



**HAL**  
open science

## Pseudo-compressibility, dispersive model and acoustic waves in shallow water flows

Anne-Sophie Bonnet-Ben Dhia, Marie-Odile Bristeau, Edwige Godlewski, Sébastien Imperiale, Anne Mangeney, Jacques Sainte-Marie

► **To cite this version:**

Anne-Sophie Bonnet-Ben Dhia, Marie-Odile Bristeau, Edwige Godlewski, Sébastien Imperiale, Anne Mangeney, et al.. Pseudo-compressibility, dispersive model and acoustic waves in shallow water flows. SEMA SIMAI Springer Series, 2021, pp.209–250. 10.1007/978-3-030-72850-2\_10 . hal-02493518v3

**HAL Id: hal-02493518**

**<https://inria.hal.science/hal-02493518v3>**

Submitted on 29 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Pseudo-compressibility, dispersive model and acoustic waves in shallow water flows

Anne-Sophie Bonnet-Ben Dhia, Marie-Odile Bristeau, Edwige Godlewski,  
Sébastien Impériale, Anne Mangeney, Jacques Sainte-Marie

**Abstract** In this paper we study a dispersive shallow water type model derived from the free surface compressible Navier-Stokes system. The compressible effects allow to capture the acoustic-like waves propagation and can be seen as a relaxation of an underlying incompressible model. The first interest of such a model is thus to capture both acoustic and water waves. The second interest lies in its numerical approximation. Indeed, at the discrete level, the pseudo-compressibility terms circumvent the resolution of an elliptic equation to capture the non-hydrostatic part of the pressure. This drastically reduces the cost of the numerical resolution of dispersive models especially in 2d and 3d.

*Keywords:* shallow water flows, dispersive models, compressible models, water and acoustic waves, projection-correction schemes, finite volumes

**Math. classification.** 65M12; 74S10; 76M12; 35L65; 35Q30; 35Q35; 76D05; 76Q05.

---

Anne-Sophie Bonnet-Ben Dhia, Inria Saclay, rue Honoré d'Estienne d'Orves F-91120 Palaiseau e-mail: Anne-Sophie.Bonnet-Bendhia@ensta-paris.fr · Marie-Odile Bristeau, Inria Paris, 2 rue Simone Iff, 75589 Paris Cedex 12 and Sorbonne University, Paris-Diderot University, CNRS, Laboratoire Jacques-Louis Lions, F-75005 Paris e-mail: Marie-Odile.Bristeau@inria.fr Edwige Godlewski, Sorbonne University, Paris-Diderot University, CNRS, Laboratoire Jacques-Louis Lions, F-75005 Paris and Inria Paris, 2 rue Simone Iff, 75589 Paris Cedex 12 e-mail: Edwige.Godlewski@upmc.fr · Sébastien Impériale, Inria Saclay rue Honoré d'Estienne d'Orves F-91120 Palaiseau and Laboratoire de mécanique des solides, Route de Saclay, F- 91120 Palaiseau e-mail: Sebastien.Imperiale@inria.fr · Anne Mangeney, Inria Paris, 2 rue Simone Iff, 75589 Paris Cedex 12 and Sorbonne University, Paris-Diderot University, CNRS, Laboratoire Jacques-Louis Lions, F-75005 Paris and Paris Diderot University, Institut de Physique du Globe de Paris, Seismology Group, 1 rue Jussieu, Paris F-75005 e-mail: Anne.Mangeney@ipgp.fr · Jacques Sainte-Marie, Inria Paris, 2 rue Simone Iff, 75589 Paris Cedex 12 and Sorbonne University, Paris-Diderot University, CNRS, Laboratoire Jacques-Louis Lions, F-75005 Paris e-mail: Jacques.Sainte-Marie@inria.fr

## 1 Presentation

The non linear shallow water model with topography [7] is widely used to describe geophysical flows and an extensive literature exists for its numerical approximation [23, 26, 10, 3, 6]. But the classical shallow water equations rely on the hydrostatic assumption and many shallow water type models taking into consideration the dispersive effects have been proposed and studied in the literature, see [24, 9, 8, 29, 28, 13, 19, 12, 2], the list being non-exhaustive.

