Journal of Computational Physics 399 (2019) 108917

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

Robust explicit relaxation technique for solving the Green-Naghdi equations ☆

Jean-Luc Guermond [b,*], Bojan Popov [b], Eric Tovar [a,b], Chris Kees [a]

[a] U.S. Army Engineer Research and Development Center, Coastal and Hydraulics Laboratory (ERDC-CHL), Vicksburg, MS 39180, USA
[b] Department of Mathematics, Texas A&M University 3368 TAMU, College Station, TX 77843, USA

ARTICLE INFO

Article history:
Received 5 February 2019
Received in revised form 27 August 2019
Accepted 28 August 2019
Available online 10 September 2019

Keywords:
Shallow water
Well-balanced approximation
Invariant domain
Green-Naghdi
Finite element method
Positivity preserving

ABSTRACT

This paper revisits an original relaxation technique introduced in Favrie and Gavrilyuk [9] for solving the Green-Naghdi equations. We propose a version of that method and a space/time approximation thereof that is scale invariant. The approximation in time is explicit and the approximation in space uses a length scale for the relaxation that is proportional to the mesh size. The new method is compatible with dry states and is provably positivity preserving under the appropriate CFL condition. The method is numerically validated against manufactured solutions and is illustrated by comparison with experimental results.

1. Introduction

The objective of the present paper is to revisit an original penalty technique introduced in Favrie and Gavrilyuk [9] for solving the Green-Naghdi modification of Saint-Venant's equations. The Green-Naghdi model was introduced in Su and Gardner [16]; this model is a better approximation of the free surface Euler equations than the shallow water model since it includes a first-order correction in terms of the shallowness parameter. In particular, the Green-Naghdi model allows for solitary wave solutions which are not supported by Saint-Venant's shallow water system. However, a major drawback of the Green-Naghdi model is that it involves third-order derivatives in space. This rules out any approximation technique that is explicit in time, since this would require the time step τ to behave like $\mathcal{O}(h^3)V^{-1}L^{-2}$, where h is the meshsize, V is a characteristic velocity scale, and L is a characteristic length scale. A large body of literature has been dedicated to the development of efficient implicit techniques to solve these equations; we refer to Duran and Marche [8] and the references therein. An innovative technique, going around the difficulty of solving a third-order partial differential equation, has been proposed in [9]. The method is based on a relaxation of the Green-Naghdi equation that makes the perturbed system hyperbolic; moreover, under reasonable assumptions, the relaxed system can be shown to converge to the Green-Naghdi equation when the relaxation parameter goes to infinity. This technique is very interesting since it yields a system of first-order conservation equations that is hyperbolic, hence opening the door for explicit time stepping. However,

☆ **Funding:** This material is based upon work supported in part by the National Science Foundation grants DMS-1619892 and DMS-1620058, by the Air Force Office of Scientific Research, USAF, under grant/contract number FA9550-18-1-0397, and by the Army Research Office under grant/contract number W911NF-19-1-0431. Permission was granted by the Chief of Engineers to publish this information.

* Corresponding author.
E-mail address: guermond@math.tamu.edu (J.-L. Guermond).

https://doi.org/10.1016/j.jcp.2019.108917
0021-9991/

this method has some deficiencies: it involves a dimensional relaxation parameter λ, it is incompatible with dry states, and no realistic approximation technique connecting the relaxation parameter λ with the time step and the mesh size has been proposed yet. The goal of this paper is to address these three questions. We propose a new version of the Favrie-Gavrilyuk relaxation technique that is scale invariant and compatible with dry states. Then, we propose a new numerical method that uses a relaxation length scale proportional to the mesh size, is explicit in time, and is provably positivity preserving (compatible with dry states) under an appropriate CFL condition. The proposed method is numerically validated against manufactured solutions and illustrated by comparison with experimental results.

The paper is organized as follows. In §2 we introduce the Green-Naghdi model and the Favrie-Gavrilyuk relaxation method. Our variation of the Favrie-Gavrilyuk technique that has all the properties listed above is described in §3. In this new model the relaxation is scale invariant, compatible with dry states, and provably positive. The model introduced in §3 is approximated in time and space in §4. Positivity is established under an appropriate CFL condition. We numerically illustrate the proposed algorithm in §5.

2. Green-Naghdi model and relaxation

In this section we review the Green-Naghdi model and the perturbation proposed in Favrie and Gavrilyuk [9].

2.1. The Green-Naghdi model

Let D be a polygonal domain in $\mathbb{R}^d$, $d\in\{1,2\}$, occupied by a body of water evolving in time under the action of gravity. Assuming that the deformation of the free surface is small compared to the water elevation, and the bottom topography z varies slowly, the problem can be well represented by Saint-Venant's shallow water model. More precisely, let $\mu=\frac{\mathsf{h}_0^2}{L^2}$ be the shallowness parameter, where h_0 is the typical water height and L is the typical horizontal lengthscale, then Saint-Venant's shallow water model is the $\mathcal{O}(\mu)$ approximation of the free surface Euler equations. Using $\boldsymbol{u}=(\mathsf{h},\boldsymbol{q})^{\mathsf{T}}$ as the dependent variable, where h is the water height and $\boldsymbol{q}$ is the flow rate, or discharge, the model is written follows:

$$\partial_t\boldsymbol{u}+\nabla\cdot\mathbb{f}(\boldsymbol{u})+\boldsymbol{b}(\boldsymbol{u},\nabla z)=0,\quad \text{a.e. }\boldsymbol{x}\in D,t\in\mathbb{R}_+. \tag{2.1}$$

The flux $\mathbb{f}(\boldsymbol{u})$ and the bathymetry source $\boldsymbol{b}(\boldsymbol{u},\nabla z)$ are given by

$$\mathbb{f}(\boldsymbol{u}):=\begin{pmatrix}\boldsymbol{q}^{\mathsf{T}}\\ \frac{1}{\mathsf{h}}\boldsymbol{q}\otimes\boldsymbol{q}+p(\boldsymbol{q})\mathbb{I}_d\end{pmatrix}\in\mathbb{R}^{(1+d)\times d},\quad \boldsymbol{b}(\boldsymbol{u},\nabla z):=\begin{pmatrix}0\\ g\mathsf{h}\nabla z\end{pmatrix}, \tag{2.2}$$

where $p(\boldsymbol{u})=\frac{1}{2}g\mathsf{h}^2$, $\mathbb{I}_d$ is the $d\times d$ identity matrix, and the mapping $z:D\ni\boldsymbol{x}\mapsto z(\boldsymbol{x})\in\mathbb{R}$ is the bottom topography, which we assume to be given. The problem (2.2) is hyperbolic, and its approximation using explicit time stepping and finite volumes, finite differences, or finite elements (either continuous or discontinuous) for the space approximation is very well documented in the literature; we refer the reader to the books of Bouchut [6] and Toro [18] for a review on this topic, and to the paper of Xing and Shu [19] for a survey on finite volume and DG methods for the shallow water system.

Unfortunately, Saint-Venant's shallow water model does not include dispersive effects, which makes the system incapable of supporting solitary waves. This defect has been addressed in Su and Gardner [16] and Green et al. [11]; the solution consists of keeping the $\mathcal{O}(\mu)$ corrections from the expansion of the free surface Euler equations. (A rigorous statement regarding this expansion can be found in Alvarez-Samaniego and Lannes [2, Thm. 6.2].) The balance equations are still (2.2) but the pressure p is modified as follows:

$$p(\boldsymbol{u})=\frac{1}{2}g\mathsf{h}^2+\frac{1}{3}\mathsf{h}^2\ddot{\mathsf{h}},\qquad \dot{\mathsf{h}}:=\partial_t\mathsf{h}+\boldsymbol{v}\cdot\nabla\mathsf{h},\qquad \ddot{\mathsf{h}}:=\partial_t\dot{\mathsf{h}}+\boldsymbol{v}\cdot\nabla\dot{\mathsf{h}}, \tag{2.3}$$

where the velocity $\boldsymbol{v}$ is defined by $\boldsymbol{v}:=\frac{1}{\mathsf{h}}\boldsymbol{q}$. Notice that due to the presence of the term $\ddot{\mathsf{h}}$, the pressure mapping $\boldsymbol{u}\mapsto p(\boldsymbol{u})$ is not a function but a second-order differential operator in space and time. Note also in passing that the term $\mathsf{h}^2\ddot{\mathsf{h}}$ in the pressure can also be rewritten as follows $-\mathsf{h}^3(\partial_t\nabla\cdot\boldsymbol{v}+\boldsymbol{v}\cdot\nabla\nabla\cdot\boldsymbol{v}-(\nabla\cdot\boldsymbol{v})^2)$, which is exactly the expression given in Su and Gardner [16, A14] a few years earlier than Green et al. [11]. Although [16] predates [11], we are going to adopt the terminology that is currently dominant in the literature and refer to (2.1)–(2.3) as the Green-Naghdi model. The model (2.1)–(2.3) is a $\mathcal{O}(\mu^2)$ approximation of the free surface Euler equations. This model is no longer hyperbolic, and it cannot be reasonably approximated using explicit time stepping since in this case the time step τ would behave like $\mathcal{O}(h^3)V^{-1}L^{-2}$, where h is the meshsize, V is a characteristic velocity, and L is a characteristic length scale.

Proposition 2.1. *Assume that $\boldsymbol{u}$ is a smooth solution to the Green-Naghdi system (2.1)–(2.3), then the following identity holds for all $t>0$ and all $\boldsymbol{x}\in D$:*

$$\partial_t\mathcal{E}(\boldsymbol{u})+\nabla\cdot\boldsymbol{\mathcal{F}}(\boldsymbol{u})=0, \tag{2.4}$$

with the notation

$$\mathcal{E}(\boldsymbol{u}) := \frac{1}{2}g\mathsf{h}^2 + g\mathsf{z}\mathsf{h} + \frac{1}{2}\mathsf{h}\boldsymbol{v}^2 + \frac{1}{6}\mathsf{h}(\dot{\mathsf{h}})^2 \tag{2.5a}$$

$$\boldsymbol{\mathcal{F}}(\boldsymbol{u}) := \boldsymbol{v}\big(\mathcal{E}(\boldsymbol{u}) + p(\boldsymbol{u})\big). \tag{2.5b}$$

Proof. Since this result is standard, we only sketch the proof.
(i) We multiply the mass conservation by $g(\mathsf{h}+z)$ and the momentum equation by $\boldsymbol{v}$, use the mass conservation equation, and add the results:

$$\partial_t(\tfrac{1}{2}g\mathsf{h}^2 + g\mathsf{z}\mathsf{h} + \tfrac{1}{2}\mathsf{h}\boldsymbol{v}^2) + \nabla\cdot(\boldsymbol{v}(\tfrac{1}{2}g\mathsf{h}^2 + g\mathsf{z}\mathsf{h} + \tfrac{1}{2}\mathsf{h}\boldsymbol{v}^2 + \tfrac{1}{2}g\mathsf{h}^2)) + \boldsymbol{v}\cdot\nabla(\tfrac{1}{3}\mathsf{h}^2\ddot{\mathsf{h}}) = 0.$$

(ii) Now, using that $\dot{\mathsf{h}} := \partial_t\mathsf{h} + \boldsymbol{v}\cdot\nabla\mathsf{h} = -\mathsf{h}\nabla\cdot\boldsymbol{v}$, we multiply the identity $\ddot{\mathsf{h}} := \partial_t\dot{\mathsf{h}} + \boldsymbol{v}\cdot\nabla\dot{\mathsf{h}}$, by $\frac{1}{3}\mathsf{h}^2\nabla\cdot\boldsymbol{v}$ and obtain

$$(\nabla\cdot\boldsymbol{v})\tfrac{1}{3}\mathsf{h}^2\ddot{\mathsf{h}} = -\tfrac{1}{3}\mathsf{h}\dot{\mathsf{h}}(\partial_t\dot{\mathsf{h}} + \boldsymbol{v}\cdot\nabla\dot{\mathsf{h}}) = -\partial_t(\tfrac{1}{6}\mathsf{h}(\dot{\mathsf{h}})^2) - \nabla\cdot(\tfrac{1}{6}\boldsymbol{v}\mathsf{h}(\dot{\mathsf{h}})^2).$$

We conclude by adding the above identities and recalling that $p(\boldsymbol{u}) = \frac{1}{2}g\mathsf{h}^2 + \frac{1}{3}\mathsf{h}^2\ddot{\mathsf{h}}$. □

2.2. Green-Naghdi with the constraint $\{q_1 = \mathsf{h}^2\}$

In this section we show that the Green-Naghdi model can be reformulated as a constrained system of conservation equations. This point of view is slightly different from that in Favrie and Gavrilyuk [9], but it should help readers who are not familiar with Hamiltonian mechanics to understand why the method proposed in [9] is indeed a good proxy for the Green-Naghdi model. We are going to review the Favrie-Gavrilyuk model in §2.3.

