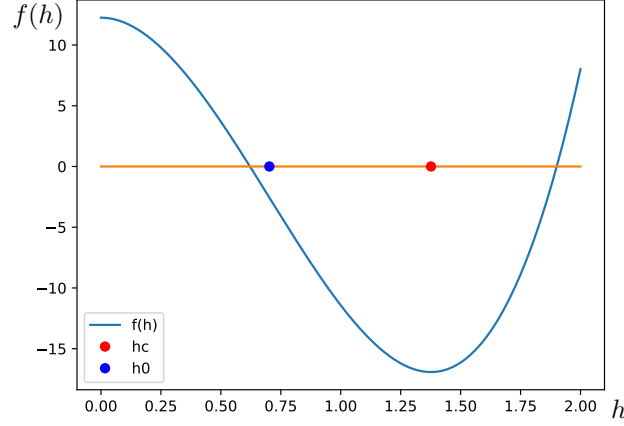and we can see that it is a minimum of the function f. An example of a function f is plotted in Figure 4.

Figure 4: An example of the root finding function $f(h)$ (6.9) with some constants C_i. The minimum h_c and the initial value of the Newton algorithm h_0 are shown.

Remark 1. *In case of $D = 0$, the minimum simplifies to*

$$h_c = -\frac{2(b(a)g - C_2)}{3g}.$$

Remark 2. *In order to find the roots of $f(h)$ the Newton-Raphson method is employed with initial value h_0 that is the positive root of*

$$f''(h) = 12Dh^2 + 12hg + 4(gb(a) - C_2),$$

given by

$$h_0 = \frac{-3g + \sqrt{9g^2 - 12D(b(a)g - C_2)}}{6D}.$$

It is easy to see that $0 \leq h_0 \leq h_c$, compare also Figure 4.

We can conclude then the following: If $f(h_c) < 0$ there exist two possible states for $W_i^*(x_{i\pm\frac{1}{2}})$, one subcritical and one supercritical. The following criterion will be used to choose one state:

1. If W_i is subcritical or supercritical, then we will choose the solution in the same regime (subcritical or supercritical) as W_i for $W_i^*(x_{i\pm\frac{1}{2}})$.

2. If W_i is transcritical, then the solution that has the same behaviour (subcritical or supercritical) as W_{i-1} will be selected for $W_i^*(x_{i-\frac{1}{2}})$ and the solution whose behaviour is the same as W_{i+1} will be selected for $W_i^*(x_{i+\frac{1}{2}})$.

19

Following the procedure described in [13], the reconstruction operator reduces to $\mathbb{P}_i(x) = W_i^*(x)$ and the first order numerical scheme reduces to:

$$W_i^{n+1} = W_i^n - \frac{\Delta t}{\Delta r}(D_{i+\frac{1}{2}}^- + D_{i-\frac{1}{2}}^+),\qquad(6.11)$$

for $W_{i-\frac{1}{2}}^+ = \mathbb{P}_i(x_{i-\frac{1}{2}})$ and $W_{i+\frac{1}{2}}^- = \mathbb{P}_i(x_{i+\frac{1}{2}})$, where we have used that $\mathbb{P}(x) = W_i^*(x)$ is a steady solution.

In the case we could not find such a stationary solution verifying (6.7) the standard trivial reconstruction is considered.

6.3   Second order well-balanced scheme

Now we consider the second order scheme for which a second order spatial reconstruction using the minmod limiter will be employed, see [13].

The cell averages of the initial condition are again computed using the mid-point rule:

$$W_i^0 = W_0(x_i), \quad \text{for all } i,$$

where $W_0(x)$ is the initial condition.

1. Obtaining the steady solution: In the same fashion as for the first order scheme, if possible, the steady state W_i^* needs to be found such that

$$W_i^*(x_i) = W_i.\qquad(6.12)$$

   After computing the constants $C_1, C_2, C_3,...,C_{N+2}$ as in (6.8), the stationary solution can be evaluated in a point $x = a$. In order to do this, the roots of the function f in (6.9) needs to be computed. As defined in (6.10), f has a minimum in h_c. Again if $f(h_c) < 0$ there exist two possible values for $W_i^*(i \pm 1, i \pm \frac{1}{2})$ and we use the same criterion as for the first order scheme in order to choose one.

2. Computing the fluctuations: After the evaluation of the stationary solution in a point $r = a$ the fluctuations $\{V_{i-1}, V_i, V_{i+1}\}$ in (6.6) are computed using the mid-point rule

$$\begin{aligned} V_{i-1} &= W_{i-1} - W_i^*(x_{i-1}), \\ V_i &= W_i - W_i^*(x_i) = 0, \\ V_{i+1} &= W_{i+1} - W_i^*(x_{i+1}). \end{aligned}$$

3. Applying the reconstruction operator: After the fluctuations are computed the *minmod* reconstruction is used to obtain the reconstruction operator (see [34])

$$Q_i(x) = V_i + minmod\left(\frac{V_i - V_{i-1}}{\Delta x}, \frac{V_{i+1} - V_{i-1}}{2\Delta x}, \frac{V_{i+1} - V_i}{\Delta x}\right)(x - x_i),$$

   where

$$minmod(a, b, c) = \begin{cases} min\{a, b, c\} & \text{if } a, b, c > 0, \\ max\{a, b, c\} & \text{if } a, b, c < 0, \\ 0 & \text{otherwise.} \end{cases}$$

4. Defining the well-balanced operator: The well-balanced reconstruction operator is given by

$$\mathbb{P}_i(x) = W_i^*(x) + Q_i(x).$$

The well-balanced property can be lost if a quadrature formula is used directly in the right part of (6.3), as the quadrature formula is in general not exact. Therefore, the semi-discrete scheme is

20

first rewritten as proposed in [13] taking into account the non-conservative part

$$\frac{dW_i}{dt} = -\frac{1}{\Delta x}\left(D_{i+\frac{1}{2}}^- + D_{i-\frac{1}{2}}^+ + \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}}\left(\mathcal{A}(\mathbb{P}_i(x))\frac{\partial}{\partial x}\mathbb{P}_i(x) - \mathcal{A}(W_i^*(x))\frac{\partial}{\partial x}W_i^*(x)\right)dx\right.$$

$$\left. + \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}}\mathcal{A}(W_i^*(x))\frac{\partial}{\partial x}W_i^*(x)dx\right).$$