Considering a two-dimensional domain  $\Omega \subset \mathbb{R}^2$  delimited by the boundary  $\Gamma = \Gamma_{in} \cup \Gamma_{out} \cup \Gamma_s$  as described in Fig. 1-(a), some of the authors have proposed a family of 2d shallow water dispersive models written under the form [2]

$$\frac{\partial h}{\partial t} + \frac{\partial(hu)}{\partial x} + \frac{\partial(hv)}{\partial y} = 0, \quad (1) \text{eq: eq1}$$

$$\frac{\partial(hu)}{\partial t} + \frac{\partial}{\partial x} \left( hu^2 + \frac{g}{2} h^2 + hp \right) + \frac{\partial(huv)}{\partial y} = -(gh + \frac{\gamma^2}{2} p) \frac{\partial z_b}{\partial x}, \quad (2) \text{?eq: eq2?}$$

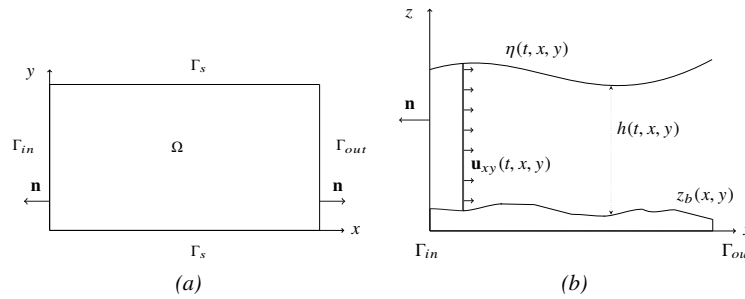
$$\frac{\partial(hv)}{\partial t} + \frac{\partial(huv)}{\partial x} + \frac{\partial}{\partial y} \left( hv^2 + \frac{g}{2} h^2 + hp \right) = -(gh + \frac{\gamma^2}{2} p) \frac{\partial z_b}{\partial y}, \quad (3) \text{?eq: eq3?}$$

$$\frac{\partial(hw)}{\partial t} + \frac{\partial(huw)}{\partial x} + \frac{\partial(hvw)}{\partial y} = \gamma p, \quad (4) \text{?eq: eq4?}$$

$$\gamma w = -h \frac{\partial u}{\partial x} + \frac{\gamma^2}{2} u \frac{\partial z_b}{\partial x} - h \frac{\partial v}{\partial y} + \frac{\gamma^2}{2} v \frac{\partial z_b}{\partial y}, \quad (5) \text{eq: eq5}$$

where  $\mathbf{u}(t, \mathbf{x}) = (u, v, w)^T$  is the velocity of the fluid with  $\mathbf{x} = (x, y)$ ,  $p$  is the non-hydrostatic part of the fluid pressure, the total pressure is given by  $p_{tot} = gh/2 + p$  and  $g$  represents the gravity acceleration. The value of the parameter  $\gamma \in \mathbb{R}$  will be discussed in Remark 1. The water depth (resp. the topography profile) is denoted  $h(t, \mathbf{x})$  (resp.  $z_b(\mathbf{x})$ ) and the free surface is defined by (see Fig. 1-(b))

$$\eta(t, \mathbf{x}) := h(t, \mathbf{x}) + z_b(\mathbf{x}). \quad (6) \text{?eq: fluid_domain?}$$



**Fig. 1** Model domain and notations, (a) view from above and (b) vertical cross section.

<pap2D\_fig:modelDomain>

For smooth solutions, the system (1)-(5) satisfies the following energy balance

$$\frac{\partial E}{\partial t} + \nabla_0 \cdot (\mathbf{u}(E + \frac{g}{2}h^2 + hp)) = 0, \quad (7) \text{?eq:eq6?}$$

with the operator  $\nabla_0 = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, 0)^T$  and

$$E = h(u^2 + v^2 + w^2)/2 + g(\eta^2 - z_b^2)/2. \quad (8) \text{?eq:EE?}$$

The system (1)-(5) defines a family  $\{\mathcal{M}_\gamma\}$  of dispersive models written in the more compact form