Upon setting $q_2 := \mathsf{h}\dot{\mathsf{h}}$, $s := -\mathsf{h}\ddot{\mathsf{h}}$, and using the mass conservation equation, the pressure equation (2.3) can be equivalently rewritten as follows (provided $\mathsf{h} > 0$):

$$p = \frac{1}{2}g\mathsf{h}^2 - \frac{1}{3}\mathsf{h}s, \qquad \frac{q_2}{\mathsf{h}} = \partial_t\mathsf{h} + \boldsymbol{v}\cdot\nabla\mathsf{h}, \qquad -s = \partial_t q_2 + \nabla\cdot(\boldsymbol{v}q_2). \tag{2.6}$$

We now introduce the additional variable q_1 by setting $q_1 := \mathsf{h}^2$. Then (2.3) can equivalently be rewritten as follows:

$$p = \frac{1}{2}g\mathsf{h}^2 - \frac{1}{3}\frac{q_1}{\mathsf{h}}s, \qquad q_1 = \mathsf{h}^2, \qquad \partial_t q_1 + \nabla\cdot(\boldsymbol{v}q_1) = q_2, \qquad \partial_t q_2 + \nabla\cdot(\boldsymbol{v}q_2) = -s. \tag{2.7}$$

Here we used that $\mathsf{h}(\partial_t\phi + \boldsymbol{v}\cdot\nabla\phi) = \partial_t(\mathsf{h}\phi) + \nabla\cdot(\boldsymbol{v}\mathsf{h}\phi)$ for any smooth function ϕ to establish that $q_2 = \partial_t q_1 + \nabla\cdot(\boldsymbol{v}q_1)$ and $-s = \partial_t q_2 + \nabla\cdot(\boldsymbol{v}q_2)$. As a result, upon setting $\boldsymbol{u} := (\mathsf{h}, \boldsymbol{q}, q_1, q_2)^{\mathsf{T}}$, and under the constraint

$$q_1 = \mathsf{h}^2, \quad \text{a.e. } (\boldsymbol{x}, t) \in D\times\mathbb{R}_+, \tag{2.8}$$

the Green-Naghdi model is equivalent to

$$\partial_t\boldsymbol{u} + \nabla\cdot\mathbb{g}(\boldsymbol{u}, s) + \boldsymbol{c}(\boldsymbol{u}, s, \nabla z) = 0, \quad \text{a.e. } (\boldsymbol{x}, t) \in D\times\mathbb{R}_+, \tag{2.9}$$

$$\mathbb{g}(\boldsymbol{u}, s) := \begin{pmatrix} \mathsf{h}\boldsymbol{v} \\ \boldsymbol{q}\otimes\boldsymbol{v} + p(\boldsymbol{u}, s)\mathbb{I}_d \\ q_1\boldsymbol{v} \\ q_2\boldsymbol{v} \end{pmatrix} \in \mathbb{R}^{(3+d)\times d}, \qquad \boldsymbol{c}(\boldsymbol{u}, s, \nabla z) := \begin{pmatrix} 0 \\ g\mathsf{h}\nabla z \\ -q_2 \\ s \end{pmatrix}, \tag{2.10}$$

where the pressure is defined by

$$p(\boldsymbol{u}, s) = \frac{1}{2}g\mathsf{h}^2 - \frac{1}{3}\frac{q_1}{\mathsf{h}}s. \tag{2.11}$$

The initial data for the system (2.9)-(2.10)-(2.11) are

$$\mathsf{h}(\boldsymbol{x}, 0) = \mathsf{h}_0(\boldsymbol{x}), \qquad \boldsymbol{v}(\boldsymbol{x}, 0) = \boldsymbol{v}_0(\boldsymbol{x}), \tag{2.12}$$

$$q_1(\boldsymbol{x}, 0) = \mathsf{h}_0^2(\boldsymbol{x}), \qquad q_2(\boldsymbol{x}, 0) = -\mathsf{h}_0^2(\boldsymbol{x})\nabla\cdot\boldsymbol{v}_0(\boldsymbol{x}), \tag{2.13}$$

where the initial condition on q_2 is obtained by observing that the mass conservation equation implies that $q_2 = -\mathsf{h}^2\nabla\cdot\boldsymbol{v}$. Here s can be interpreted as the Lagrange multiplier associated with the constraint $q_1 = \mathsf{h}^2$. The above argument can be summarized as follows.

Proposition 2.2. *Assuming enough smoothness in time and space, the Green-Naghdi model is equivalent to* (2.9)-(2.10)-(2.11) *on the manifold* $\{q_1 = \mathsf{h}^2\}$.

The question addressed in this paper is that of relaxing the constraint $\{q_1 = \mathsf{h}^2\}$ so that the system (2.9)-(2.10)-(2.11) augmented with the relaxed constraint becomes hyperbolic, remains compatible with dry states, and is scale invariant.

2.3. The Favrie-Gavrilyuk model

In Favrie and Gavrilyuk [9], the authors propose a relaxation technique for the constraint $\{q_1 = \mathsf{h}^2\}$ based on an Hamiltonian mechanics point of view. Their approach consists of observing that the Green-Naghdi model admits a variational formulation from the following master Lagrangian:

$$\mathcal{L}(\mathsf{h}, \boldsymbol{q}) = \int_D \left(\frac{1}{2}\mathsf{h}^{-1}\boldsymbol{q}^2 - \frac{1}{2}g\mathsf{h}^2 + \frac{1}{6}\mathsf{h}\dot{\mathsf{h}}^2 \right) \mathrm{d}x.$$

Then the authors modify the Lagrangian functional through an augmented Lagrangian technique with a relaxation parameter λ that scales like the square of a velocity. Three possible perturbations are proposed in [9], but there is only one that is unconditionally hyperbolic and has a phase velocity that converges to that of the Green-Naghdi model as the relaxation parameter λ goes to infinity. The model in question consists of adding two new scalar variables (q_1, q_2), just like in (2.10), and expressing the balance equations as follows:

$$\partial_t \boldsymbol{u} + \nabla\cdot\mathbb{g}\left(\boldsymbol{u}, \lambda(\tfrac{q_1}{\mathsf{h}^2} - 1)\right) + \boldsymbol{c}\left(\boldsymbol{u}, \lambda(\tfrac{q_1}{\mathsf{h}^2} - 1), \nabla z\right) = 0, \quad \text{a.e. } \boldsymbol{x} \in D, t \in \mathbb{R}_+. \tag{2.14}$$

Notice that, referring to (2.10) for the definition of $\mathbb{g}$ and slightly abusing the notation, the pressure is defined in (2.14) by

$$p(\boldsymbol{u}) := \frac{1}{2}g\mathsf{h}^2 - \frac{1}{3}\frac{q_1}{\mathsf{h}}\lambda\left(\frac{q_1}{\mathsf{h}^2} - 1\right). \tag{2.15}$$

In conclusion, by comparing (2.9)-(2.10)-(2.11) with (2.14)-(2.15), we observe that the Hamiltonian mechanic strategy adopted in [9] to relax the constraint $\{q_1 = \mathsf{h}^2\}$ simply consists of replacing s in (2.10)-(2.11) by $\lambda(\frac{q_1}{\mathsf{h}^2} - 1)$, where $\lambda > 0$ is some ad hoc large dimensional parameter scaling like the square of a velocity.

Although the argumentation and the numerical results reported in [9] are convincing, we see three problems with the current state of this formulation. First, the quantity λ is dimensional and somewhat ad hoc. No real insight is given in [9] on how λ should be chosen. Second, there is no mechanism in the definition of the pressure (2.15) that keeps the pressure and the wave speed finite when the water height becomes small. More precisely, letting $\overline{p}(\mathsf{h}, q_1) := -\frac{1}{3}\lambda\left(\frac{q_1}{\mathsf{h}^2} - 1\right)\frac{q_1}{\mathsf{h}}$; we observe that $\overline{p}(\mathsf{h}, q_1)$ may grow unboundedly as $\mathsf{h} \to 0$ unless there is some mechanism that guarantees that $q_1^2\mathsf{h}^{-3}$ is uniformly bounded. Moreover, it is shown in [9] (and again in Proposition 3.5 below) that the eigenvalues of the (one-dimensional) system are $v := \frac{1}{\mathsf{h}}q$ and $v \pm \sqrt{g\mathsf{h} + \frac{1}{3}\lambda\frac{q_1^2}{\mathsf{h}^4}}$. Hence, the system is indeed hyperbolic, but to ensure that the wave speed stays bounded when $\mathsf{h} \to 0$, the ratio $q_1^2\mathsf{h}^{-4}$ must be bounded uniformly with respect to h. But, it does not seem that this type of estimate can be deduced a priori from (2.10)-(2.11). Actually, we have tested this model numerically and observed that none of these conditions are satisfied. In other words, the model is not compatible with dry states, which from our point of view, is a major drawback. Third, it is not clear how the time and space approximation of (2.10)-(2.11) should be done to construct a convergent approximation of the Green-Naghdi model. In particular, it is not clear how the relaxation parameter λ should be connected to the approximation parameters (time step and mesh size). It is the purpose of the present paper is to address these questions and propose an alternative formulation to (2.14)-(2.15).

3. Scale invariant relaxation

The purpose of this section is to introduce a relaxation of the Green-Naghdi model (2.9)-(2.10)-(2.11) (or (2.1)–(2.3)) with the following properties: (i) be scale independent; (ii) be hyperbolic; (iii) be compatible with dry states.

3.1. Pressure and energy

To simplify the computations we adopt the same notation as in [9] and introduce the quantities $\eta := \frac{q_1}{\mathsf{h}}$, $\omega := \frac{q_2}{\mathsf{h}}$. We also introduce a small length scale ϵ, which we will later think of as being the mesh size when the model is approximated in space. Let $\overline{\lambda}$ be a non-dimensional number of order one. Let $\Gamma \in C^2(\mathbb{R}; [0, \infty))$ be some smooth non-negative function with the constraints $\Gamma(1) = 0$ and $\Gamma'(1) = 0$. We then replace s in (2.10)-(2.11) by $-\overline{\lambda}g\frac{\mathsf{h}^2}{\epsilon}\Gamma'(\frac{\eta}{\mathsf{h}})$. Notice that $g\frac{\mathsf{h}^2}{\epsilon}$ scales like the square of a velocity, which is what one should expect since $-\overline{\lambda}g\frac{\mathsf{h}^2}{\epsilon}\Gamma'(\frac{\eta}{\mathsf{h}})$ should be an ansatz for $-\mathsf{h}\ddot{\mathsf{h}}$. As will be established in Proposition 3.1, where we derive an energy (entropy) equality, the definition of the associated pressure should be $\widetilde{p}(\mathsf{h}, \eta) := -\frac{1}{3}\frac{\overline{\lambda}g}{\epsilon}\eta^3\partial_x(x^{-2}\Gamma(x))_{|x=\eta\mathsf{h}^{-1}}$. In conclusion, setting $\boldsymbol{u} := (\mathsf{h}, \boldsymbol{q}, q_1, q_2)^\intercal$, and recalling that $\eta := \frac{q_1}{\mathsf{h}}$, the problem we consider in the rest of this paper consists of solving the following set of equations:

$$\partial_t \mathsf{h} + \nabla\cdot(\mathsf{h}\boldsymbol{v}) = 0, \tag{3.1a}$$

$$\partial_t \boldsymbol{q} + \nabla\cdot(\boldsymbol{q}\otimes\boldsymbol{v} + p\mathbb{I}) + g\mathsf{h}\nabla z = 0, \tag{3.1b}$$

$$\partial_t q_1 + \nabla\cdot(q_1\boldsymbol{v}) = q_2, \tag{3.1c}$$
$$\partial_t q_2 + \nabla\cdot(q_2\boldsymbol{v}) = -\overline{\lambda} g \tfrac{\mathsf{h}^2}{\epsilon}\Gamma'(\tfrac{\eta}{\mathsf{h}}), \tag{3.1d}$$

for a.e. $\boldsymbol{x}\in D$, $t\in\mathbb{R}_+$, with the pressure defined by

$$p(\boldsymbol{u}) := \tfrac{1}{2}g\mathsf{h}^2 + \widetilde{p}(\mathsf{h},\eta), \tag{3.2a}$$
$$\widetilde{p}(\mathsf{h},\eta) := -\frac{1}{3}\frac{\overline{\lambda} g}{\epsilon}\eta^3\partial_x(x^{-2}\Gamma(x))_{|x=\eta\mathsf{h}^{-1}}, \tag{3.2b}$$

and the initial data

$$\mathsf{h}(\boldsymbol{x},0) = \mathsf{h}_0(\boldsymbol{x}), \qquad \boldsymbol{v}(\boldsymbol{x},0) = \boldsymbol{v}_0(\boldsymbol{x}), \tag{3.3a}$$
$$q_1(\boldsymbol{x},0) = \mathsf{h}_0^2(\boldsymbol{x}), \qquad q_2(\boldsymbol{x},0) = -\mathsf{h}_0^2(\boldsymbol{x})\nabla\cdot\boldsymbol{v}_0(\boldsymbol{x}). \tag{3.3b}$$

Proposition 3.1. *Assume that $\boldsymbol{u}$ is a smooth solution to (3.1)-(3.2)-(3.3). Assume $\epsilon>0$ is constant (i.e., does not depend on $\boldsymbol{x}$ and t). Then the following identity holds for all $t>0$ and all $\boldsymbol{x}\in D$:*

$$\partial_t E(\boldsymbol{u}) + \nabla\cdot\boldsymbol{F}(\boldsymbol{u}) = 0,$$

with the entropy pair $(E,\boldsymbol{F})$ defined by

$$E(\boldsymbol{u}) := \frac{1}{2}g\mathsf{h}^2 + gz\mathsf{h} + \frac{1}{2}\mathsf{h}\boldsymbol{v}^2 + \frac{1}{6}\mathsf{h}\omega^2 + \frac{\overline{\lambda} g}{3\epsilon}\mathsf{h}^3\Gamma(\frac{\eta}{\mathsf{h}}), \tag{3.4a}$$
$$\boldsymbol{F}(\boldsymbol{u}) := \boldsymbol{v}(E(\boldsymbol{u}) + p(\boldsymbol{u})). \tag{3.4b}$$

Proof. (i) As usual, we multiply the mass conservation by $g(\mathsf{h}+z)$ and the momentum equation by $\boldsymbol{v}$, use the mass conservation equation, and add the results:

$$\partial_t(\tfrac{1}{2}g\mathsf{h}^2 + gz\mathsf{h} + \tfrac{1}{2}\mathsf{h}\boldsymbol{v}^2) + \nabla\cdot(\boldsymbol{v}(\tfrac{1}{2}g\mathsf{h}^2 + gz\mathsf{h} + \tfrac{1}{2}\mathsf{h}\boldsymbol{v}^2 + \tfrac{1}{2}g\mathsf{h}^2)) + \boldsymbol{v}\cdot\nabla\widetilde{p}(\mathsf{h},\eta) = 0.$$

(ii) Now, we multiply (3.1c) by $\mathsf{h}\Gamma'(\frac{\eta}{\mathsf{h}})$ and we obtain

$$\mathsf{h}\Gamma'(\tfrac{\eta}{\mathsf{h}})\,(\partial_t(\mathsf{h}\eta) + \nabla\cdot(\mathsf{h}\eta\boldsymbol{v})) = \mathsf{h}^2\Gamma'(\tfrac{\eta}{\mathsf{h}})\,(\partial_t\eta + \boldsymbol{v}\cdot\nabla\eta) = \mathsf{h}^2\Gamma'(\tfrac{\eta}{\mathsf{h}})\omega.$$

Using the chain rule we obtain

$$\partial_t\Gamma(\tfrac{\eta}{\mathsf{h}}) = \Gamma'(\tfrac{\eta}{\mathsf{h}})(\tfrac{1}{\mathsf{h}}\partial_t\eta - \tfrac{\eta}{\mathsf{h}^2}\partial_t\mathsf{h}), \qquad \boldsymbol{v}\cdot\nabla\Gamma(\tfrac{\eta}{\mathsf{h}}) = \Gamma'(\tfrac{\eta}{\mathsf{h}})(\tfrac{1}{\mathsf{h}}\boldsymbol{v}\cdot\nabla\eta - \tfrac{\eta}{\mathsf{h}^2}\boldsymbol{v}\cdot\nabla\mathsf{h}).$$