Once this equivalent form is obtained, we use that W_i^* is a stationary solution in the second integral and then employ the mid-point rule for the first integral without losing the well-balanced property what leads to

$$\frac{dW_i}{dt} = -\frac{1}{\Delta x}\left(D_{i+\frac{1}{2}}^- + D_{i-\frac{1}{2}}^+ + \mathcal{A}(\mathbb{P}_i(x_i))minmod\left(\frac{V_i - V_{i-1}}{\Delta x}, \frac{V_{i+1} - V_{i-1}}{2\Delta x}, \frac{V_{i+1} - V_i}{\Delta x}\right)\right), \quad (6.13)$$

for $W_{i-\frac{1}{2}}^+ = \mathbb{P}_i(x_{i-\frac{1}{2}})$ and $W_{i+\frac{1}{2}}^- = \mathbb{P}_i(x_{i+\frac{1}{2}})$. The discretization in time is performed with a Runge-Kutta TVD method of order 2, see [19].

Remark 3. *The extension to higher-order is straightforward: Although not implemented in the present paper, a third order well-balanced scheme will be based on the two point Gaussian quadrature formula for computing the averages. In the first step, we need to find the constants C_j, $j = 1, ..., N+2$ such that*

$$\frac{1}{2}W_i^*(x_a, C_1, ..., C_{N+2}) + \frac{1}{2}W_i^*(x_b, C_1, ..., C_{N+2}) = W_i,$$

where x_a and x_b are the two quadrature points and $W_i^*(x, C_1, ..., C_{N+2})$ represents the stationary solution given by the constants C_j evaluated in x. Then we follow the steps considering a third order reconstruction operator (e.g. CWENO reconstruction) and using again the two point Gaussian quadrature.

6.4 Spatial discretization

In order to completely define the scheme, what remains is to define the form of the fluctuations $D_{i+\frac{1}{2}}^\pm$ and the non-conservative terms in (6.3), for which we use a path-consistent scheme based on segments in the conservative variables as family of paths joining two states:

$$\Psi(s; W_l, W_r) = \begin{pmatrix} \Psi_U(s; W_l, W_r) \\ \Psi_b(s; W_l, W_r) \end{pmatrix} = \begin{pmatrix} U_l + s(U_r - U_l) \\ b_l + s(b_r - b_l) \end{pmatrix}, \quad s \in [0,1],$$

and a PVM-like method [9] corresponding to a choice in (6.4) of

$$D_{i+\frac{1}{2}}^\pm = \frac{1}{2}\left(F(U_r) - F(U_l) + B_{i+\frac{1}{2}}(U_r - U_l) - S_{i+\frac{1}{2}}(b_r - b_l)\right. \tag{6.14}$$

$$\left. \pm Q_{i+\frac{1}{2}}(U_r - U_l - \mathcal{A}_{i+\frac{1}{2}}^{-1}S_{i+\frac{1}{2}}(b_r - b_l))\right),$$

where

$$\mathcal{A}_{i+\frac{1}{2}} = \begin{pmatrix} A_{i+\frac{1}{2}} & -S_{i+\frac{1}{2}} \\ 0 & 0 \end{pmatrix}$$

is a generalized Roe matrix [33] so that $A_{i+\frac{1}{2}} = J_{i+\frac{1}{2}} + B_{i+\frac{1}{2}}$ and $S_{i+\frac{1}{2}}$ have to verify

$$A_{i+\frac{1}{2}} = \int_0^1 A(U_l + s(U_r - U_l))\,ds, \tag{6.15}$$

$$J_{i+\frac{1}{2}}(U_r - U_l) = F(U_r) - F(U_l), \tag{6.16}$$

21

$$B_{i+\frac{1}{2}} = \int_0^1 B(U_l + s(U_r - U_l)) \, ds, \tag{6.17}$$

$$S_{i+\frac{1}{2}} = \int_0^1 S(U_l + s(U_r - U_l)) \, ds, \tag{6.18}$$

and the polynomial viscosity matrix is $Q_{i+\frac{1}{2}} = P(A_{i+\frac{1}{2}})$, for polynomial P. The source term evaluates to

$$S_{i+\frac{1}{2}} = \begin{pmatrix} 0 \\ -g\frac{h_l+h_r}{2} \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

In the case of the model SWLME, it can be shown that the system (6.16) leads to an evaluation of the Jacobian

$$J_{i+\frac{1}{2}} = \frac{\partial F}{\partial U}(h_R, u_{m,R}, \alpha_{1,R}, ..., \alpha_{N,R}), \tag{6.19}$$

at the intermediate values

$$h_R = \frac{h_l + h_r}{2}, \quad u_{m,R} = \frac{\sqrt{h_l}u_{m,l} + \sqrt{h_r}u_{m,r}}{\sqrt{h_l} + \sqrt{h_r}},$$

and

$$\alpha_{j,R} = \frac{\sqrt{h_l}h_r\alpha_{j,r} + \sqrt{h_r}h_l\alpha_{j,l}}{\sqrt{h_l}h_r + \sqrt{h_r}h_l}, \quad j = \{1, ..., N\}.$$

Remark 4. *We point out that (6.19) is a generalization of the mean values that are obtained with the Roe matrix for the usual Shallow Water equations.*

From (6.17) we obtain that $B_{i+\frac{1}{2}}$ is an evaluation of the non-conservative terms

$$B_{i+\frac{1}{2}} = diag(0, 0, -u_{m,b}, ..., -u_{m,b}), \tag{6.20}$$

at values

$$u_{m,b} = \begin{cases} \dfrac{h_r^2 u_r + h_l^2 u_l + h_l h_r \left[(u_l - u_r)log\left(\frac{h_r}{h_l}\right) - (u_r + u_l)\right]}{(h_r - h_l)^2} & \text{if } h_r \neq h_l, \\ \dfrac{u_r + u_l}{2} & \text{if } h_r = h_l. \end{cases}$$

Setting $A_{i+\frac{1}{2}} = J_{i+\frac{1}{2}} + B_{i+\frac{1}{2}}$ and $S_{i+\frac{1}{2}}$, it can be shown that $\mathcal{A}_{i+\frac{1}{2}}$ is a Roe matrix in the sense of [33].