$$\frac{\partial h}{\partial t} + \nabla_0 \cdot (h\mathbf{u}) = 0, \quad (9) \text{todeq:nh1}$$

$$\frac{\partial(h\mathbf{u})}{\partial t} + \nabla_0 \cdot (h\mathbf{u} \otimes \mathbf{u}) + \nabla_0(\frac{g}{2}h^2) + \nabla_{sw}^\gamma p = -gh\nabla_0 z_b, \quad (10) \text{?todeq:nh2?}$$

$$\text{div}_{sw}^\gamma(\mathbf{u}) = 0, \quad (11) \text{todeq:nh3}$$

where the shallow water versions of the gradient and divergence operators are defined by

$$\nabla_{sw}^\gamma f = \begin{pmatrix} h\frac{\partial f}{\partial x} + f\frac{\partial \zeta}{\partial x} \\ h\frac{\partial f}{\partial y} + f\frac{\partial \zeta}{\partial y} \\ -\gamma f \end{pmatrix}, \quad (12) \text{todeq:gradsw}$$

$$\text{div}_{sw}^\gamma(\mathbf{w}) = \frac{\partial(hw_1)}{\partial x} + \frac{\partial(hw_2)}{\partial y} - w_1\frac{\partial \zeta}{\partial x} - w_2\frac{\partial \zeta}{\partial y} + \gamma w_3, \quad (13) \text{todeq:divsw}$$

for  $\mathbf{w} = (w_1, w_2, w_3)^T$  and

$$\zeta = h + \frac{\gamma^2}{2}z_b. \quad (14) \text{?todeq:def_zeta?}$$

Whereas  $\zeta$  depends on  $\gamma$ , for the sake of simplicity, we have adopted a simplified notation and  $\zeta_\gamma$  is replaced by  $\zeta$ .

The model studied in this paper consists in a compressible version of the model (9)-(11) where the divergence free constraint (11) is replaced by an evolution equation – a relaxed version of (11) – modeling the propagation of acoustic-type waves.

*Remark 1* The value of the parameter  $\gamma$  is discussed in [2]. Here we just recall the two extreme hydraulic regimes that can be represented by shallow water models. First the case where  $u \ll \sqrt{gh}$  i.e. the fluid velocity is very small compared to the water wave velocity or equivalently the Froude number is very low. In this situation the value  $\gamma = \sqrt{3}$  is well adapted since  $\mathcal{M}_{\sqrt{3}}$  corresponds to the well known Green-Naghdi model [24]. Another typical situation is the case of advection dominated flows –  $u$  cannot be neglected with respect to  $\sqrt{gh}$  – where the value  $\gamma = 2$  is more appropriate.

(rem:gamma)

The numerical analysis of the system (9)-(11) is studied in [2] and a numerical scheme based on a projection-correction scheme [14] has been proposed. Since the

model (9)-(11) appears as an extension of the classical Saint-Venant system, the hyperbolic part is treated using a finite volume approach – explicit in time – coupled with the resolution of a saddle point problem – implicit in time – corresponding to an elliptic-type equation for the contribution of the dispersive terms.

Because of the divergence free constraint (11) used to approximate the non-hydrostatic part of the pressure  $p$ , an implicit treatment is natural (see Sec. 3.2) but it significantly increases the computational costs. Indeed, an explicit in time scheme constrained by a CFL condition is required for the approximation of the hyperbolic part implying small time steps but simple computations of the numerical fluxes. Whereas the dispersive terms are obtained through the resolution of an elliptic equation for the whole domain. Therefore, for the numerical approximation of the model (9)-(11) over a 2d geometrical domain discretized with  $N$  cells, at each time step we have to compute  $O(N)$  numerical fluxes and to perform the resolution of a linear symmetric problem. For a stationary linear symmetric problem having at our disposal a good preconditioner, the resolution cost can be estimated as  $O(N \log N)$  computations but in our situation, the matrices depend on time – and hence have to be built at each time step – and we do not have any high-performance preconditioner. Hence the computational costs can be estimated as  $O(N^{3/2})$ , the resolution of the elliptic part becoming very limitative.