This in turn implies that

$$\begin{aligned}
\mathsf{h}\Gamma'(\tfrac{\eta}{\mathsf{h}})\big(\partial_t(\mathsf{h}\eta) + \nabla\cdot(\mathsf{h}\eta\boldsymbol{v})\big) &= \mathsf{h}^2\Gamma'(\tfrac{\eta}{\mathsf{h}})\,(\partial_t\eta + \boldsymbol{v}\nabla\cdot\eta) = \mathsf{h}^3\Gamma'(\tfrac{\eta}{\mathsf{h}})\tfrac{1}{\mathsf{h}}\,(\partial_t\eta + \boldsymbol{v}\nabla\cdot\eta)\\
&= \mathsf{h}^3\left(\partial_t\Gamma(\tfrac{\eta}{\mathsf{h}}) + \boldsymbol{v}\cdot\nabla\Gamma(\tfrac{\eta}{\mathsf{h}})\right) + \mathsf{h}^3\tfrac{\eta}{\mathsf{h}^2}\Gamma'(\tfrac{\eta}{\mathsf{h}})(\partial_t\mathsf{h} + \boldsymbol{v}\cdot\nabla\mathsf{h})\\
&= \mathsf{h}^3\left(\partial_t\Gamma(\tfrac{\eta}{\mathsf{h}}) + \boldsymbol{v}\cdot\nabla\Gamma(\tfrac{\eta}{\mathsf{h}})\right) - \mathsf{h}^3\Gamma'(\tfrac{\eta}{\mathsf{h}})\tfrac{\eta}{\mathsf{h}^2}\mathsf{h}\nabla\cdot\boldsymbol{v}\\
&= \mathsf{h}\left(\partial_t(\mathsf{h}^2\Gamma(\tfrac{\eta}{\mathsf{h}})) + \boldsymbol{v}\cdot\nabla(\mathsf{h}^2\Gamma(\tfrac{\eta}{\mathsf{h}})) - \Gamma(\tfrac{\eta}{\mathsf{h}})(\partial_t\mathsf{h}^2 + \boldsymbol{v}\cdot\nabla\mathsf{h}^2)\right) - \mathsf{h}^2\eta\Gamma'(\tfrac{\eta}{\mathsf{h}})\nabla\cdot\boldsymbol{v}\\
&= \partial_t(\mathsf{h}^3\Gamma(\tfrac{\eta}{\mathsf{h}})) + \nabla\cdot(\mathsf{h}^3\Gamma(\tfrac{\eta}{\mathsf{h}})\boldsymbol{v}) + (2\mathsf{h}^3\Gamma(\tfrac{\eta}{\mathsf{h}}) - \mathsf{h}^2\eta\Gamma'(\tfrac{\eta}{\mathsf{h}}))\nabla\cdot\boldsymbol{v}.
\end{aligned}$$

Hence, assuming that ϵ is constant, and using $\widetilde{p}(\mathsf{h},\eta) = \frac{1}{3}\frac{\overline{\lambda} g}{\epsilon}\left(2\mathsf{h}^3\Gamma(\frac{\eta}{\mathsf{h}}) - \mathsf{h}^2\eta\Gamma'(\frac{\eta}{\mathsf{h}})\right)$, we have established that

$$\partial_t(\frac{\overline{\lambda} g}{3\epsilon}\mathsf{h}^3\Gamma(\frac{\eta}{\mathsf{h}})) + \nabla\cdot(\frac{\overline{\lambda} g}{3\epsilon}\mathsf{h}^3\Gamma(\frac{\eta}{\mathsf{h}})\boldsymbol{v}) + \widetilde{p}(\mathsf{h},\eta)\nabla\cdot\boldsymbol{v} = \frac{\overline{\lambda} g}{3\epsilon}\mathsf{h}^2\Gamma'(\frac{\eta}{\mathsf{h}})\omega.$$

Note in passing that this computation justifies our definition of $\widetilde{p}(\mathsf{h},\eta)$ in (3.2b).

(iii) Finally, we multiply (3.1d) by $\frac{1}{3}\omega$ and obtain

$$\partial_t(\tfrac{1}{6}\mathsf{h}\omega^2) + \nabla\cdot(\tfrac{1}{6}\mathsf{h}\omega^2\boldsymbol{v}) = -\frac{\overline{\lambda} g}{3\epsilon}\mathsf{h}^2\Gamma'(\frac{\eta}{\mathsf{h}})\omega.$$

Now, we add the three identities obtained above, and the result follows readily. □

By comparing (3.4a) with (2.5a), and recalling that ω is meant to be an approximation of $\dot{\mathsf{h}}$, we observe $\frac{\overline{\lambda}g}{3\epsilon}\mathsf{h}^3\Gamma(\frac{\eta}{\mathsf{h}})$ is an extra energy induced by the relaxation. This quantity must be positive and should be minimum when $\eta=\dot{\mathsf{h}}$. Hence the design constraints for the energy functional Γ are $\Gamma(x)\geq 0$ for all $x\in\mathbb{R}$ and $\Gamma'(1)=0$. We can also set $\Gamma(1)=0$ since the energy level can be defined up to a constant. More precise statements regarding Γ are made in §3.3.

Corollary 3.2. *Let $\boldsymbol{u}$ be a smooth solution to (3.1)-(3.2)-(3.3). Assume $\epsilon>0$ is constant (i.e., does not depend on $\boldsymbol{x}$ and t). Let $T>0$. Assume that the boundary conditions for (3.1)-(3.2) are such that $F(\boldsymbol{u})\cdot\boldsymbol{n}_{\partial D}=0$ for all $t\in(0,T)$. Then, there is $c(\boldsymbol{u}_0)$ such that*

$$\int_0^T\int_D \mathsf{h}^3\Gamma(\eta\mathsf{h}^{-1})\,\mathrm{d}x\,\mathrm{d}t\leq c(\boldsymbol{u}_0)\epsilon.$$

Remark 3.3 *(Choice of source).* One may wonder whether changing the form of the source term in (3.1d) could change the behavior of the algorithm. For instance one could let the source term be $-\frac{\overline{\lambda}g}{\epsilon}\mathsf{h}^\alpha\eta^{2-\alpha}S(\frac{\eta}{\mathsf{h}})$ for some $\alpha\in\mathbb{R}$. One can then derive an energy as in Proposition 3.1 by introducing $\Gamma(x):=\int_1^x\xi^{2-\alpha}S(\xi)\,\mathrm{d}\xi$, since this definition implies that $\mathsf{h}^\alpha\eta^{2-\alpha}S(\frac{\eta}{\mathsf{h}})=\mathsf{h}^2\Gamma'(\frac{\eta}{\mathsf{h}})$. This argument shows that the form $-\frac{\overline{\lambda}g}{\epsilon}\mathsf{h}^2\Gamma'(\frac{\eta}{\mathsf{h}})$ for the source term is quite general.

Remark 3.4 *(Positivity of η).* Any reasonable solution of (3.1)-(3.2) satisfies $\mathsf{h}\geq 0$, but whether it also satisfies $\eta\geq 0$ is not clear. A possible way to enforce this additional property consists of modifying the system (3.1)-(3.2) as follows:

$$\partial_t\mathsf{h}+\nabla\cdot(\mathsf{h}\boldsymbol{v})=0, \tag{3.5a}$$
$$\partial_t\boldsymbol{q}+\nabla\cdot(\boldsymbol{q}\otimes\boldsymbol{v}+p\mathbb{I})+g\mathsf{h}\nabla z=0, \tag{3.5b}$$
$$\partial_t(q_1)+\nabla\cdot(q_1\boldsymbol{v})=q_2 r(\mathsf{h},\eta), \tag{3.5c}$$
$$\partial_t(q_2)+\nabla\cdot(q_2\boldsymbol{v})=-\overline{\lambda}g\tfrac{\mathsf{h}^2}{\epsilon}\Gamma'(\tfrac{\eta}{\mathsf{h}})r(\mathsf{h},\eta), \tag{3.5d}$$

where $r(\mathsf{h},\eta)$ is any smooth function satisfying the following properties $r(\mathsf{h},\mathsf{h})=1$, $r(\mathsf{h},0)=0$, for all $\mathsf{h}\in\mathbb{R}_+$. A possible candidate is $r(\mathsf{h},\eta)=\frac{2\eta^2}{\eta^2+\mathsf{h}^2}$.

3.2. Hyperbolicity of the model

We now verify whether the new model (3.1)-(3.2) is hyperbolic for all admissible states of the system.

Proposition 3.5 *(Hyperbolicity). Let $\mathbb{k}(\boldsymbol{u})$ be the conservative flux of the system (3.1)-(3.2). For any unit vector $\boldsymbol{n}\in\mathbb{R}^d$, the $d+3$ eigenvalues of the Jacobian matrix of the flux $\mathbb{k}(\boldsymbol{u})\boldsymbol{n}$ are $\mu_k=\boldsymbol{v}\cdot\boldsymbol{n}$, $k\in\{2{:}d+2\}$ and*

$$\mu_1=\boldsymbol{v}\cdot\boldsymbol{n}-\sqrt{g\mathsf{h}+\partial_{\mathsf{h}}\widetilde{p}(\mathsf{h},\eta)},\quad \mu_{d+3}=\boldsymbol{v}\cdot\boldsymbol{n}+\sqrt{g\mathsf{h}+\partial_{\mathsf{h}}\widetilde{p}(\mathsf{h},\eta)}. \tag{3.6}$$

The system (3.1)-(3.2) is hyperbolic iff the following holds for all $\eta\in\mathbb{R}$ and all $\mathsf{h}\geq 0$:

$$g\mathsf{h}\left(1+\frac{1}{3}\frac{\overline{\lambda}}{\epsilon}\eta\left(x^3\partial_{xx}(x^{-2}\Gamma(x))\right)_{|x=\eta\mathsf{h}^{-1}}\right)\geq 0. \tag{3.7}$$

Proof. We introduce the notation $v:=\frac{1}{\mathsf{h}}\boldsymbol{q}\cdot\boldsymbol{n}$ and $\boldsymbol{q}^\perp:=\boldsymbol{q}-(\boldsymbol{q}\cdot\boldsymbol{n})\boldsymbol{q}$, $\boldsymbol{v}^\perp:=\frac{1}{\mathsf{h}}\boldsymbol{q}^\perp$. Then we make a change of coordinates so that the conserved variable can be re-written $\boldsymbol{u}=(\mathsf{h},(\boldsymbol{q}\cdot\boldsymbol{n}),\boldsymbol{q}^\perp,q_1,q_2)^{\mathsf{T}}$. Recalling the notation $q_1=\mathsf{h}\eta$, we introduce the function $\overline{p}(\mathsf{h},q_1)$ such that $\overline{p}(\mathsf{h},q_1):=\widetilde{p}(\mathsf{h},\eta)$. The Jacobian matrix of the flux $\mathbb{k}(\boldsymbol{u})\boldsymbol{n}$ is then given by

$$\begin{pmatrix} 0 & 1 & \mathbf{0}^{\mathsf{T}} & 0 & 0\\ gh-v^2+\partial_{\mathsf{h}}\overline{p} & 2v & \mathbf{0}^{\mathsf{T}} & \partial_{q_1}\overline{p} & 0\\ -v\boldsymbol{v}^\perp & \boldsymbol{v}^\perp & v & 0 & 0\\ -v\eta & \eta & \mathbf{0}^{\mathsf{T}} & v & 0\\ -vw & w & 0 & 0 & v\end{pmatrix}.$$

Denoting by μ a generic eigenvalue, the characteristic polynomial is

$$(\mu-v)^{1+d}((\mu-v)^2-(g\mathsf{h}+\eta\partial_{q_1}\overline{p}+\partial_{\mathsf{h}}\overline{p})).$$

The eigenvalues are v, with multiplicity $d+1$, and $v-\sqrt{g\mathsf{h}+\eta\partial_{q_1}\overline{p}+\partial_{\mathsf{h}}\overline{p}}$ and $v+\sqrt{g\mathsf{h}+\eta\partial_{q_1}\overline{p}+\partial_{\mathsf{h}}\overline{p}}$, each with multiplicity 1. We now investigate the sign of $g\mathsf{h}+\eta\partial_{q_1}\overline{p}+\partial_{\mathsf{h}}\overline{p}$. Using that $\widetilde{p}(\mathsf{h},\eta):=\overline{p}(\mathsf{h},q_1)$, we observe that $\partial_{\mathsf{h}}\widetilde{p}=\eta\partial_{q_1}\overline{p}+\partial_{\mathsf{h}}\overline{p}$.

Using the change of variable $\mathsf{h}(\tau)=\tau^{-1}$, i.e., $\tau(\mathsf{h})=\mathsf{h}^{-1}$, we set $\widehat{\widetilde{p}}(\tau,\eta):=\widetilde{p}(\tau^{-1},\eta)$, and recalling (3.2b), this gives $\widehat{\widetilde{p}}(\tau,\eta)=-\frac{1}{3}\frac{\overline{\lambda}g}{\epsilon}\partial_\tau(\tau^{-2}\Gamma(\eta\tau))$. Using $\partial_{\mathsf{h}}\widetilde{p}(\mathsf{h},\eta)=-\tau^2\partial_\tau\widehat{\widetilde{p}}(\tau,\eta)$, this in turn implies that the following inequality

$$\begin{aligned}g\mathsf{h}+\partial_{\mathsf{h}}\widetilde{p}(\mathsf{h},\eta)&=g\mathsf{h}+\frac{1}{3}\frac{\overline{\lambda}g}{\epsilon}\tau^2\partial_{\tau\tau}(\tau^{-2}\Gamma(\eta\tau))\\&=g\mathsf{h}(1+\frac{\overline{\lambda}}{3\epsilon}\eta\times(x^3\partial_{xx}(x^{-2}\Gamma(x)))_{|x=\eta\mathsf{h}^{-1}})\geq 0,\end{aligned}$$

is a necessary and sufficient condition for hyperbolicity; whence (3.7). □

3.3. Construction of Γ

In this section we construct an energy functional Γ that satisfies the requirement identified in the previous sections. Smooth solutions of (3.1) satisfy the positivity constraints $\mathsf{h}\geq 0$, but we have not been able to establish that η stays nonnegative; we actually conjecture that this may not be the case. Hence it is important that $\Gamma(\eta\mathsf{h}^{-1})$ be well defined for $\eta\in\mathbb{R}$.