For the polynomial viscosity matrix $Q_{i+\frac{1}{2}}$ an HLL-like method that correspond to choosing a polynomial approximation of the matrix Q as $P(x) = a_0 + a_1 x$ in (6.14) is used, see [12] for more details. The coefficients are given as

$$a_0 = \frac{S_r|S_l| - S_l|S_r|}{S_r - S_l}, \quad a_1 = \frac{|S_r| - |S_l|}{S_r - S_l},$$

where S_r and S_l are the maximum and the minimum eigenvalue of $A_{i+\frac{1}{2}}$, respectively.

Remark 5. *The eigenvalues of $A_{i+\frac{1}{2}}$ are computed numerically. However, it is possible to use the Cardano's formula to obtain exact eigenvalues.*

7 Numerical tests

In this section several tests with increasing complexity are considered to validate the results obtained starting from steady state initial conditions with the well-balanced first and second order schemes for the SWLME. Subsequently, we use a transient dam-break problem to compare the SWLME with the results obtained for the HSWME and the βHSWME, see [21]. For implementation details used in all examples of this section we refer to the implementation [20].

7.1 Well-balanced property

The first four test cases are intended to show that the scheme is effectively well-balanced. A 1000-
point uniform mesh, free boundary conditions and a CFL number of 0.5 are used. In all cases we
exemplarily use $N = 8$ moments and $g = 9.812$.

Test 1: Lake at rest

For the lake at rest, a zero velocity profile corresponding to water at rest with the following bottom topography is used in the spatial domain $[-1, 1]$

$$b_0(x) = \begin{cases} 2 - x^2 & \text{if } -0.5 < x < 0.5, \\ 1.75 & \text{otherwise,} \end{cases} \tag{7.1}$$

and therefore

$$W_0(x) = (h_0(x), u_{m,0}(x)h_0(x), \alpha_{1,0}(x)h_0(x), ..., \alpha_{N,0}(x)h_0(x)) = (3 - b_0(x), 0, 0, ..., 0). \tag{7.2}$$

The initial condition is shown in Figure 5. In Table 1 we observe that the well-balanced and
also the non well-balanced schemes of first and second order capture well the lake at rest. This
is due to the fact that straight lines are used as the paths in the numerical scheme. This is a
parameterization of the stationary solutions [4, 28]. For the first order test case, even the standard
non well-balanced scheme gives the right solution

Scheme (1000 cells)	$\|\Delta h\|_1$ (1st)	$\|\Delta u\|_1$ (1st)	$\|\Delta h\|_1$ (2nd)	$\|\Delta u\|_1$ (2nd)
Well-balanced	0.00	8.16e-16	0.00	8.16e-16
Non well-balanced	0.00	7.12e-16	4.51e-15	1.75e-14

Table 1: Well-balanced vs non well-balanced schemes: L^1 errors $\|\Delta \cdot \|_1$ at time $t = 0.5$ for the SWLME model with initial conditions (7.1) and (7.2).
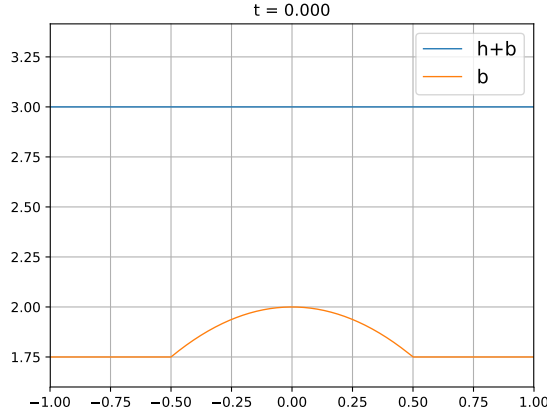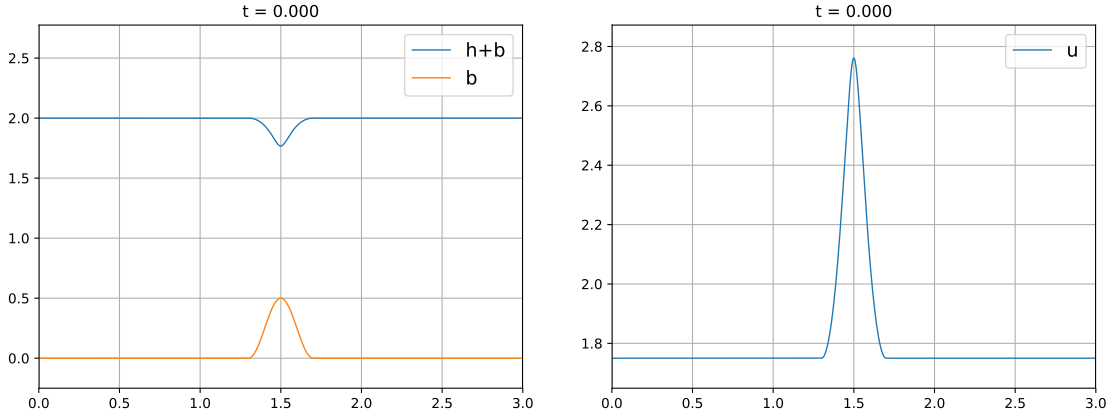

Figure 5: Initial condition for the lake at rest (7.1) and (7.2).

Test 2: Subcritical stationary solution

We consider a subcritical stationary solution as initial condition in the spatial domain $[0, 3]$, similar to [11]. The bottom topography is chosen as

$$b_0(x) = \begin{cases} 0.25(1 + cos(5\pi(x + 0.5))) & \text{if } 1.3 < x < 1.7, \\ 0 & \text{otherwise.} \end{cases} \tag{7.3}$$

As $W_0(x)$ we take the subcritical stationary solution such that $C_1 = 3.5$, $C_2 = 17.56957396120237$ and $C_i = 0$ for $i \in \{3, ..., N+2\}$. The initial condition is shown in Figure 6. In Table 2 we observe that our well-balanced schemes of first and second order capture well the subcritical stationary solution while the non well-balanced schemes do not. The non well-balanced scheme shows a clear error whereas the well-balanced scheme is exact up to almost machine prevision.