In this paper we propose, starting from the compressible Navier-Stokes equations, a modified version of (9)-(11) allowing to propagate both water and acoustic-type waves. The proposed model consists in modifying Eq. (11) in order to include compressibility effects. The new formulation has another advantage since it is possible to discretize it with a fully explicit time scheme and the computational costs are asymptotically  $O(N/\sqrt{\varepsilon})$ ,  $\sqrt{\varepsilon}$  being a parameter that will be precised later. Even if the parameter  $\varepsilon$  can be small, in 2d cases or with fine meshes we have  $\varepsilon N \gg 1$  and hence  $O(N/\sqrt{\varepsilon}) \ll O(N^{3/2})$ .

This paper is organized as follows. First starting from the 3d compressible Navier-Stokes equations, we derive a 2d shallow water model where the acoustic waves – that can be seen as pseudo-compressibility effects – are considered. Then a numerical scheme – explicit in time – is proposed for this 2d model and its properties are studied. Some stability properties – especially a discrete entropy equality – for the proposed scheme are established in the 1d context. Finally for a well known test case, an illustration comparing the implicit strategy and the resolution of the pseudo-compressible model are presented and the associated computational costs are given.

## 2 A compressible and dispersive model in shallow water context

?(sec:NSF)? In this section, we derive a shallow water approximation of the 3d compressible Navier-Stokes with free surface. The model obtained in Prop. 5 propagates both water and acoustic waves and its dispersive properties are studied. Finally, considering the

acoustic velocity is very large compared to the gravity wave velocity, we propose a new formulation as a pseudo-compressible shallow water dispersive model.

## 2.1 The compressible Navier-Stokes-Fourier system

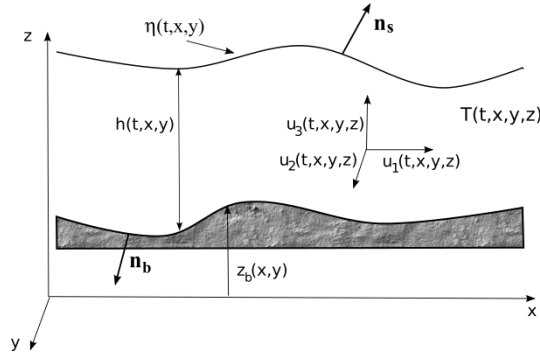
We consider the classical compressible Navier-Stokes system describing a free surface gravitational 3d flow over a bottom topography  $z_b(x, y)$  (see Fig. [Åž2](#)),

$$\frac{\partial \tilde{\rho}}{\partial t} + \nabla \cdot (\tilde{\rho} \mathbf{U}) = 0, \quad (15) \quad \text{eq:mass_cons}$$

$$\frac{\partial (\tilde{\rho} \mathbf{U})}{\partial t} + \nabla \cdot (\tilde{\rho} \mathbf{U} \otimes \mathbf{U}) + \nabla \tilde{p} - \nabla \cdot \sigma = \tilde{\rho} \mathbf{g}, \quad (16) \quad \text{eq:NSF_2d3}$$

$$\frac{\partial}{\partial t} \left( \tilde{\rho} \frac{|\mathbf{U}|^2}{2} + \tilde{\rho} \tilde{e} \right) + \nabla \cdot \left( \left( \tilde{\rho} \frac{|\mathbf{U}|^2}{2} + \tilde{\rho} \tilde{e} + \tilde{p} - \sigma \right) \mathbf{U} \right) = -\nabla \cdot Q_{\tilde{T}} + \tilde{\rho} \mathbf{g} \cdot \mathbf{U}, \quad (17) \quad \text{eq:NSF_energy}$$

where  $\mathbf{U} = (u_1, u_2, u_3)^T$  is the velocity,  $\tilde{\rho}$  is the density,  $\tilde{p}$  is the fluid pressure,  $\sigma$  is the viscosity stress and  $\mathbf{g} = (0, 0, -g)^T$  represents the gravity forces. The internal specific energy is denoted by  $\tilde{e}$ , the temperature by  $\tilde{T}$ . The symbol  $\nabla$  denotes  $\nabla = \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)^T$ . In the following, we will also use the notation  $\mathbf{v} = (u_1, u_2)^T$  for the horizontal velocity and  $\nabla_{x,y}$  corresponds to the projection of  $\nabla$  on the horizontal plane i.e.  $\nabla_{x,y} = \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right)^T$ . The square norm of the velocity vector is  $|\mathbf{U}|^2 = u_1^2 + u_2^2 + u_3^2$ .