One way to proceed consists of adapting the augmented Lagrangian approach. Looking at (2.6), we replace the source s by a function $s(\mathsf{h},\eta)$, so that $\widetilde{p}(\mathsf{h},\eta)=-\frac{1}{3}\mathsf{h}s(\mathsf{h},\eta)$ and $\frac{\overline{\lambda}g}{\epsilon}\mathsf{h}^2\Gamma'(\frac{\eta}{\mathsf{h}})=s(\mathsf{h},\eta)$. This means that $\widetilde{p}(\mathsf{h},\eta)=-\frac{1}{3}\frac{\overline{\lambda}g}{\epsilon}\mathsf{h}\mathsf{h}^2\Gamma'(\frac{\eta}{\mathsf{h}})$. But, in order for the energy (entropy) argument from Proposition 3.1 to hold, one must also have $\widetilde{p}(\mathsf{h},\eta)=-\frac{1}{3}\frac{\overline{\lambda}g}{\epsilon}\eta^3\partial_x(x^{-2}\Gamma(x))_{x=\eta\mathsf{h}^{-1}}$. This is possible only if $\mathsf{h}^3\Gamma'(\frac{\eta}{\mathsf{h}})=\eta^3\partial_x(x^{-2}\Gamma(x))_{x=\eta\mathsf{h}^{-1}}$. This gives the ODE

$$\Gamma'(x)=x^3\partial_x(x^{-2}\Gamma(x)),\quad \Gamma(1)=0. \tag{3.8}$$

The solution is $\Gamma(x)=\alpha(1-x)^2$ with $\alpha\in\mathbb{R}$. The additional condition $\Gamma(x)\geq 0$ for all $x\in\mathbb{R}$ implies that $\alpha>0$. Unfortunately, this definition of Γ does not satisfy uniformly the hyperbolicity requirement (3.7); more precisely, $\partial_{xx}(x^{-2}\Gamma(x))=2\alpha x^{-4}(3-2x)$. Hence, hyperbolicity can be lost when $x>\frac{3}{2}$, i.e., $\eta>\frac{3}{2}\mathsf{h}$. In order to remedy to this problem, we redefine $\Gamma(x)$ for $x\geq 1$ by enforcing

$$x^\gamma\partial_{xx}(x^{-2}\Gamma(x))=2\alpha,\quad \Gamma(1)=0,\quad \Gamma'(1)=0, \tag{3.9}$$

with any γ such that $\gamma\geq 1$. Notice that if $\gamma<1$, then $|\partial_x(x^{-2}\Gamma(x))|\sim x^{1-\gamma}$ when $|x|\to\infty$, and $\widetilde{p}(\mathsf{h},\eta)$ behaves like $\eta^{4-\gamma}\mathsf{h}^{1-\gamma}$. But, (under the reasonable assumption that η stays bounded), saying that $x\to+\infty$ is equivalent to saying that $\mathsf{h}\to 0$. This argument shows that $\widetilde{p}(\mathsf{h},\eta)$ grows unboundedly as $\mathsf{h}\to 0$ and $\eta\mathsf{h}^{-1}\to\infty$ if $\gamma<1$. Note also that $\partial_x(\alpha x^{-2}(1-x)^2)=2\alpha$; hence (3.9) implies that Γ'' is continuous at 1; i.e., $\Gamma\in C^2(\mathbb{R};[0,\infty))$. Assuming that $\gamma\notin\{1,2\}$, the solution to (3.9) is $\Gamma(x)=\frac{2\alpha}{(\gamma-1)(\gamma-2)}x^{4-\gamma}+ax^3+bx^2$, with $a=\frac{2\alpha}{\gamma-1}$, $b=-\frac{2\alpha}{\gamma-2}$.

In the numerical tests reported at the end of the paper we use $\gamma=4$ and $\alpha=3$; this gives the following definition

$$\Gamma(x)=\begin{cases}3(1-x)^2 & \text{if } x\leq 1,\\ (1+2x)(1-x)^2 & \text{if } 1\leq x.\end{cases}$$

Notice that $\Gamma(x)\geq 0$ for any $x\in\mathbb{R}$, $\Gamma(1)=0$, $\Gamma'(1)=0$, and $\Gamma\in C^2(\mathbb{R};[0,\infty))$. The source term $\mathsf{h}^2\Gamma'(\eta\mathsf{h}^{-1})$ in the balance equation for $\mathsf{h}\omega$ and the pressure $\widetilde{p}(\mathsf{h},\eta)$ are given by

$$\mathsf{h}^2\Gamma'(\eta\mathsf{h}^{-1})=\begin{cases}6(\eta\mathsf{h}-\mathsf{h}^2), & \text{if } \eta\leq\mathsf{h}\\ 6(\eta^2-\eta\mathsf{h}), & \text{if } \mathsf{h}\leq\eta,\end{cases} \tag{3.10}$$

$$\widetilde{p}(\mathsf{h},\eta)=-\frac{\overline{\lambda}g}{3\epsilon}\times\begin{cases}6\mathsf{h}(\eta\mathsf{h}-\mathsf{h}^2), & \text{If } \eta\leq\mathsf{h}\\ 2(\eta^3-\mathsf{h}^3), & \text{If } \mathsf{h}\leq\eta.\end{cases} \tag{3.11}$$

The above definition of Γ implies that

$$\partial_{\mathsf{h}}\widetilde{p}(\mathsf{h},\eta)=g\mathsf{h}\frac{\overline{\lambda}}{3\epsilon}\times\begin{cases}6\mathsf{h}+12(\mathsf{h}-\eta), & \text{If } \eta\leq\mathsf{h},\\ 6\mathsf{h}, & \text{If } \mathsf{h}\leq\eta.\end{cases} \tag{3.12}$$

Hence the hyperbolicity condition (3.7) is satisfied for any pair $(\eta,\mathsf{h})\in\mathbb{R}\times\mathbb{R}_+$.

Remark 3.6 *(Choices in [9]).* Recall that in Favrie and Gavrilyuk [9], the source term for the conservation of q_2 is $-s:=-\lambda((\eta\mathsf{h}^{-1}-1)$, and, following (2.11), the pressure is defined by $\widetilde{p}(\mathsf{h},\eta)=-\frac{1}{3}\eta s$. We then observe that unless there exists an a priori upper bound on $\eta\mathsf{h}^{-1}$ when $\mathsf{h}\to 0$, the source term and the pressure may behave badly in dry regions. Contrary to [9], the choice that we make for the pressure in (2.6) when $\eta\leq\mathsf{h}$ is $\widetilde{p}(\mathsf{h},\eta)=-\frac{1}{3}\mathsf{h}s$ with $s=\frac{\overline{\lambda}g}{3\epsilon}\mathsf{h}^2\Gamma'(\frac{\eta}{\mathsf{h}})$; i.e., we use (2.6) to define the pressure instead of (2.11). We have observed numerically that the definition proposed in [9] is incompatible with dry states, whereas the definition that we propose is.

Remark 3.7 *(Uniform hyperbolicity).* Notice that the definition $\Gamma(x) := (1+2x)(1-x)^2$ gives a model that is uniformly hyperbolic with respect to (η, h) (see (3.12)), but we have observed that this model does not behave properly with dry states when $\eta \leq \mathsf{h}$. Instead, using $\Gamma(x) = 3(1-x)^2$ when $\eta \leq \mathsf{h}$ implies that $\mathsf{h}^2\Gamma'(\eta\mathsf{h}^{-1}) = 6\mathsf{h}(\eta - \mathsf{h})$ and $\widetilde{p}(\mathsf{h}, \eta) = -\frac{\overline{\lambda}g}{3\epsilon}6\mathsf{h}^2(\eta - \mathsf{h})$; hence $(\mathsf{h}^2\Gamma'(\eta\mathsf{h}^{-1}))/\mathsf{h}$ and $\widetilde{p}(\mathsf{h}, \eta)/\mathsf{h}^2$ stay bounded when $\mathsf{h} \to 0$ i.e., with our definition the dispersion effects vanish when $\eta \leq \mathsf{h}$ and $\mathsf{h} \to 0$.

4. Approximation

We describe in this section an explicit finite element technique to solve the proposed model. We use the method described in Guermond et al. [13] for the approximation in space and time. Without going through all the details we now recall the main features of the method.

4.1. Finite element setting

Given a shape-regular family of matching meshes $(\mathcal{T}_h)_{h>0}$, given a reference finite element $(\widehat{K}, \widehat{P}, \widehat{\Sigma})$, we introduce the continuous finite element space

$$P(\mathcal{T}_h) := \{v \in C^0(D;\mathbb{R}) \mid v_{|K} \circ T_K \in \widehat{P},\ \forall K \in \mathcal{T}_h\}, \tag{4.1}$$

where $T_K : \widehat{K} \to K$ is the geometric bijective transformation that maps the reference element $\widehat{K}$ to the current element K. We then set $\boldsymbol{P}(\mathcal{T}_h) := [P(\mathcal{T}_h)]^{3+d}$ to approximate the conserved variable $\boldsymbol{u} := (\mathsf{h}, \boldsymbol{q}, q_1, q_2)^{\mathsf{T}}$ in space. The approximation of the bottom topography is done in $P(\mathcal{T}_h)$. The global shape functions in $P(\mathcal{T}_h)$ are denoted by $\{\varphi_i\}_{i\in\mathcal{V}}$, i.e., $\dim(P(\mathcal{T}_h)) = \mathrm{card}(\mathcal{V})$. For any $i \in \mathcal{V}$, we set $\mathcal{I}(i) := \{j \in \mathcal{V} \mid \varphi_i\varphi_j \not\equiv 0\}$ and refer to $\mathcal{I}(i)$ as the stencil of the shape function φ_i. Finally, we define the following three quantities which will play an important role in the rest of the paper:

$$m_i = \int_D \varphi_i(\boldsymbol{x})\,\mathrm{d}x, \qquad \boldsymbol{c}_{ij} := \int_D \varphi_i \nabla\varphi_j\,\mathrm{d}x, \qquad \boldsymbol{n}_{ij} := \frac{\boldsymbol{c}_{ij}}{\|\boldsymbol{c}_{ij}\|_{\ell^2}} \qquad i,j \in \mathcal{V}. \tag{4.2}$$

Owing to the partition of unity property we have $\sum_{j\in\mathcal{V}} \boldsymbol{c}_{ij} = \boldsymbol{0}$. Notice that $\boldsymbol{c}_{ij} = -\boldsymbol{c}_{ji}$ if either φ_i or φ_j is zero on ∂D and $\sum_{i\in\mathcal{V}} \boldsymbol{c}_{ij} = \boldsymbol{0}$ if φ_j is zero on ∂D. These properties are essential to establish conservation.

4.2. Dependent and auxiliary variables representation

The bottom topography is henceforth approximated by $z_h = \sum_{i\in\mathcal{V}} \mathsf{Z}_i\varphi_i \in \mathcal{P}(\mathcal{T}_h)$. The approximate conserved variable $\boldsymbol{u}_h := (\mathsf{h}_h, \boldsymbol{q}_h, q_{1,h}, q_{2,h})^{\mathsf{T}}$ is represented as follows: $\boldsymbol{u}_h = \sum_{i\in\mathcal{V}} \mathbf{U}_i\varphi_i$, where $\mathbf{U}_i := (\mathsf{H}_i, \mathbf{Q}_i, \mathsf{Q}_{1,i}, \mathsf{Q}_{2,i})^{\mathsf{T}}$. Here $\mathsf{h}_h = \sum_{i\in\mathcal{V}} \mathsf{H}_i\varphi_i$ is the water height, $\boldsymbol{q}_h = \sum_{i\in\mathcal{V}} \mathbf{Q}_i\varphi_i$ is the discharge, and $q_{1,h} = \sum_{i\in\mathcal{V}} \mathsf{Q}_{1,i}\varphi_i$, $q_{2,h} = \sum_{i\in\mathcal{V}} \mathsf{Q}_{1,i}\varphi_i$ are the approximations of the two auxiliary variables $q_1 = \mathsf{h}\eta$ and $q_2 = \mathsf{h}\omega$.

The relaxation parameter ϵ is chosen to be proportional to the local mesh size. In practice we set $\mathsf{E}_i = m_i^{\frac{1}{d}}$ and $\epsilon_h = \sum_{i\in\mathcal{V}} \mathsf{E}_i\varphi_i$, where $m_i := \int_D \varphi_i\,\mathrm{d}x$ from (4.2) is proportional to the volume of the support of the shape function φ_i, and E_i is an approximation of the diameter of the support.

Just like in Kurganov and Petrova [15, Eq. (2.17)], Chertock et al. [7, Eq. (3.10)], [4,13, §5.1], we define the approximate velocity $\boldsymbol{v}_h$ and the approximate auxiliary quantities η_h and ω_h by invoking some regularization to limit the effects of roundoff errors in dry regions. After setting $\mathsf{H}_{0,\max} = \operatorname{ess\,sup}_{\boldsymbol{x}\in D} \mathsf{h}_0(\boldsymbol{x})$, where h_0 is the initial water height, and we define the following regularized quantities:

$$\mathsf{H}_i^\delta := \left(\frac{2\mathsf{H}_i}{\mathsf{H}_i^2 + \max(\mathsf{H}_i, \delta\mathsf{H}_{0,\max})^2}\right)^{-1}, \quad \mathbf{V}_i := \frac{\mathbf{Q}_i}{\mathsf{H}_i^\delta}, \quad \mathsf{N}_i := \frac{\mathsf{Q}_{1,i}}{\mathsf{H}_i^\delta}, \tag{4.3}$$

where δ is a dimensionless parameter based on machine precision. In the applications reported at the end of the paper we take $\delta = 10^{-7}$; this number is slightly larger than the square root of double precision machine epsilon. Note that with the above definition we have $\mathbf{V}_i = \frac{1}{\mathsf{H}_i}\mathbf{Q}_i$ if $\mathsf{H}_i \geq \delta\mathsf{H}_{0,\max}$; hence, the proposed regularization is active only in situations where genuine dry states occur.

4.3. The generic method and graph viscosity

Since all the strong stability preserving Runge Kutta methods are composed of convex combinations of the forward Euler method, we restrict the presentation of the scheme to the forward Euler method time stepping.