Scheme (1000 cells)	$\|\Delta h\|_1$ (1st)	$\|\Delta u\|_1$ (1st)	$\|\Delta h\|_1$ (2nd)	$\|\Delta u\|_1$ (2nd)
Well-balanced	9.16e-16	1.79e-15	1.42e-15	3.24e-15
Non well-balanced	2.48e-6	5.08e-6	3.21e-5	8.40e-5

Table 2: Well-balanced vs non well-balanced schemes: L^1 errors $\|\Delta \cdot \|_1$ at time $t = 0.5$ for the SWLME model with initial condition (7.3).

Figure 6: Initial condition for the subcritical stationary solution (7.3).

Test 3: Transcritical stationary solution

Next, we consider a transcritical stationary solution using an initial condition in $[0,3]$ similar to [11]. The bottom topography is chosen as

$$b_0(x) = \begin{cases} 0.25(1 + cos(5\pi(x + 0.5))) & \text{if } 1.3 < x < 1.7, \\ 0 & \text{otherwise.} \end{cases} \tag{7.4}$$

As $W_0(x)$ we take the transcritical stationary solution

$$W_0(x) = \begin{cases} W_*(x) & \text{if } x < 1.5 \\ W^*(x) & \text{if } x > 1.5 \end{cases} \tag{7.5}$$

where W_* and W^* are the subcritical and supercritical stationary solutions such that $C_1 = 2.5$, $C_2 = 21, 15525$ and $C_i = 0$ for $i \in \{3, ..., N+2\}$. The initial condition is shown in Figure 7. In Table 3 we observe that our well-balanced schemes of first and second order capture well the transcritical stationary solution while the non well-balanced schemes do not. Again, the non well-balanced schemes result in a large error while the well-balanced schemes achieve a very accurate steady state solution.

24

Scheme (1000 cells)	$\|\Delta h\|_1$ (1st)	$\|\Delta u\|_1$ (1st)	$\|\Delta h\|_1$ (2nd)	$\|\Delta u\|_1$ (2nd)
Well-balanced	3.53e-14	2.95e-13	3.53e-14	2.98e-13
Non well-balanced	1.46e-5	1.22e-4	3.07e-4	1.12e-3

Table 3: Well-balanced vs non well-balanced schemes: L^1 errors $\|\Delta \cdot \|_1$ at time $t = 0.5$ for the SWLME model with initial condition (7.4) and (7.5).
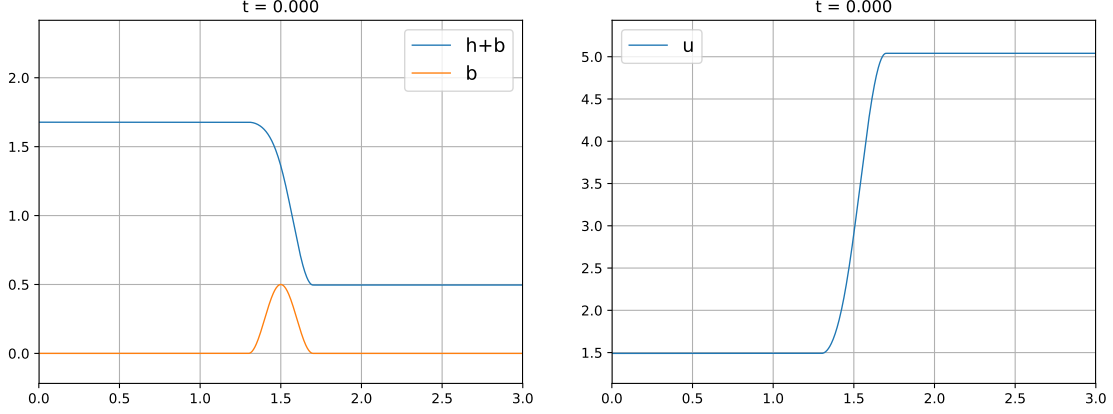

Figure 7: Initial condition for the subcritical stationary solution (7.4) and (7.5).

Test 4: Subcritical stationary solution with non zero moments

Lastly, we consider the following initial condition in $[0, 3]$ that is a subcritical stationary solution with non-vanishing coefficients α_i. The bottom topography is chosen as

$$b_0(x) = \begin{cases} 0.25(1 + cos(5\pi(x + 0.5))) & \text{if } 1.3 < x < 1.7 \\ 0 & \text{otherwise} \end{cases} \quad (7.6)$$

As $W_0(x)$ we use the subcritical stationary solution such that $C_1 = 3.5$, $C_2 = 21,15525$ and $C_i = 0.25$ for $i \in \{3, ..., N + 2\}$. The initial condition is shown in Figure 8. In Table 4 we observe that our well-balanced schemes of first and second order capture well the subcritical stationary solution while the non well-balanced schemes do not. Even in this test case with non-zero higher-order coefficients α_i the well-balanced scheme is much more accurate than the standard non well-balanced version.

Scheme	$\|\Delta h\|_1$, 1st	$\|\Delta u\|_1$ (1st)	$\|\Delta \alpha_i\|_1$ (1st)	$\|\Delta h\|_1$ (2nd)	$\|\Delta u\|_1$ (2nd)	$\|\Delta \alpha_i\|_1$ (2nd)
wb	4.00e-15	9.71e-15	4.45e-15	2.56e-15	7.66e-15	5.04e-15
Non wb	3.11e-6	6.65e-6	6.98e-7	3.98e-5	1.04e-4	2.52e-5

Table 4: Well-balanced (WB) vs non well-balanced schemes: L^1 errors $\|\Delta \cdot \|_1$ at time $t = 0.5$ for the SWLME model with initial condition (7.6).