[\(fig:notations3d\\_rho\)](#) **Fig. 2** Flow domain with water height  $h(t, \mathbf{x})$ , free surface  $\eta(t, \mathbf{x})$  and bottom  $z_b(\mathbf{x})$ .

The term  $\tilde{\rho} \mathbf{g} \cdot \mathbf{U} = -\tilde{\rho} g u_3$  in (17) prevents this equation from being directly a local energy conservation law. But multiplying the mass conservation (15) by  $z$  we get the identity

$$\frac{\partial (z \tilde{\rho})}{\partial t} + \nabla \cdot (z \tilde{\rho} \mathbf{U}) = \tilde{\rho} u_3. \quad (18) \quad \text{eq:def_w}$$

Computing the integral along the vertical axis of relation (18) and using the boundary conditions (24),(22) – that are described below – one obtains

$$\frac{\partial}{\partial t} \int_{z_b}^{\eta} g z \tilde{\rho} dz + \nabla_{x,y} \cdot \int_{z_b}^{\eta} g z \tilde{\rho} \mathbf{v} dz = \int_{z_b}^{\eta} g \tilde{\rho} u_3 dz, \quad (19) \text{eq:def\_w\_sw}$$

which is the integrated local conservation of gravitational potential energy.

Regarding constitutive equations, we assume that the fluid is Newtonian i.e. the viscous part of the Cauchy stress depends linearly on the velocity and is given by

$$\boldsymbol{\sigma} = \xi \nabla \cdot \mathbf{U} \mathbb{I} + 2\mu D(\mathbf{U}),$$

where  $\mu$  is the viscosity coefficient,  $\xi$  is the second viscosity and  $D(\mathbf{U}) = (\nabla \mathbf{U} + (\nabla \mathbf{U})^T)/2$ . The heat flux  $Q_{\tilde{T}}$  obeys the Fourier law  $Q_{\tilde{T}} = -\tilde{\lambda} \nabla \tilde{T}$ , which explains the name "Navier-Stokes-Fourier" which is often given to system (15)-(17),  $\tilde{\lambda}$  being the heat conductivity.

Among the thermodynamic variables  $\tilde{\rho}$ ,  $\tilde{p}$ ,  $\tilde{T}$ ,  $\tilde{e}$ , only two of them are independent. Indeed, we have a state law under the form

$$f(\tilde{\rho}, \tilde{T}, \tilde{p}) = 0, \quad (20) \text{eq:NSF\_state\_eq}$$

where  $f$  is a real valued function. We give some examples below. Moreover, the thermodynamic variables are linked by the Gibbs identity

$$d\tilde{e} = \frac{\tilde{p}}{\tilde{\rho}^2} d\tilde{\rho} + \tilde{T} ds, \quad (21) \text{eq:thermo\_id}$$

where  $s$  is the specific entropy of the fluid. Classically, in order to have a convenient entropy dissipation one has to assume that  $-s$  is a convex function of  $1/\tilde{\rho}$ ,  $\tilde{e}$ .

### 2.1.1 Boundary conditions at the bottom

Let  $\mathbf{n}_b$  and  $\mathbf{n}_s$  be the unit outward normals at the bottom and at the free surface respectively, defined by (see Fig 2)

$$\mathbf{n}_b = \frac{1}{\sqrt{1 + |\nabla_{x,y} z_b|^2}} \begin{pmatrix} \nabla_{x,y} z_b \\ -1 \end{pmatrix}, \quad \mathbf{n}_s = \frac{1}{\sqrt{1 + |\nabla_{x,y} \eta|^2}} \begin{pmatrix} -\nabla_{x,y} \eta \\ 1 \end{pmatrix}.$$

On the bottom we prescribe an impermeability condition

$$\mathbf{U} \cdot \mathbf{n}_b = 0, \quad (22) \text{eq:bottom}$$

and a friction condition given e.g. by a Navier law

$$((\boldsymbol{\sigma} - p\mathbb{I}) \cdot \mathbf{n}_b) \cdot \mathbf{t}_i = -\kappa \mathbf{U} \cdot \mathbf{t}_i, \quad i = 1, 2, \quad (23) \text{eq:fric}$$

with  $\kappa$  a Navier coefficient and  $(\mathbf{t}_i, i = 1, 2)$  two tangential vectors.