Let $\boldsymbol{u}_0$ be the initial data for (3.1)-(3.2), see (3.3). Let $\boldsymbol{u}_h^0 = \sum_{i\in\mathcal{V}} \mathbf{U}_i^0\varphi_i \in \boldsymbol{P}(\mathcal{T}_h)$ be some reasonable approximation of $\boldsymbol{u}_0$. Let τ be the time step at the current time t_n, $n \in \mathbb{N}$, and let us set $t_{n+1} := t_n + \tau$, (actually τ depends on n; we discuss

below how to estimate τ at every time step). Let $\boldsymbol{u}_h^n := \sum_{i\in\mathcal{V}} \mathbf{U}_i^n \varphi_i \in \boldsymbol{P}(\mathcal{T}_h)$ be the space approximation of $\boldsymbol{u}$ at time t_n and let us denote $\boldsymbol{u}_h^{n+1} := \sum_{i\in\mathcal{V}} \mathbf{U}_i^{n+1}\varphi_i$ the update of $\boldsymbol{u}$ at t_{n+1}. For all $i\in\mathcal{V}$ and all $j\in\mathcal{I}(i)$, we now define the following numerical flux and source:

$$\mathbf{F}_{ij}^n = \mathbf{U}_j^n(\mathbf{V}_j^n\cdot\boldsymbol{c}_{ij}) + \Big(0, \big(\widetilde{\mathsf{P}}_j(\mathbf{U}_j^n) + g\mathsf{H}_i^n(\mathsf{H}_j^n + Z_j)\big)\boldsymbol{c}_{ij}^{\mathsf{T}}, 0, 0\Big)^{\mathsf{T}}, \tag{4.4}$$

$$\mathbf{S}_i^n = \big(0, \mathbf{0}^{\mathsf{T}}, \mathsf{Q}_{2,i}^n, -\mathsf{S}_{2,i}(\mathbf{U}_i^n)\big)^{\mathsf{T}}, \tag{4.5}$$

where, for well-balancing purposes, the pressure $\widetilde{\mathsf{P}}_j$ and the source term $\mathsf{S}_{2,i}^n$ are defined as follows:

$$\widetilde{\mathsf{P}}_j(\mathbf{U}_j) = -\frac{\overline{\lambda}g}{3\mathsf{E}_j}\times\begin{cases} 6\mathsf{H}_j(\mathsf{Q}_{1,j}-\mathsf{H}_j^2), & \text{if } \mathsf{Q}_{1,j}\le\mathsf{H}_j^2\\ 2\frac{(\mathsf{Q}_{1,j}-\mathsf{H}_j^2)}{\mathsf{H}_j^\delta}(\mathsf{N}_j^2+\mathsf{Q}_{1,j}+\mathsf{H}_j^2), & \text{if } \mathsf{H}_j^2\le\mathsf{Q}_{1,j}.\end{cases} \tag{4.6}$$

$$\mathsf{S}_{2,i}(\mathbf{U}_i) = \frac{\overline{\lambda}g}{\mathsf{E}_i}\times\begin{cases} 6(\mathsf{Q}_{1,i}-\mathsf{H}_i^2), & \text{if } \mathsf{Q}_{1,i}\le\mathsf{H}_i^2\\ 6\mathsf{N}_j\frac{(\mathsf{Q}_{1,i}-\mathsf{H}_i^2)}{\mathsf{H}_i^\delta}, & \text{if } \mathsf{H}_i^2\le\mathsf{Q}_{1,i}.\end{cases} \tag{4.7}$$

The quantity $\mathbf{U}_i^{n+1}$, $i\in\mathcal{V}$, is updated as follows:

$$m_i\frac{\mathbf{U}_i^{n+1}-\mathbf{U}_i^n}{\tau} = \sum_{j\in\mathcal{I}(i)} -\mathbf{F}_{ij}^n + \sum_{j\in\mathcal{I}(i)\setminus\{i\}}\Big((d_{ij}^n-\mu_{ij}^n)(\mathbf{U}_j^{*,i,n}-\mathbf{U}_i^{*,j,n}) + \mu_{ij}^n(\mathbf{U}_j^n-\mathbf{U}_i^n)\Big) + m_i\mathbf{S}_i^n, \tag{4.8}$$

where the artificial viscosity coefficients d_{ij}^n and μ_{ij}^n are yet to be defined and are assumed to satisfy the following properties: $d_{ij}^n = d_{ji}^n$, $\mu_{ij}^n = \mu_{ji}^n$, $d_{ij}^n \ge \mu_{ij}^n \ge 0$, $i\neq j$. The states $\mathbf{U}_j^{*,i,n}$ and $\mathbf{U}_i^{*,j,n}$ are defined as follows:

$$(\mathsf{H}_i^{*,j}, \mathbf{Q}_i^{*,j}) := \frac{\mathsf{H}_i^{*,j}}{\mathsf{H}_i^\delta}(\mathsf{H}_i, \mathbf{Q}_i), \quad (\mathsf{Q}_{1,i}^{*,j}, \mathsf{Q}_{2,i}^{*,j}) := \Big(\frac{\mathsf{H}_i^{*,j}}{\mathsf{H}_i^\delta}\Big)^2(\mathsf{Q}_{1,i}, \mathsf{Q}_{2,i}), \tag{4.9a}$$

$$\mathsf{H}_i^{*,j} := \max(0, \mathsf{H}_i + \mathsf{Z}_i - \max(\mathsf{Z}_i, \mathsf{Z}_j)), \quad \forall i\in\mathcal{V},\ j\in\mathcal{I}(i). \tag{4.9b}$$

The quantity $\mathsf{H}_i^{*,j}$, introduced in Audusse et al. [3], is the so-called hydrostatic reconstruction of the water height; it makes the above method well-balanced with respect to rest states, as established in [4,13]. We refer the reader to the seminal work of Bermúdez and Vázquez [5, Def. 1] and Greenberg and Le Roux [12] where the idea of well-balanced schemes for steady states at rest has been introduced.

4.4. Viscosity

In this section we give a definition of the artificial viscosity coefficients μ_{ij}^n, d_{ij}^n that makes the method positive.

In principle, we could follow the methodology introduced in Azerad et al. [4, §3.2] or Guermond et al. [13, §4.1-§4.2] to estimate d_{ij}^n. But here, for the time being, we follow a more pragmatic approach since the Riemann problem associated with the system (3.1)-(3.2) is more complicated than that of the simple shallow water system. We set

$$\mu_{ij}^{\mathrm{L},n} := \max(|\boldsymbol{V}_i^n\cdot\boldsymbol{n}_{ij}|\|\boldsymbol{c}_{ij}\|_{\ell^2}, |\boldsymbol{V}_j^n\cdot\boldsymbol{n}_{ji}|\|\boldsymbol{c}_{ji}\|_{\ell^2}), \tag{4.10}$$

$$d_{ij}^{\mathrm{L},n} := \max(\mu_{ij}^{\mathrm{L},n}, \max(\lambda_{ij}^n\|\boldsymbol{c}_{ij}\|_{\ell^2}, \lambda_{ji}^n\|\boldsymbol{c}_{ji}\|_{\ell^2})), \tag{4.11}$$

where

$$\lambda_{ij}^n = \max(|\boldsymbol{V}_i^n\cdot\boldsymbol{n}_{ij} - (g\mathsf{H}_i^n+\theta_i^n)^{\frac12}|, |\boldsymbol{V}_j^n\cdot\boldsymbol{n}_{ij} + (g\mathsf{H}_j^n+\theta_j^n)^{\frac12}|) \tag{4.12}$$

with $\theta_i^n := \partial_{\mathsf{h}}\widetilde{\mathsf{p}}(\mathsf{H}_i^n, \mathsf{N}_i^n)\big(\frac{\mathsf{E}_i}{\max(\mathsf{E}_i, \mathsf{H}_i^n)}\big)^2$. Notice that when $\mathsf{H}_i^n \le \mathsf{E}_i$, the expressions $\boldsymbol{V}_i^n\cdot\boldsymbol{n}_{ij} - (g\mathsf{H}_i^n+\theta_i^n)^{\frac12}$ and $\boldsymbol{V}_i^n\cdot\boldsymbol{n}_{ij} + (g\mathsf{H}_i^n+\theta_i^n)^{\frac12}$ are the two extreme eigenvalues of the relaxed system (3.1)-(3.2) as established in Proposition 3.5. In the limit $\mathsf{H}_i^n \gg \mathsf{E}_i$, we recover the two extreme eigenvalues of the shallow water model. The objective of the ratio $\big(\frac{\mathsf{E}_i}{\max(\mathsf{E}_i,\mathsf{H}_i^n)}\big)^2$ is to account for dispersion effects in the viscosity in the regions that are almost dry.

When using $\mu_{ij}^{\mathrm{L},n}$ and $d_{ij}^{\mathrm{L},n}$ in lieu of μ_{ij}^n and d_{ij}^n in (4.8) one obtains a method that is first-order accurate in space only. One can make the method formally second-order accurate (at least in the L^1-norm) by reducing the viscosity as explained in Azerad et al. [4, §4.3]. One first introduces a smoothness indicator based on the water height:

$$\alpha_i^n := \frac{|\sum_{j\in\mathcal{I}(i)\setminus\{i\}}\beta_{ij}(\mathsf{H}_j^n-\mathsf{H}_i^n)|}{\sum_{j\in\mathcal{I}(S_i)\setminus\{i\}}|\beta_{ij}(\mathsf{H}_j^n-\mathsf{H}_i^n)|}, \tag{4.13}$$

where the parameters β_{ij} are positive and chosen to make the method linearity-preserving; that is, $\sum_{j\in\mathcal{I}(i)\setminus\{i\}}\beta_{ij}(\mathsf{H}_j^n-\mathsf{H}_i^n)=0$ if the restriction of $\sum_{j\in\mathcal{I}(i)}\mathsf{H}_j^n\varphi_j(\boldsymbol{x})$ over the support of φ_i is linear. For instance, one can take $\beta_{ij}=-\int_D\nabla\varphi_i\cdot\nabla\varphi_j\,\mathrm{d}x$ if the mesh satisfies an acute angle condition. One can also use the mean-value coordinates, e.g., Floater [10, Eq. 5.1]. Then one set

$$\mu_{ij}^{\mathrm{H},n}:=\max(\psi(\alpha_i^n),\psi(\alpha_j^n))\mu_{ij}^{\mathrm{L},n},\qquad d_{ij}^{\mathrm{H},n}:=\max(\psi(\alpha_i^n),\psi(\alpha_j^n))d_{ij}^{\mathrm{L},n},\tag{4.14}$$

where $\psi:[0,1]\to[0,1]$ is any Lipschitz function such that $\psi(1)=1$. The coefficients $\mu_{ij}^{\mathrm{H},n}$ and $d_{ij}^{\mathrm{H},n}$ are then used in lieu of μ_{ij}^n and d_{ij}^n in (4.8).

4.5. Well-balancing and positivity

We now establish that the algorithm (4.8) is well-balanced, as in Azerad et al. [4], beginning with a precise definition of exactly well-balanced schemes.

Definition 4.1 *(Exact rest).* A numerical state $(\mathsf{h}_h,\boldsymbol{q}_h,q_{1,h},q_{2,h})$ is said to be at exact rest (or exactly at rest) if $\boldsymbol{q}_h=\boldsymbol{0}$, $q_{2,h}=0$, $\mathsf{H}_i^2=\mathsf{Q}_{1,i}^n$, for all $i\in\mathcal{V}$, and if the approximate water height h_h and the approximate bathymetry map z_h satisfy the following alternative for all $i\in\mathcal{V}$: for all $j\in\mathcal{I}(i)$, either $\mathsf{H}_j=\mathsf{H}_i=0$ or $\mathsf{H}_j+\mathsf{Z}_j=\mathsf{H}_i+\mathsf{Z}_i$.

Definition 4.2 *(Exactly well-balanced).* A mapping $\boldsymbol{S}:\boldsymbol{P}(\mathcal{T}_h)\to\boldsymbol{P}(\mathcal{T}_h)$ is said to be an exactly well-balanced scheme if $\boldsymbol{S}(\boldsymbol{u}_h)=\boldsymbol{u}_h$ when $\boldsymbol{u}_h$ is an exact rest state.

Definition 4.3 *(Positivity).* Let us denote $\mathsf{h}_h(\boldsymbol{u}_h)=\sum_{i\in\mathcal{V}}\mathsf{H}_i(\boldsymbol{u}_h)\varphi$ the water height of $\boldsymbol{u}_h$ for any $\boldsymbol{u}_h\in\boldsymbol{P}(\mathcal{T}_h)$. A mapping $\boldsymbol{S}:\boldsymbol{P}(\mathcal{T}_h)\to\boldsymbol{P}(\mathcal{T}_h)$ is said to be an exactly well-balanced scheme if $\mathsf{H}_i(\boldsymbol{u}_h)\ge0$, for all $i\in\mathcal{V}$, implies that $\mathsf{H}_i(\boldsymbol{S}(\boldsymbol{u}_h))\ge0$ for all $i\in\mathcal{V}$.