7.2 Comparison between the SWLME, HSWME and βHSWME

In the following two tests, the results for the new SWLME model are compared with other hyperbolic models, HSWME and βHSWME, for which a Roe matrix was derived and explicitly given in the appendix A. These tests will be done in the spatial domain $[-0.4, 0.4]$ with $g = 1$ and $N = 8$. We consider a flat bottom topography ($b_x = 0$) and neglect friction terms. In this test case, the well-balanced property is of no interest, we therefore only compare the standard first and second order schemes.
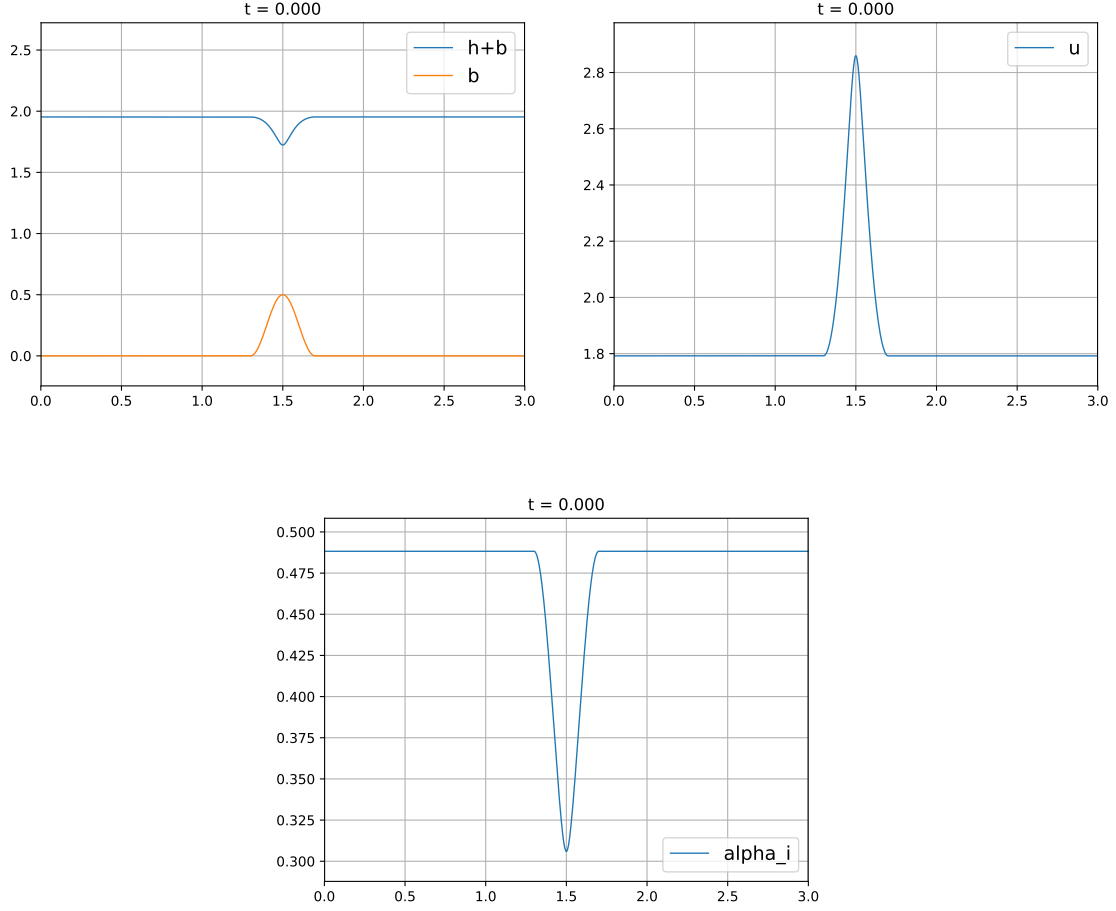
Figure 8: Initial condition for the subcritical stationary solution (7.6).

Test 5: transient model comparison with standard dam-break test
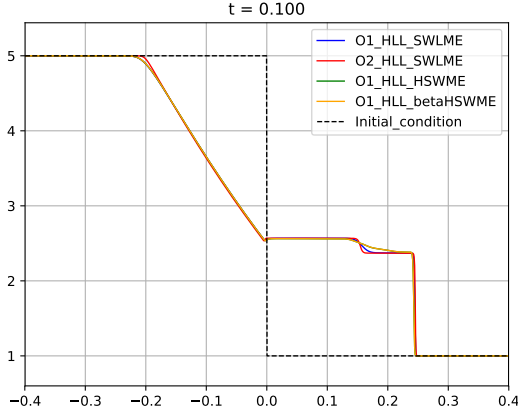
We are going to consider the following dam-break initial condition taken from [21] without friction terms

$$W_0(x) = (h_0(x), u_{m,0}(x)h_0(x), \alpha_{1,0}(x)h_0(x), ..., \alpha_{N,0}(x)h_0(x)), \tag{7.7}$$

where $u_{m,0}(x) = 0.25$, $\alpha_{1,0}(x) = -0.25$, $\alpha_{N,0}(x) = 0.25$, $\alpha_{i,0}(x) = 0, i \in \{2, ..., N-1\}$, and

$$h_0(x) = \left\{ \begin{array}{ll} 5 & \text{if } x < 0, \\ 1 & \text{if } x > 0. \end{array} \right. \tag{7.8}$$

Figure 9 shows the numerical results obtained with the first and second order scheme for the SWLME, and the first order schemes for the HSWME and the βHSWME. The results for the second order schemes applied to the latter two models are quantitatively the same as the first order results and thus omitted here. We can conclude that the results obtained are quite similar for all models in the variables h, u and α_1. As expected, the second order scheme captures the rarefaction waves better. We point out that the speed of the shock that travels from the left to the right is slightly higher in the case of the SWLME than in the other two models because in (5.8) we observe that all the α_i are taken into account for the maximum and minimum eigenvalues while in the HSWME and βHSWME only α_1 is contributing.
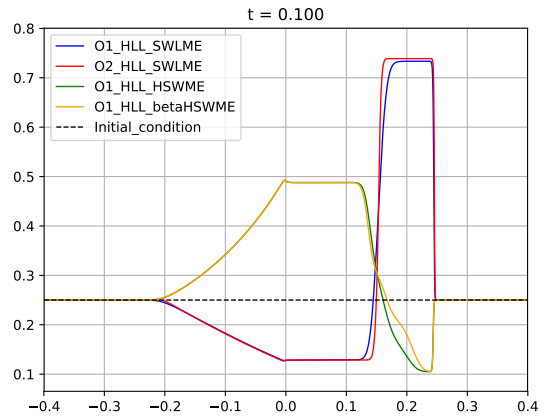
(a) Water height h.
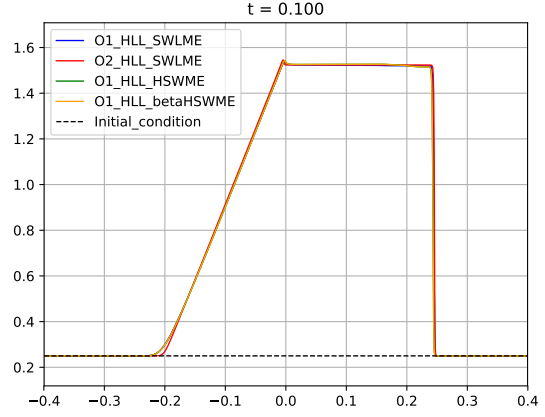
(b) Velocity u.
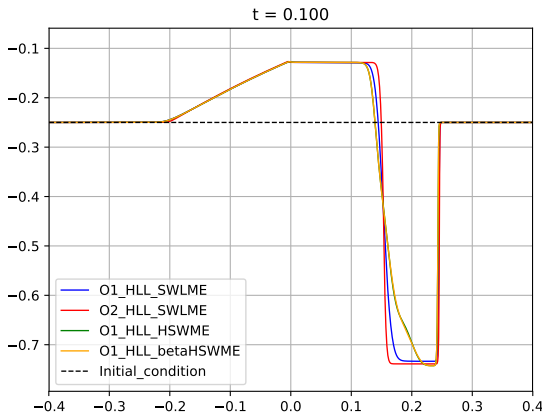
(c) First coefficient α_1.