*Remark 2* The formulation of the two boundary conditions (22),(23) means that the fluid remains in contact with the topography. Besides, we assume throughout the paper that the total pressure remains non-negative.

### 2.1.2 Boundary conditions at free surface

On the free surface  $z = \eta(t, x, y)$ , we use the kinematic boundary condition

$$\frac{\partial \eta}{\partial t} + \mathbf{v}(t, x, y, \eta) \cdot \nabla_{x,y} \eta - u_3(t, x, y, \eta) = 0, \quad (24) \text{eq:free_surf}$$

and the no stress condition

$$(\sigma - p\mathbb{I}) \cdot \mathbf{n}_s = -p^a(t, x, y)\mathbf{n}_s + W(t, x, y)\mathbf{t}_s, \quad (25) \text{eq:bound3ns}$$

where  $p^a(t, x, y)$ ,  $W(t, x, y)$  are two given external forcings,  $p^a$  (resp.  $W$ ) mimics the effects of the atmospheric pressure (resp. the wind blowing at the free surface) and  $\mathbf{t}_s$  is a given unit horizontal vector. Throughout the paper we assume  $p^a = cst$ ,  $W = 0$ .

### 2.1.3 Boundary conditions for the temperature

The heat flux in Eq. (17) requires to define boundary conditions for the temperature. Moreover when the state law (20) will be precised, the definition of the temperature at each boundary will be mandatory. We can choose either Neumann or Dirichlet conditions namely at the bottom

$$\lambda \frac{\partial \tilde{T}}{\partial \mathbf{n}_b} = F\tilde{T}_b^0, \quad (26) \text{?BC:neumann_bottom?}$$

or

$$\tilde{T}_b = \tilde{T}_b^0, \quad (27) \text{?BC:dirichlet_bottom?}$$

and at the free surface

$$\lambda \frac{\partial \tilde{T}}{\partial \mathbf{n}_s} = F\tilde{T}_s^0, \quad (28) \text{?BC:neumann_surface?}$$

or

$$\tilde{T}_s = \tilde{T}_s^0, \quad (29) \text{?BC:dirichlet_surface?}$$

where  $F\tilde{T}_b^0$ ,  $F\tilde{T}_s^0$  are two given temperature fluxes and  $\tilde{T}_b^0$ ,  $\tilde{T}_s^0$  are two given temperatures.



## 2.2 Thermodynamic considerations

In the following proposition, we propose a formulation of the compressible Euler system – corresponding to the system (15)-(17) with  $\lambda = \xi = \mu = 0$  – where the acoustic speed explicitly appears. The system is deduced from the compressible Navier-Stokes system (15)-(17) with the boundary conditions (22)-(25).

**Proposition 1** *Considering a state law under the form*

$$\tilde{p} = f(\tilde{\rho}, \tilde{T}), \quad (30) \quad \text{eq:state\_law\_p}$$

*the compressible Euler system can be rewritten under the form*

$$\frac{\partial \tilde{\rho}}{\partial t} + \nabla \cdot (\tilde{\rho} \mathbf{U}) = 0, \quad (31) \quad \text{eq:NSF\_ac1}$$

$$\frac{\partial (\tilde{\rho} \mathbf{U})}{\partial t} + \nabla \cdot (\tilde{\rho} \mathbf{U} \otimes \mathbf{U}) + \nabla \tilde{p} = \tilde{\rho} \mathbf{g}, \quad (32) \quad \text{eq:NSF\_ac2}$$

$$\frac{\partial (\tilde{\rho} \tilde{p})}{\partial t} + \nabla \cdot (\tilde{\rho} \mathbf{U} \tilde{p}) + \tilde{\rho}^2 \tilde{c}^2 \nabla \cdot \mathbf{U} = 0, \quad (33) \quad \text{eq:NSF\_ac3}$$

where the sound speed  $\tilde{c}$  is defined below by (47).