Proposition 4.4. *Let $\boldsymbol{S}:\boldsymbol{u}_h^h\mapsto\boldsymbol{S}(\boldsymbol{u}_h^n):=\boldsymbol{u}_h^{n+1}$ be the scheme defined in (4.8).* (i) *The scheme is exactly well-balanced with either $\mu_{ij}^n:=\mu_{ij}^{\mathrm{L},n}$ or $\mu_{ij}^n:=\mu_{ij}^{\mathrm{H},n}$;* (ii) *The scheme with the viscosity (4.10) is positivity preserving provided that $\frac{\tau}{m_i}(\sum_{j\in\mathcal{I}(i)\setminus\{i\}}d_{ij}^n)\le\frac12$;* (iii) *Let k_ψ be the Lipschitz constant of ψ. With the notation $x_-=\max(0,-x)$, the scheme with the viscosity (4.14) is positivity preserving provided that $\frac{\tau}{m_i}(\sum_{j\in\mathcal{I}(i)\setminus\{i\}}d_{ij}^n+(\boldsymbol{c}_{ij}\cdot\boldsymbol{V}_j^n)_-)\le\frac12$ and $\frac{\tau}{m_i}\max_{j\in\mathcal{I}(i)\setminus\{i\}}(\boldsymbol{c}_{ij}\cdot V_j)_-\le\frac{1}{4k_\psi c_\sharp^2}$ where $c_\sharp=\max_{i\in\mathcal{V}}\mathrm{card}(\mathcal{I}(i))$.*

Proof. (i) We proceed as in the proof of Azerad et al. [4, Lem. 4.1]. Since $\boldsymbol{u}_h^n$ is exactly at rest, we have $\mathbf{U}_i^{*,j}=\mathbf{U}_j^{*,i}$ and $\mu_{ij}^n=0$ for all $i\in\mathcal{V}$ and all $j\in\mathcal{I}(i)$; hence, the viscous term in (4.8) is zero. Moreover, the source term $\boldsymbol{S}_i^n$ defined in (4.5) is zero since $q_{2,h}=0$ and, according to (4.7), $\mathsf{S}_{2,i}(\mathbf{U}_i^n)=\boldsymbol{0}$ because $\mathsf{Q}_{1,i}^n=(\mathsf{H}_i^n)^2$ for all $i\in\mathcal{V}$. Similarly, according to (4.6) $\widetilde{\mathsf{P}}_j(\mathbf{U}_j^n)=0$ for all $j\in\mathcal{I}(i)$. In conclusion we have $m_i\frac{\mathbf{U}_i^{n+1}-\mathbf{U}_i^n}{\tau}=\sum_{j\in\mathcal{I}(i)}-\Big(0,\big(g\mathsf{H}_i^n(\mathsf{H}_j^n+Z_j)\big)\boldsymbol{c}_{ij}^{\mathsf{T}},0,0\Big)^{\mathsf{T}}$. Using that $\sum_{j\in\mathcal{I}(i)}\boldsymbol{c}_{ij}=0$ we infer that $\sum_{j\in\mathcal{I}(i)}\mathsf{H}_i(\mathsf{H}_j+\mathsf{Z}_j)\boldsymbol{c}_{ij}=\sum_{j\in\mathcal{I}(i)}\mathsf{H}_i(\mathsf{H}_j+\mathsf{Z}_j-\mathsf{H}_i-\mathsf{Z}_i)\boldsymbol{c}_{ij}$. Consider $j\in\mathcal{I}(i)$, then according to Definition 4.1, either $\mathsf{H}_i=0$ and $\mathsf{H}_j=0$, or $\mathsf{H}_j+\mathsf{Z}_j-\mathsf{H}_i-\mathsf{Z}_i=0$; whence $\mathbf{U}_i^{n+1}=\mathbf{U}_i^n$, i.e., $\boldsymbol{u}_h^{n+1}=\boldsymbol{u}_h^n$. (ii) The proof of the second assertion is exactly the same as that of [4, Prop. 4.2]. (iii) The proof of the third assertion is exactly the same as that of [4, Prop. 4.4]. □

5. Numerical illustrations

We finish this note by illustrating the performance of the proposed method. We first validate the algorithm (4.8) with analytical solutions, then we make comparisons with experiments involving wetting and drying. All the numerical simulations are performed in one space dimension. The gravity constant is always taken to be $g=9.81\ \mathrm{m\,s^{-2}}$.

5.1. Implementation details

All the simulations reported in the paper are done with continuous linear finite elements in one space dimension on uniform grids, $d=1$. The time stepping is done with the third-order, three stages, strong stability preserving Runge-Kutta method, SSP RK(3,3). The time step is estimated at the beginning of each time step by using the expression

$$\tau=\mathrm{CFL}\times\max_{i\in\mathcal{V}}\frac{m_i}{\sum_{j\in\mathcal{I}(i)\setminus\{i\}}d_{ij}^{\mathrm{L},n}}.\tag{5.1}$$

In all the tests, the relaxation is done with $\overline{\lambda}=1$ and $\mathsf{E}_i=m_i^{\frac{1}{d}}$, for all $i\in\mathcal{V}$. In (4.8) we use the second-order viscosities $\mu_{ij}^{\mathrm{H},n}$ and $d_{ij}^{\mathrm{H},n}$ defined in (4.14), with the definition

Table 1
Convergence tests. Solitary wave (5.3). Final time $t = 50$ s. CFL $= 0.05$.

I	E_1	Rate	E_2	Rate	E_∞	Rate	Δ_∞	Rate	Θ
100	2.18E-01	–	2.14E-01	–	2.50E-01	–	1.44E-04	–	–
200	7.26E-02	1.59	6.49E-02	1.72	7.60E-02	1.72	1.02E-04	0.51	0.76
400	1.57E-02	2.21	1.27E-02	2.36	1.59E-02	2.25	6.01E-05	0.76	0.81
800	4.42E-03	1.83	3.29E-03	1.94	4.55E-03	1.81	3.00E-05	1.00	0.95
1600	2.02E-03	1.13	1.50E-03	1.13	1.65E-03	1.46	1.32E-05	1.19	0.90
3200	1.01E-03	1.00	7.75E-04	0.95	8.38E-04	0.98	6.20E-06	1.09	0.92
6400	4.97E-04	1.02	3.98E-04	0.96	4.31E-04	0.96	3.15E-06	0.98	0.99
12800	2.48E-04	1.00	2.03E-04	0.97	2.23E-04	0.95	1.49E-06	1.08	1.01

$$\psi(\alpha) := \left(\frac{(\alpha - \alpha_0)_+}{1 - \alpha_0} \right)^p, \tag{5.2}$$

with $p = 3$ and $\alpha_0 = \frac{1}{2}$.

5.2. A solitary wave with flat bottom topography

We consider a solution to the Green-Naghdi system (i.e., (2.1)–(2.3)) that is a solitary wave:

$$\mathsf{h}(x,t) = \mathsf{h}_1 + \frac{\mathsf{h}_2 - \mathsf{h}_1}{(\cosh(r(x - x_0 - ct))^2}, \qquad q(x,t) = c(\mathsf{h} - \mathsf{h}_1), \tag{5.3}$$

with $c = \sqrt{g\mathsf{h}_2}$ and $r = \sqrt{3(\mathsf{h}_2 - \mathsf{h}_1)/(4\mathsf{h}_2\mathsf{h}_1^2)}$. The bathymetry is flat $z = 0$.

We do the simulation in the domain $D = (0, 1000\text{ m})$ with $\mathsf{h}_2 = 11$ m and $\mathsf{h}_1 = 10$ m. The wave is located at $x_0 = 200$ m when $t = 0$. The simulations are run until $t = 50$ s; the wave travels approximately 524.4 m during this time. We estimate the accuracy of the method by computing the following quantities at the final time:

$$E_q := \frac{\|\mathsf{h}_h - \mathsf{h}\|_{L^q(D)}}{\|\mathsf{h}\|_{L^q(D)}}, \quad q \in \{1, 2, \infty\}, \qquad \Delta_\infty := \frac{\|\mathsf{h}_h - \eta_h\|_{L^\infty(D)}}{\|\mathsf{h}\|_{L^\infty(D)}}, \tag{5.4}$$

where h_h and η_h are the finite element approximations of h and η, respectively. We show in Table 1 E_1, E_2, E_∞, and Δ_∞ for various meshsizes. We observe that the convergence rate is first-order in all the norms. This is compatible with the fact that the system (3.1) is an $O(\epsilon)$-perturbation of the Green-Naghdi system and we have defined the relaxation parameter to be $\epsilon = \mathcal{O}(h)$ where h is the local meshsize. In the last column labeled "Θ" in Table 1 we show the ratio of the elapsed time between two consecutive meshes divided by 2^{1+d} (here $d = 1$). More precisely, denoting by $t_{\text{cpu}}(h)$ and $t_{\text{cpu}}(h/2)$ the elapsed time for the meshes of meshsize h and $h/2$, respectively, the ratio reported in the table is

$$\Theta := \frac{t_{\text{cpu}}(h/2)}{2^{1+d} t_{\text{cpu}}(h)}. \tag{5.5}$$

We see that Θ approaches 1 as $h \to 0$ which is the optimal ratio for hyperbolic systems. If the Green-Naghdi system were approximated with an explicit method this ratio would behave like 4 (i.e., we would have $\frac{t_{\text{cpu}}(h/2)}{t_{\text{cpu}}(h)} \sim 2^{d+3}$).

5.3. Stationary soliton wave with topography

We now consider steady state solutions to the Green-Naghdi model (2.1)–(2.3) in one space dimension:

$$\partial_x(\mathsf{h}v) = 0, \tag{5.6a}$$

$$\partial_x \left(\mathsf{h}v^2 + \frac{1}{2}g\mathsf{h}^2 + \frac{1}{3}\mathsf{h}^2\ddot{\mathsf{h}} \right) + g\mathsf{h}\partial_x z = 0, \tag{5.6b}$$

where $v = \frac{1}{\mathsf{h}}q$ is the velocity and $\dot{\mathsf{h}} = v\partial_x\mathsf{h}$, $\ddot{\mathsf{h}} = v\partial_x\dot{\mathsf{h}}$. We give a brief derivation of the Bernoulli relation for the Green-Naghdi model which will help us to find steady state solutions of (5.6). Equation (5.6a) implies that the flow discharge $q = \mathsf{h}v$ is constant. We divide (5.6b) by $g\mathsf{h}$ and replace v by $\frac{q}{\mathsf{h}}$ to obtain

$$-\frac{q^2}{g\mathsf{h}^3}\partial_x\mathsf{h} + \partial_x(\mathsf{h} + z) + \frac{q^2}{3g\mathsf{h}}\partial_x\left(\mathsf{h}^2 \frac{1}{\mathsf{h}}\partial_x(\frac{1}{\mathsf{h}}\partial_x\mathsf{h}) \right) = 0. \tag{5.7}$$

Then using the identity $\frac{1}{\mathsf{h}}\partial_x(\mathsf{h}\partial_x(\frac{1}{\mathsf{h}}\partial_x\mathsf{h})) = \partial_x(\frac{1}{\mathsf{h}}\partial_{xx}\mathsf{h} - \frac{1}{2}(\frac{1}{\mathsf{h}}\partial_x\mathsf{h})^2)$, the above equation can be rewritten as follows:

Table 2
Convergence tests. Steady-state wave with bathymetry (5.10). Final time $t = 100$ s. CFL $= 0.1$.

I	E_1	Rate	E_2	Rate	E_∞	Rate	Δ_∞	Rate	Θ
100	7.17E-03	–	1.60E-02	–	7.66E-02	–	6.71E-04	–	–
200	9.65E-04	2.89	2.45E-03	2.71	1.60E-02	2.26	2.81E-04	1.26	0.93
400	1.45E-04	2.73	3.22E-04	2.92	2.17E-03	2.88	1.30E-04	1.11	0.92
800	5.45E-05	1.42	1.14E-04	1.50	5.45E-04	1.99	6.61E-05	0.98	0.92
1600	2.40E-05	1.18	5.65E-05	1.01	4.11E-04	0.41	3.59E-05	0.88	0.92
3200	1.09E-05	1.13	2.63E-05	1.11	2.06E-04	0.99	1.91E-05	0.91	1.02
6400	7.20E-06	0.60	1.31E-05	1.00	9.23E-05	1.16	1.01E-05	0.91	0.70
12800	3.58E-06	1.01	6.14E-06	1.09	3.56E-05	1.37	5.17E-06	0.97	0.98

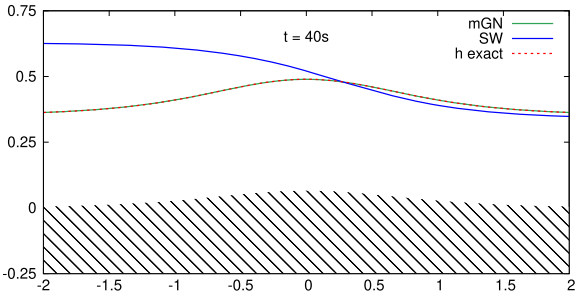

Fig. 1. Comparison of the numerical steady state profiles of the modified Green-Naghdi model (3.1)–(3.2) and shallow water model at $t = 40$ s.

$$\partial_x \left(\frac{q^2}{2gh^2} + h(x) + z(x) + \frac{q^2}{6gh^2} \left(2h\partial_{xx}h - (\partial_x h)^2 \right) \right) = 0. \tag{5.8}$$

Thus, we have the following Bernoulli relation

$$\frac{q^2}{2gh^2} + h(x) + z(x) + \frac{q^2}{6gh^2} \left(2h\partial_{xx}h - (\partial_x h)^2 \right) = C_{\text{Ber}}, \tag{5.9}$$

where the Bernoulli constant C_{Ber} is obtained by looking at the limit of the above identity when $x \to \pm\infty$ assuming that all the derivatives vanish at infinity; this gives $C_{\text{Ber}} = h_0 + \frac{q^2}{2gh_0^2}$. We now use (5.9) to define the bathymetry. For simplicity, we take the exact solution of the water height to be $h(x) = h_0 + a h_0 (\cosh(r(x - x_0)))^{-2}$, with $h_0 = \sqrt{\frac{3a}{4r^2(1+a)}}$. The parameters at our disposal are the flow rate q, the amplitude a, and the solitary wave width r. We then use (5.9) to define the bottom topography

$$z(x) = \frac{-3\sqrt{3}a^{\frac{3}{2}} g + 8q^2 r^2 \sqrt{(1+a)r^2}}{6g\sqrt{(1+a)r^2}} (\cosh(r(x - x_0)))^{-2}. \tag{5.10}$$

In our numerical experiments, we consider $D = (-10, 10\text{ m})$, and set $q = 1\text{ m}^2\,\text{s}^{-1}$, $a = 0.2$ m, $r = 1$ m, and $x_0 = 0$. We initialize the water height with the exact profile defined above and initialize the flow rate with a small perturbation of the exact rate with $q_0 = 0.95\text{ m}^2\,\text{s}^{-1}$. The flow rate is enforced at the inflow boundary. The water height is enforced at the inflow boundary and at the outflow boundary. We stop the simulations at $t = 100$ s and compute the error indicators and complexity indicators defined in (5.4)–(5.5). The results are reported in Table 2. The same observations as in §5.2 can be made. The method converges with the expected rates and the computational complexity is optimal, i.e., $\Theta \to 1$ and $h \to 0$.

In the Fig. 1, we compare the steady state profiles of both the modified Green-Naghdi model (2.1)–(2.3) and the shallow water model (i.e., (2.1)-(2.2) with $p(\boldsymbol{u}) = \frac{1}{2}gh$) at time $t = 40$ s. For the shallow water solution we enforce the flow rate at the inflow boundary and the water height at the outflow boundary. We see that the modified Green-Naghdi model gives the solitary wave profile while the shallow water model produces an expansion profile.