(d) Last coefficient α_8.

Figure 9: Results obtained with the different models for the standard dam-break initial condition (7.7) and (7.8) for variables h, u, α_1, α_8 at t=0.1.

Test 6: transient model comparison with square root velocity profile

For the last test, we consider the following dam-break initial condition:

$$W_0(x) = (h_0(x), u_{m,0}(x)h_0(x), \alpha_{1,0}(x)h_0(x), ..., \alpha_{N,0}(x)h_0(x)), \tag{7.9}$$

where we use a square root initial velocity profile (4.1) $u(0, x, \zeta) = u_m(0, x) + \sum_{j=1}^{N} \alpha_j(0, x)\phi_j(\zeta) = \sqrt{\zeta}$, such that the initial variables can be computed according to (4.6) and (4.7) as $u_{m,0}(x) = 1$ and

$$\begin{array}{llll} \alpha_{1,0}(x) = -\frac{3}{5}, & \alpha_{2,0}(x) = -\frac{1}{7}, & \alpha_{3,0}(x) = -\frac{1}{15}, & \alpha_{4,0}(x) = -\frac{3}{77}, \\ \alpha_{5,0}(x) = -\frac{1}{39}, & \alpha_{6,0}(x) = -\frac{1}{55}, & \alpha_{7,0}(x) = -\frac{3}{221}, & \alpha_{8,0}(x) = -\frac{1}{95}. \end{array} \tag{7.10}$$

The initial water height is chosen as

$$h_0(x) = \left\{ \begin{array}{ll} 5 & \text{if } x < 0, \\ 1 & \text{if } x > 0. \end{array} \right. \tag{7.11}$$

In Figure 10 we show the numerical results obtained with the first and second order scheme for the SWLME, and the first order schemes for the HSWME and the βHSWME. We can conclude

27

that the results obtained are quite similar for all of them in the variables h, u and α_1. This is not the case for the variable α_8 where we can see that both the HSWME and the βHSWME result in strong oscillations. In comparison, the new SWLME is more stable than the other two models. Again the second order scheme captures the rarefaction waves better. Note that the emerging instability is not the result of an unstable high-order scheme, as the solutions for HSWME and βHSWME are even unstable with the first order scheme, while the SWLME yields stable results for both schemes. We point out that in this test the difference between the speed of the shock is even higher in the SWLME than in the other test because this time all the α_i have a non-zero initial value.
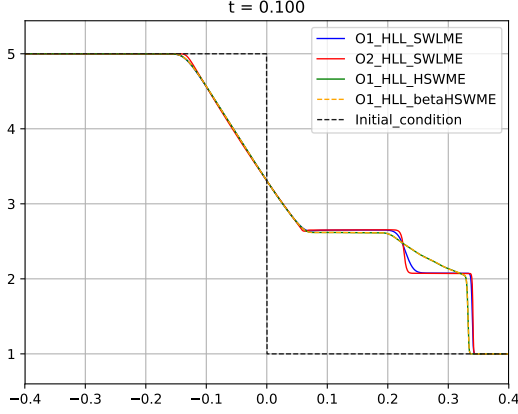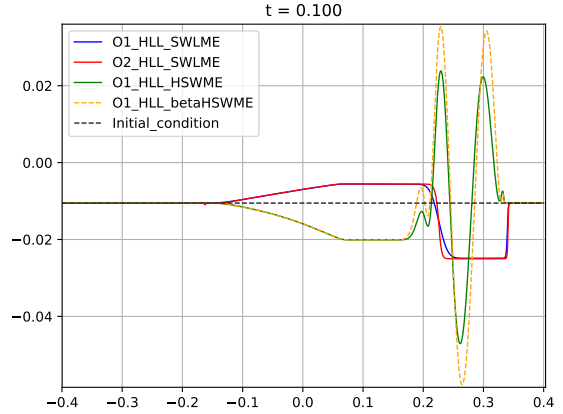

(a) Water height h.
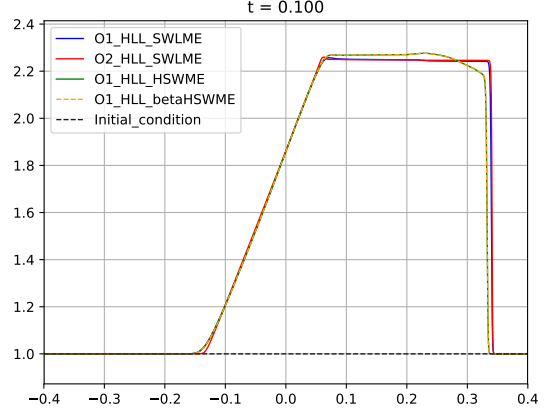
(b) Velocity u.
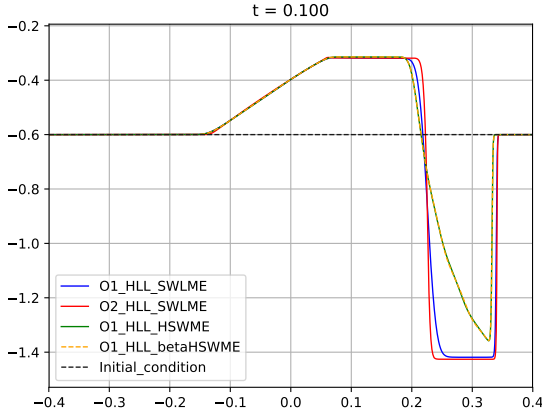
(c) First coefficient α_1.

(d) Last coefficient α_8.

Figure 10: Results obtained with the different models for the dam-break with square root velocity profile initial condition (7.9) and (7.11) for variables h, u, α_1, α_8 at t=0.1.

8 Conclusion

In this paper, we analytically and numerically investigate steady states of Shallow Water Moment Equations (SWME). After showing that the steady states for the SWME with $N = 1$ are extensions of the standard Shallow Water Equations (SWE), we pointed out that the case for arbitrary N poses difficulties due to the loss of hyperbolicity and the structure of the model. The analysis was generalized with the help of a newly derived model called Shallow Water Linearized Moment

Equations (SWLME), based on a linearization during the derivation. The concise derivation of the SWLME allowed to prove hyperbolicity and to fully characterize its eigenstructure analytically. This information was used to define a first order and a second order well-balanced numerical scheme preserving the steady states of the model numerically up to machine precision. Numerical results for lake-at-rest, subcritical, and transcritical initial conditions showed the success of the numerical scheme. Additionally, we compared the new SWLME model to other existing shallow water moment models, obtaining very similar solutions for the standard dam-break test. The solution for a more complex velocity profile seems more stable with the new SWLME model while existing models show emerging instabilities.

The current work is a major step towards a better understanding of shallow water moment models and opens up many possibilities for future work and applications. Next steps could be a detailed stability analysis of the models including the right hand side friction terms, which were neglected in this paper, or the design of proper implicit numerical scheme for potentially stiff friction terms. An extension towards well-balanced schemes of higher-order is possible following the construction of the second order scheme in this paper.

Acknowledgements

A   HSWME and β-HSWME models

The HSWME and β-HSWME models are derived and explicitly given in [21]. For our numerical schemes, we can write these two models in the form (6.1) where the conservative flux is given by

$$F^{HSWME}(U) = F^{\beta HSWME}(U) = \begin{pmatrix} hu_m \\ hu_m^2 + g\frac{h^2}{2} + \frac{1}{3}h\alpha_1^2 \\ 2hu_m\alpha_1 \\ \frac{2}{3}h\alpha_1^2 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

the non-conservative matrix is given by

$$B^{HSWME}(U) = \begin{pmatrix} 0 & 0 & & & & \\ 0 & 0 & & & & \\ & -u_m & \frac{3}{5}\alpha_1 & & & \\ & -\alpha_1 & u_m & \frac{4}{7}\alpha_1 & & \\ & & \frac{2}{5}\alpha_1 & \ddots & \ddots & \\ & & & \ddots & \ddots & \frac{N+1}{2N+1}\alpha_1 \\ & & & & \frac{N-1}{2N-1}\alpha_1 & u_m \end{pmatrix},$$

$$B^{\beta HSWME}(U) = \begin{pmatrix} 0 & 0 & & & & & \\ 0 & 0 & & & & & \\ -u_m & \frac{3}{5}\alpha_1 & & & & & \\ -\alpha_1 & u_m & \frac{4}{7}\alpha_1 & & & & \\ & \frac{2}{5}\alpha_1 & \ddots & & \ddots & & \\ & & \ddots & & \ddots & & \frac{N+1}{2N+1}\alpha_1 \\ & & & & \beta_N + \frac{N-1}{2N-1}\alpha_1 & u_m \end{pmatrix},$$

with $\beta_N = \frac{N^2-N}{2N^2+N-1}\alpha_1$ the parameter of the β-HSWME model. The source term S is the same as for the SWLME model.

The respective terms for the generalized Roe scheme from (6.15), (6.16), (6.17) and (6.18) can be obtained by:

$$J^{HSWME}_{i+\frac{1}{2}} = J^{\beta HSWME}_{i+\frac{1}{2}} = \frac{\partial F}{\partial U}(h_R, u_{m,R}, \alpha_{1,R}, ..., \alpha_{N,R}),$$

using (6.19), and

$$B^{HSWME}_{i+\frac{1}{2}} = B^{HSWME}(u_{m,b}, \alpha_{1,b}), \quad B^{\beta HSWME}_{i+\frac{1}{2}} = B^{\beta HSWME}(u_{m,b}, \alpha_{1,b}),$$

where the Roe averages are given by

$$u_{m,b} = \begin{cases} \frac{h_r^2 u_{m,r} + h_l^2 u_{m,l} + h_l h_r \left[(u_{m,l} - u_{m,r})log\left(\frac{h_r}{h_l}\right) - (u_{m,r} + u_{m,l}) \right]}{(h_r - h_l)^2} & \text{if } h_r \neq h_l, \\ \frac{u_{m,r} + u_{m,l}}{2} & \text{if } h_r = h_l, \end{cases}$$

and

$$\alpha_{1,b} = \begin{cases} \frac{h_r^2 \alpha_{1,r} + h_l^2 \alpha_{1,l} + h_l h_r \left[(\alpha_{1,l} - \alpha_{1,r})log\left(\frac{h_r}{h_l}\right) - (\alpha_{1,r} + \alpha_{1,l}) \right]}{(h_r - h_l)^2} & \text{if } h_r \neq h_l, \\ \frac{\alpha_{1,r} + \alpha_{1,l}}{2} & \text{if } h_r = h_l. \end{cases}$$

References

[1] F. Alcrudo and F. Benkhaldoun. Exact solutions to the Riemann problem of the shallow water equations with a bottom step. *Computers & Fluids*, 30(6):643–671, 2001. 1

[2] E. Audusse, F. Bouchut, M.-O. Bristeau, R. Klein, and B. Perthame. A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows. *SIAM Journal on Scientific Computing*, 25(6):2050–2065, 2004. 2

[3] J. P. Berberich, P. Chandrashekar, and C. Klingenberg. High order well-balanced finite volume methods for multi-dimensional systems of hyperbolic balance laws. *arXiv preprint arXiv:1903.05154*, 2019. 2

[4] A. Bermudez and M. E. Vázquez-Cendón. Upwind methods for hyperbolic conservation laws with source terms. *Computers & Fluids*, 23(8):1049–1071, 1994. 2, 23