5.4. Experiment 1, Run-up

We now consider a test case based on the laboratory experiments reported in Synolakis [17] which demonstrates the ability of the proposed method to handle dry states. It also demonstrates the dispersive properties of the Green-Naghdi model. We study the propagation, shoaling, and run-up of a solitary wave over a plane beach with a constant slope of s. The bathymetry is defined by

$$z(x) = \max\{sx, -h_0\}. \tag{5.11}$$

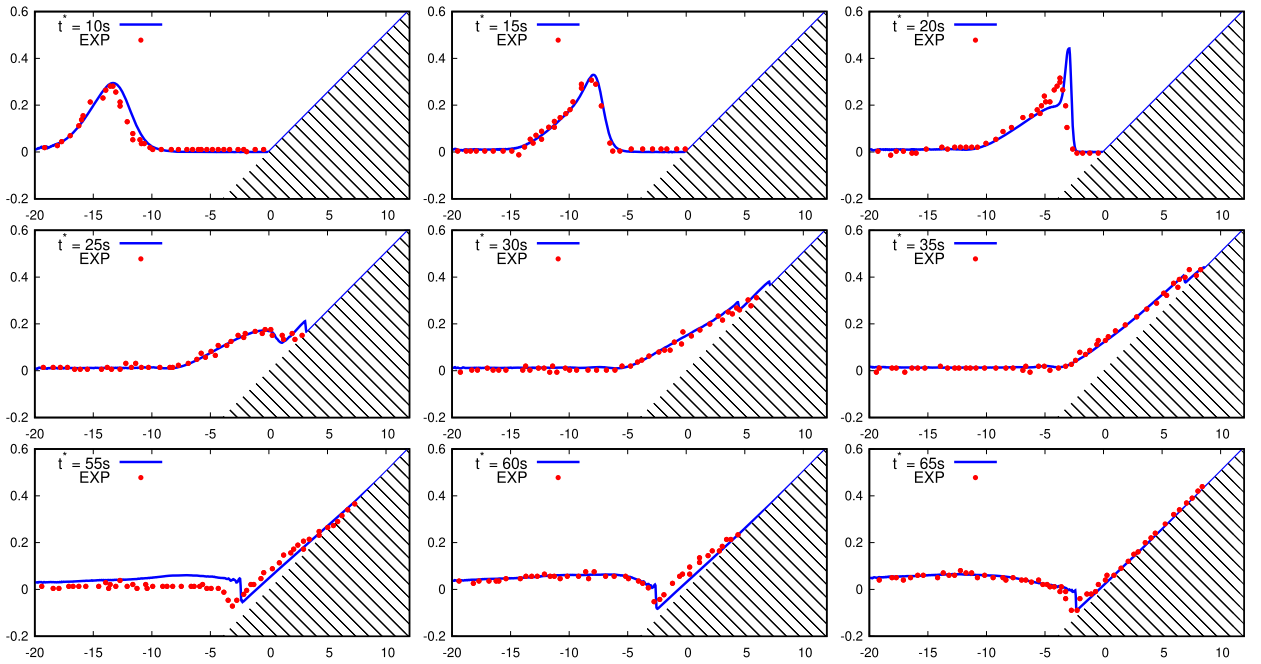

Fig. 2. Water elevation (in meters). Comparison of numerical results with experimental data (red dots) at $t^* = \{10, 15, 20, 25, 30, 35, 55, 60, 65\}$. CFL $= 0.08$. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

Let $\tilde{\mathsf{h}}(x)$ and $\tilde{u}(x)$ be the water height and velocity of an exact solitary wave:

$$\tilde{\mathsf{h}}(x) = \mathsf{h}_0 + \frac{a\mathsf{h}_0}{(\cosh(r(x - x_0)))^2}, \qquad \tilde{u}(x) = c\frac{\tilde{\mathsf{h}}(x) - \mathsf{h}_0}{\tilde{\mathsf{h}}(x)}, \tag{5.12}$$

with wave speed $c = \sqrt{g\mathsf{h}_0(1+a)}$ and width $r = \sqrt{\frac{3a\mathsf{h}_0}{4\mathsf{h}_0^3(1+a)}}$. We initialize the water height and discharge by setting

$$\mathsf{h}(x, 0) = \max\{\tilde{\mathsf{h}}(x) - \mathsf{h}_0 - z(x), 0\}, \qquad q(x, 0) = \tilde{u}(x)\mathsf{h}(x, 0). \tag{5.13}$$

The computations are done in the domain $D = (-35, 15\ \text{m})$. The slope of the beach is $s = 1/19.85$. The reference water height is $\mathsf{h}_0 = 1$ m, and we initiate the solitary wave with amplitude $a = 0.28$ m at half the wavelength from the toe of the beach; that is, we set $x_0 = -\frac{\mathsf{h}_0}{s} - \frac{L}{2}$ with $L = \frac{2}{k}\operatorname{arccosh}(0.05^{-\frac{1}{2}})$ where the wave number k is defined to be $k = \left(\frac{3a}{4\mathsf{h}_0^3}\right)^{\frac{1}{2}}$. The above data have been collected from the website for Basilisk [1].

In this experiment some energy loss is due to friction at the bottom of the channel. We account for this loss of discharge by adopting Glauckler-Manning's friction law; that is to say, we add the following source term in the momentum equation (3.1b): $-gn^2\mathsf{h}^{-\gamma}\boldsymbol{q}\|\boldsymbol{v}\|_{\ell^2}$. The parameter n is Glauckler-Manning's roughness coefficient; it has units $\text{m}^{\frac{\gamma-2}{2}}$ s. Here we take $\gamma = \frac{4}{3}$ and $n = 0.016\ \text{m}^{-\frac{1}{3}}$ s which best describes the surface condition of the smooth glass beach in the laboratory experiments. The friction term is treated explicitly as proposed in [13]: the expression (4.5) for the source term is replaced by

$$\mathbf{S}_i^n = \Big(0, \frac{2gn^2\mathbf{Q}_i^n\|\mathbf{V}_i^n\|_{\ell^2}}{(\mathsf{H}_i^n)^\gamma + \max\left((\mathsf{H}_i^n)^\gamma, 2\tau gn^2\|\mathbf{V}_i^n\|_{\ell^2}\right)}, \mathsf{Q}_{2,i}^n, -\mathsf{S}_i(\mathbf{U}_i^n)\Big)^{\mathsf{T}}. \tag{5.14}$$

We show in Fig. 2, the results of our computation done on the domain D with 1200 mesh points at various times; note that we use the non-dimensional time $t^* = t\sqrt{g/\mathsf{h}_0}$ which coincides with the experimental data. We see that our numerical results agree with the experimental data (red dots) both away from the beach and on it. (Note in passing that the experimental data at $t^* = 55$ look somewhat dubious.)

In Fig. 3 we compare the experiments at $t^* = \{10, 15, 20\}$ with the numerical approximation of the Green-Naghdi model (top row) and the numerical approximation of the shallow water model (bottom row). The shallow water approximation is obtained by using $\overline{\lambda} = 0$ in (4.8). We observe that the Green-Naghdi model captures well the propagation of the solitary wave away from the beach as opposed to the shallow water model which produces a shock. Note also that the Green-Naghdi model fits far better the experiments than the shallow water model.

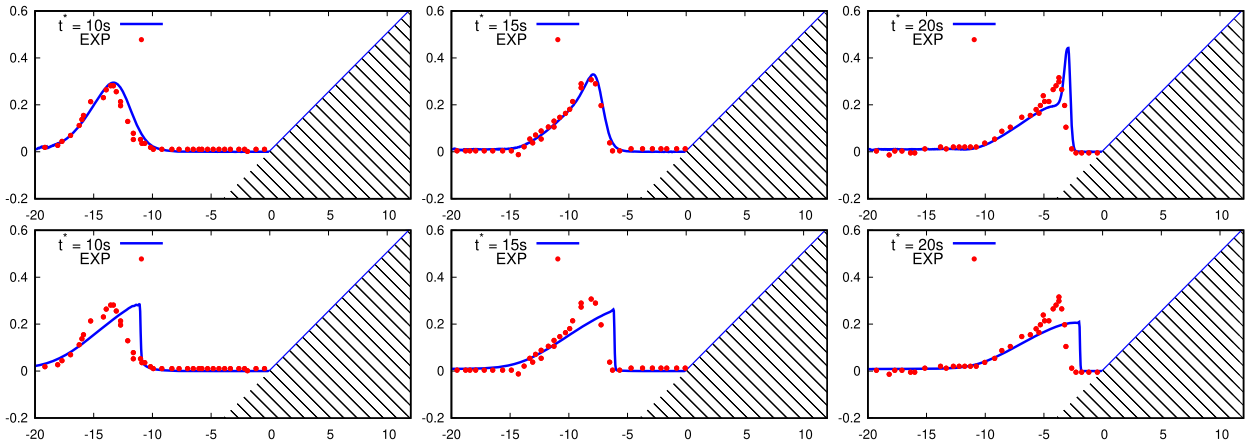

Fig. 3. Water elevation (in meters). Comparison between the (modified) Green-Naghdi model (top row) and shallow water model (bottom row) at $t^* = \{10, 15, 20\}$.

5.5. Experiment 2, Seawall

We now consider the experiments of Hsiao and Lin [14] performed at the Tainan Hydraulic Laboratory in Taiwan. In this is set of experiments a solitary wave over-tops a seawall. This test case highlights the wetting/drying process along with the dispersive characteristics of the proposed model. The numerical computations are based on the Type 1 test in [14] where the propagating solitary wave breaks on the sloping beach before reaching the wall. This test consists of a solitary wave with amplitude of 0.7 m and reference height $\mathsf{h}_0 = 0.2$ m. We initialize the water height as in the previous section, but with $a = 0.35$ m and at $x_0 = 5.9$ m. The seawall bathymetry in meters (found at [1]) is defined by

$$z(x) + \mathsf{h}_0 = \begin{cases} 0, & x < 10, \\ \frac{3.6}{20} + \frac{0.076}{13.9-13.6}(x - 13.6), & 13.6 \le x \le 13.9, \\ \frac{3.6}{20} + 0.076, & 13.9 \le x \le 13.948, \\ \frac{3.6}{20} + 0.076 - \frac{0.076-0.022}{14.045-13.948}(x - 13.948), & 13.948 \le x \le 14.045, \\ \frac{1}{20}(x - 10), & \text{otherwise.} \end{cases} \tag{5.15}$$

Note that the bathymetry must be shifted by 3 meters (to the left) to match the experimental data reported in Hsiao and Lin [14]. We account for the loss of discharge through friction at the bottom of the channel by adopting Glauckler-Manning's friction law, and we use $n = 0.012$ m$^{-\frac{1}{3}}$ s for the surface roughness.

In the experiment, several gauges were placed along the basin to capture the free surface elevation over the time period $t \in [0, 12\ s]$. With the terminology used in [14], we compare the numerical results with the measurements at 7 gauges at the following the locations: G1($x = 5.900$ m), G3($x = 9.664$ m, G10($x = 9.644$ m), G22($x = 10.462$ m), G28($x = 10.732$ m), G37($x = 11.005$ m), G14($x = 11.120$ m). Notice that the gauges on "dry land" contain data only for the water height and not the water height plus elevation. Our sketch of the setup with the positions of the gauges can be seen in the top left panel of Fig. 4. The computations are done in $D = (0, 12$ m) with a mesh composed of 1200 points. We show in the top right panel and in rows 2, 3, and 4 of Fig. 4, the temporal series over the period $t \in [0, 12\ s]$ of $\mathsf{h} + z$ at the gauges G1, G3, G10, G22, and h at the gauges G28, G37, G40 (blue line). The experimental data are reported with red dots. Our results compare well with the experiments and seem to be of a quality comparable to the numerical results reported in [14, Fig. 8],

We finally compare in Fig. 5 the numerical results obtained with the Green-Naghdi model (top row) and the numerical results obtained with the shallow water model (bottom row). The shallow water computations are done by using $\overline{\lambda} = 0$ in (4.8). The dispersive effects of the Green-Naghdi model are clearly visible. Note again that the Green-Naghdi model fits far better the experiments than the shallow water model since it accounts for the dispersive effects.

5.6. CFL behavior

We finish with a discussion on the CFL number. Using the formula (3.6) for the eigenvalues of the hyperbolic system (3.1) together with (3.12), we infer that the largest eigenvalue behaves like $\sqrt{g\mathsf{H}_{0,\max}}\sqrt{2\overline{\lambda}\frac{\mathsf{H}_{0,\max}}{h}}$, where h is the characteristic meshsize. Denoting by $\mathrm{CFL}_{\mathrm{sw}} := \tau\sqrt{g\mathsf{H}_{0,\max}}h^{-1}$ the CFL number for the shallow water equations, the above estimate means that the algorithm is stable provided $\mathrm{CFL}_{\mathrm{sw}}\sqrt{2\overline{\lambda}\frac{\mathsf{H}_{0,\max}}{h}} \lesssim 1$. That is to say, we should have stability in some reasonable sense provided $\mathrm{CFL}_{\mathrm{sw}} \lesssim \sqrt{\frac{h}{2\overline{\lambda}\mathsf{H}_{0,\max}}}$.