[5] R. Bernetti, V. A. Titarev, and E. F. Toro. Exact solution of the Riemann problem for the shallow water equations with discontinuous bottom geometry. *Journal of Computational Physics*, 227(6):3212–3243, 2008. 1

[6] Z. Cai, Y. Fan, and R. Li. Globally hyperbolic regularization of Grad's moment system in one dimensional space. *Commun. Math. Sci.*, 11(2):547–571, 2013. 11

[7] A. Canestrelli, A. Siviglia, M. Dumbser, and E. F. Toro. Well-balanced high-order centred schemes for non-conservative hyperbolic systems. Applications to shallow water equations with fixed and mobile bed. *Advances in Water Resources*, 32(6):834–844, 2009. 2

[8] M. J. Castro-Díaz, T. Chacón Rebollo, E. D. Fernández-Nieto, and C. Parés. On well-balanced finite volume methods for nonconservative nonhomogeneous hyperbolic systems. *SIAM Journal on Scientific Computing*, 29(3):1093–1126, 2007. 2

[9] M. J. Castro-Díaz and E. D. Fernández-Nieto. A class of computationally fast first order finite volume solvers: PVM methods. *SIAM Journal on Scientific Computing*, 34(4):A2173–A2196, 2012. 15, 21

[10] M. J. Castro-Díaz, J. M. Gallardo, J. A. López-García, and C. Parés. Well-balanced high order extensions of Godunov's method for semilinear balance laws. *SIAM Journal on Numerical Analysis*, 46(2):1012–1039, 2008. 2, 17

[11] M. J. Castro-Díaz, J. A. López-García, and C. Parés. High order exactly well-balanced numerical methods for shallow water systems. *Journal of Computational Physics*, 246:242–264, 2013. 23, 24

[12] M. J. Castro-Díaz, T. Morales de Luna, and C. Parés. Well-balanced schemes and path-conservative numerical methods. In *Handbook of Numerical Analysis*, volume 18, pages 131–175. Elsevier, 2017. 2, 22

[13] M. J. Castro-Díaz and C. Parés. Well-balanced high-order finite volume methods for systems of balance laws. *Journal of Scientific Computing*, 82(2):48, 2020. 2, 16, 17, 18, 20, 21

[14] T. Chacón Rebollo, A. Dominguez Delgado, and E. D. Fernández-Nieto. A family of stable numerical solvers for the shallow water equations with source terms. *Computer methods in applied mechanics and engineering*, 192(1-2):203–225, 2003. 2

[15] M. Christen, J. Kowalski, and P. Bartelt. RAMMS: Numerical simulation of dense snow avalanches in three-dimensional terrain. *Cold Regions Science and Technology*, 63(1-2):1–14, 2010. 1

[16] R. V. Craster and O. K. Matar. Dynamics and stability of thin liquid films. *Reviews of modern physics*, 81(3):1131, 2009. 1

[17] J. Garres-Díaz, T. Morales de Luna, M. J. Castro, and J. Koellermeier. Shallow water moment models for bedload transport problems. *submitted*. 1

[18] C. W. Gear and I. Kevrekidis. Projective methods for stiff differential equations: Problems with gaps in their eigenvalue spectrum. *SIAM J. Sci. Comput.*, 24:1091–1106, 2003. 7

[19] S. Gottlieb and C. W. Shu. Total variation diminishing Runge-Kutta schemes. *Mathematics of Computation*, 67(221):73–85, 1998. 21

[20] J. Koellermeier and E. Pimentel-García. Software for: Steady states and well-balanced schemes for shallow water moment equations with topography. Zenodo, 2020. http://doi.org/10.5281/zenodo.4274991. 22

[21] J. Koellermeier and M. Rominger. Analysis and numerical simulation of hyperbolic shallow water moment equations. *Commun. Comp. Phys.*, 28((3)):1038–1084, 2020. 1, 7, 8, 9, 10, 11, 12, 14, 16, 22, 26, 29

[22] J. Koellermeier and M. Torrilhon. Simplified hyperbolic moment equations. In *Proceedings of the 16th International Conference on Hyperbolic Problems*, 2016. 11

[23] J. Kowalski and M. Torrilhon. Moment approximations and model cascades for shallow flow. *Communications in Computational Physics*, 25, 2019. 1, 3, 5, 6, 7, 8

[24] P. Lafitte and G. Samaey. Asymptotic-preserving projective integration schemes for kinetic equations in the diffusion limit. *SIAM Journal on Scientific Computing*, 34:A579–A602, 2012. 7

[25] P. G. LeFloch and M. D. Thanh. The Riemann problem for the shallow water equations with discontinuous topography. *Communications in Mathematical Sciences*, 5(4):865–885, 2007. 1

[26] 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(2):474–499, 2006. 2

[27] C. Parés. Numerical methods for nonconservative hyperbolic systems: a theoretical framework. *SIAM Journal on Numerical Analysis*, 44(1):300–321, 2006. 2

[28] C. Parés and M. J. Castro-Díaz. On the well-balance property of Roe's method for nonconservative hyperbolic systems. Applications to shallow-water systems. *ESAIM: mathematical modelling and numerical analysis*, 38(5):821–852, 2004. 23

[29] G. Rosatti and L. Begnudelli. The Riemann problem for the one-dimensional, free-surface shallow water equations with a bed step: Theoretical analysis and numerical simulations. *Journal of Computational Physics*, 229(3):760–787, 2010. 1

[30] G. Russo and A. Khe. High order well-balanced finite volume schemes for systems of balance laws. 2008. 2

[31] J. B. Schijf and Johan C. Schönfled. Theoretical considerations on the motion of salt and fresh water. In *Minnesota International Hydraulic Convention*. IAHR, 1953. 1

[32] D. Serre. *Systems of Conservation Laws 1: Hyperbolicity, Entropies, Shock Waves*. Cambridge University Press, Cambridge, 1999. 9

[33] I. Toumi. A weak formulation of Roe's approximate Riemann solver. *Journal of Computational Physics*, 102(2):360–373, 1992. 21, 22

[34] B. Van Leer. Towards the ultimate conservative difference scheme I. the quest of monotonicity. In *Proceedings of the Third International Conference on Numerical Methods in Fluid Mechanics*, pages 163–168. Springer, 1973. 20