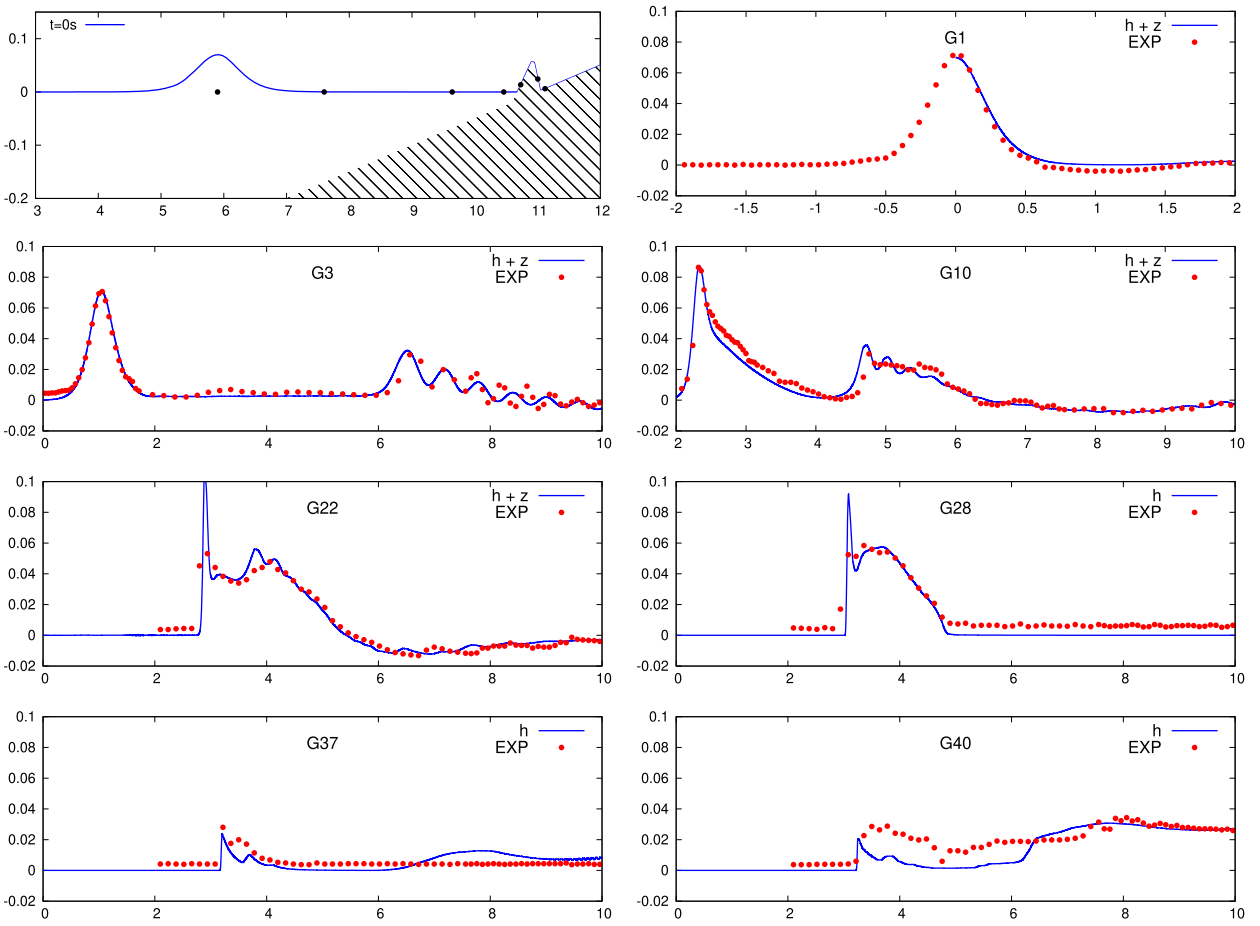

Fig. 4. (Top left) Initial setup for the seawall problem (in meters) with the location of the 7 experimental gauges; from left to right: G1, G3, G10, G22, G28, G37, G40. (Other panels) Temporal series over the period $t \in [0, 12\ \text{s}]$ of $\mathrm{h} + z$ at the gauges G1, G3, G10, G22, and h at the gauges G28, G37, G40 (blue line) compared with the experimental data (red dots). CFL $= 0.065$.

Table 3
Dependence of CFL w.r.t. the meshsize. Error in the L^1-norm on the water height.

CFL	Number of grid points							
	100	200	400	800	1600	3200	6400	12800
1.000	NaN	–	–	–	–	–	–	–
0.710	NaN	NaN	–	–	–	–	–	–
0.500	3.63E-01	NaN	NaN	–	–	–	–	–
0.350	6.03E-03	1.36E-03	3.14E-03	NaN	–	–	–	–
0.250	5.39E-03	1.21E-03	4.96E-04	7.93E-03	NaN	–	–	–
0.177	–	1.19E-03	4.95E-04	1.46E-04	3.72E-02	3.88E-01	–	–
0.125	–	–	4.89E-04	1.47E-04	8.58E-05	6.97E-02	NaN	–
0.088	–	–	–	1.47E-04	8.58E-05	1.19E-04	1.14E-01	4.58E-01
0.062	–	–	–	–	8.58E-05	3.60E-05	1.86E-04	1.58E-01
0.044	–	–	–	–	–	3.60E-05	2.41E-05	2.04E-04
0.031	–	–	–	–	–	–	2.41E-05	1.33E-05
0.022	–	–	–	–	–	–	–	1.34E-05

We illustrate the slight dependency of the CFL number (as introduced in (5.1)) with respect to the mesh size by simulating the solitary wave problem with flat bottom topography described in §5.2 over the time interval $[0, 1\ \text{s}]$ in the computational domain $D = (0, 400\ \text{m})$. We show in Table 3 the relative error on the water height in the L^1-norm at $t = 1$ s for various mesh sizes and various values of the CFL number. The number reported in the leftmost column in row $k + 3$ is equal to $\frac{1}{(\sqrt{2})^k}$, $k \in \{0{:}11\}$. The symbols NaN mean that the computation could not finish. The symbols $-$ mean that the computation has not been done. We observe that indeed the algorithm performs optimally when $\text{CFL} \sim \sqrt{\frac{h}{\mathrm{H}_{0,\max}}}$.

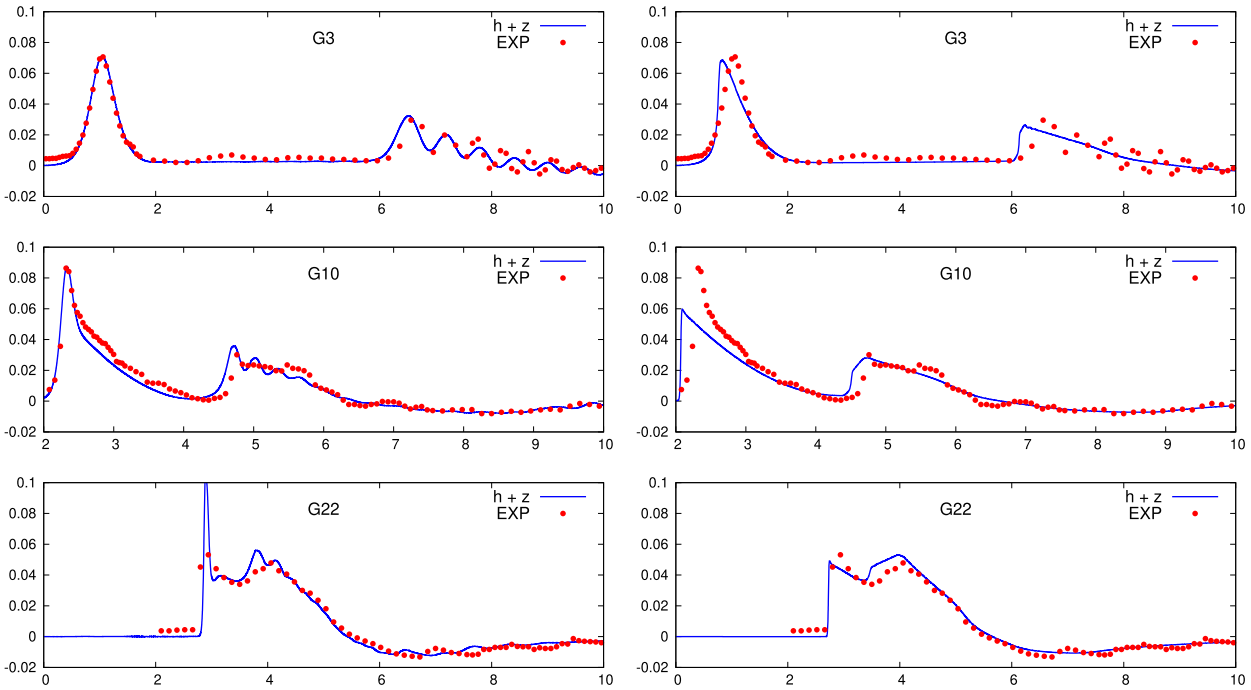

Fig. 5. Green-Naghdi (left column) vs. shallow water (right column). Temporal series over the period $t \in [0, 12\ \text{s}]$ of $\mathrm{h} + z$ at the gauges G3, G10, G22.

Table 4
L^∞-norm of the error on the water height w.r.t. meshsize; $\epsilon = 1$ m in rows 3 and 5; $\epsilon = h$ in rows 4 and 6.

CFL	ϵ	Number of grid points							
		100	200	400	800	1600	3200	6400	12800
0.2	1 m	2.82E-01	9.75E-02	1.78E-02	4.40E-03	2.59E-03	2.79E-03	2.96E-03	2.88E-03
	h	2.77E-01	9.27E-02	1.84E-02	4.84E-03	1.71E-03	6.46E-01	NaN	NaN
0.05	1 m	2.82E-01	9.84E-02	1.77E-02	4.49E-03	2.60E-03	2.76E-03	2.90E-03	2.90E-03
	h	2.78E-01	9.55E-02	1.89E-02	4.91E-03	1.70E-03	8.36E-04	4.33E-04	2.24E-04

Considering that the Green-Naghdi equations are just a $\mathcal{O}(\mu^2)$ approximation of the free surface Euler equations, as suggested in Favrie and Gavrilyuk [9], one could use the relaxation technique (3.1) as an alternative model of the free surface Euler equations by fixing the relaxation parameter ϵ. In that case, the CFL number becomes independent of the meshsize. We then propose to use (3.1) with the relaxation parameter $\epsilon = 1$ m. To illustrate, this idea, we simulate again the solitary wave problem with flat bottom topography over the time interval $[0, 50\ \text{s}]$ in the computational domain $D = (0, 1000\ \text{m})$. We show in Table 4 the relative error on the water height measured in the L^∞-norm at $t = 50$ s for various meshsizes. We compare the results obtained by the method with $\epsilon = 1$ m with the results obtained by the method described in the paper with $\epsilon = h$, where h is the meshsize. We observe that the relative error is almost the same for both methods up to 800 grid points. Beyond this resolution the relative error produced by the method with $\epsilon = 1$ m is almost constant (3×10^{-3}), whereas it decreases like $\mathcal{O}(\tau)$ for the other method. Two series of computations are done, one with CFL $= 0.2$ and the other with CFL $= 0.05$. The result for the method with $\epsilon = 1$ m are independent of the CFL number, whereas, as expected, the original method crashes on fine meshes with CFL $= 0.2$. Since 1% relative error is usually acceptable in many engineering applications, it is reasonable in practice to use the proposed method with fixed ϵ.

6. Summary

In this work we introduced a method for approximating solutions of the Green-Naghdi equations. The approach combines a new hyperbolic system of equations, which is a relaxed form of the full Green-Naghdi equations, with a new finite element method. The finite element method extends an approach for the shallow water equations presented in Guermond et al. [13], and we show that it is well-balanced, positivity preserving, and compatible with dry states for the new modified Green-Naghdi system.

The new relaxed hyperbolic model is closely related to the augmented Lagrangian relaxation of Favrie and Gavrilyuk [9], but our relaxation method is scale invariant and compatible with dry states, properties that are desirable for coastal engineering applications. Furthermore, we provide a functional form for the relaxation parameter that results in convergence of

solutions to the Green-Naghdi model as the space and time discretizations are refined. We demonstrate the expected first-order convergence rate to analytical solutions of the Green-Naghdi model, both for flat and smoothly varying bathymetry.

We also provide model validation based on experimental data for wave run-up and wave over-topping of a seawall, which demonstrates both close agreement with experiment data and distinct improvements over the basic shallow water approximation. The validation problems raise the question of whether it is necessary to require convergence of the method to the Green-Naghdi solution for practical applications or whether a fixed (mesh-independent) relaxation parameter can be defined for engineering applications.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Basilisk software. URL http://www.basilisk.fr/src/green-naghdi.h.
[2] B. Alvarez-Samaniego, D. Lannes, Large time existence for 3D water-waves and asymptotics, Invent. Math. 171 (3) (2008) 485–541.
[3] E. Audusse, F. Bouchut, M.-O. Bristeau, R. Klein, B. Perthame, A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows, SIAM J. Sci. Comput. 25 (6) (2004) 2050–2065.
[4] P. Azerad, J.-L. Guermond, B. Popov, Well-balanced second-order approximation of the shallow water equation with continuous finite elements, SIAM J. Numer. Anal. 55 (6) (2017) 3203–3224.
[5] A. Bermúdez, M.E. Vázquez, Upwind methods for hyperbolic conservation laws with source terms, Comput. Fluids 23 (8) (1994) 1049–1071.
[6] F. Bouchut, Nonlinear Stability of Finite Volume Methods for Hyperbolic Conservation Laws and Well-Balanced Schemes for Sources, Frontiers in Mathematics, Birkhäuser Verlag, Basel, 2004.
[7] A. Chertock, S. Cui, A. Kurganov, T. Wu, Well-balanced positivity preserving central-upwind scheme for the shallow water system with friction terms, Int. J. Numer. Methods Fluids 78 (6) (2015) 355–383.
[8] A. Duran, F. Marche, A discontinuous Galerkin method for a new class of Green-Naghdi equations on simplicial unstructured meshes, Appl. Math. Model. 45 (2017) 840–864.
[9] N. Favrie, S. Gavrilyuk, A rapid numerical method for solving Serre-Green-Naghdi equations describing long free surface gravity waves, Nonlinearity 30 (7) (2017) 2718–2736.
[10] M.S. Floater, Generalized barycentric coordinates and applications, Acta Numer. 24 (2015) 161–214.
[11] A.E. Green, N. Laws, P.M. Naghdi, On the theory of water waves, Proc. R. Soc. Lond. Ser. A 338 (1974) 43–55.
[12] J.M. Greenberg, A.-Y. Le Roux, A well-balanced scheme for the numerical processing of source terms in hyperbolic equations, SIAM J. Numer. Anal. 33 (1) (1996) 1–16.
[13] J.-L. Guermond, M. Quezada de Luna, B. Popov, C. Kees, M. Farthing, Well-balanced second-order finite element approximation of the shallow water equations with friction, SIAM J. Sci. Comput. 40 (6) (2018) A3873–A3901.
[14] S.-C. Hsiao, T.-C. Lin, Tsunami-like solitary waves impinging and overtopping an impermeable seawall: experiment and RANS modeling, Coast. Eng. 57 (1) (2010) 1–18.
[15] A. Kurganov, G. Petrova, A second-order well-balanced positivity preserving central-upwind scheme for the Saint-Venant system, Commun. Math. Sci. 5 (1) (2007) 133–160.
[16] C.H. Su, C.S. Gardner, Korteweg-de Vries equation and generalizations. III. Derivation of the Korteweg-de Vries equation and Burgers equation, J. Math. Phys. 10 (1969) 536–539.
[17] C.E. Synolakis, The runup of solitary waves, J. Fluid Mech. 185 (1987) 523–545.
[18] E. Toro, Shock-Capturing Methods for Free-Surface Shallow Flows, John Wiley, 2001.
[19] Y. Xing, C.-W. Shu, A survey of high order schemes for the shallow water equations, J. Math. Study 47 (3) (2014) 221–249.