The system (31)-(33) is completed with the boundary conditions (22),(24) and (25) which becomes

$$\tilde{p}(t, x, y, \eta) = p^a(t, x, y) = cst. \quad (34) \quad \text{eq:ps}$$

Smooth solutions of the system (31)-(33) satisfy the energy balance

$$\frac{\partial}{\partial t} \left( \tilde{\rho} \frac{|\mathbf{U}|^2}{2} + \tilde{\rho} \tilde{e} \right) + \nabla \cdot \left( \left( \tilde{\rho} \frac{|\mathbf{U}|^2}{2} + \tilde{\rho} \tilde{e} + \tilde{p} \right) \mathbf{U} \right) = \tilde{\rho} \mathbf{g} \cdot \mathbf{U}, \quad (35) \quad \text{eq:NSF\_ac4}$$

where the internal energy  $\tilde{e}$  satisfies the equation

$$\frac{\partial (\tilde{\rho} \tilde{e})}{\partial t} + \nabla \cdot (\tilde{\rho} \tilde{e} \mathbf{U}) = -\tilde{p} \nabla \cdot \mathbf{U}. \quad (36) \quad \text{eq:NSF\_internal\_energy\_3}$$

Notice that in this proposition we have kept the same notations even if we have switched from the Navier-Stokes to the Euler system.

(prop:NS\_acc)

For the Euler system (31)-(33) – and also for the Navier-Stokes system – a crucial point is the duality relation between the gradient and divergence operators which writes

$$\int_{\Omega \times [z_b, \eta]} \tilde{p} \nabla \cdot \mathbf{V} dx dz = \int_{\partial(\Omega \times [z_b, \eta])} \tilde{p} \mathbf{V} \cdot \mathbf{n} ds - \int_{\Omega \times [z_b, \eta]} \mathbf{V} \cdot \nabla \tilde{p} dx dz. \quad (37) \quad \text{eq:duality\_euler}$$

It will be important to have, in the shallow water context, a relation analogous to (37), see (95) below.

**Proof (Proof of Prop. 1)** The main point of this proof is the derivation of Eq. (33).

Taking the scalar product of Eq. (16) by  $\mathbf{U}$  yields the kinetic energy equation

$$\frac{\partial}{\partial t} \left( \tilde{\rho} \frac{|\mathbf{U}|^2}{2} \right) + \nabla \cdot \left( \left( \tilde{\rho} \frac{|\mathbf{U}|^2}{2} + \tilde{p} - \sigma \right) \mathbf{U} \right) = \tilde{p} \nabla \cdot \mathbf{U} - \sigma : D(\mathbf{U}) + \tilde{\rho} \mathbf{g} \cdot \mathbf{U}. \quad (38) \quad \text{eq:NSF_kin_energy}$$

Subtracting (38) to (17) gives the equation for the internal energy

$$\frac{\partial(\tilde{\rho}\tilde{e})}{\partial t} + \nabla \cdot (\tilde{\rho}\tilde{e}\mathbf{U}) = -\tilde{p}\nabla \cdot \mathbf{U} + \sigma : D(\mathbf{U}) - \nabla \cdot Q_{\tilde{T}}, \quad (39) \quad \text{?eq:NSF_internal_energy-1?}$$

or equivalently

$$\tilde{\rho} \frac{D\tilde{e}}{Dt} = -\tilde{p}\nabla \cdot \mathbf{U} + \sigma : D(\mathbf{U}) - \nabla \cdot Q_{\tilde{T}}, \quad (40) \quad \text{eq:NSF_internal_energy_2}$$

with the classical notation  $D/Dt \equiv \partial/\partial t + \mathbf{U} \cdot \nabla$ . We can write the continuity equation (15) as

$$\tilde{\rho} \frac{D\tilde{\rho}}{Dt} + \tilde{\rho}^2 \nabla \cdot \mathbf{U} = 0. \quad (41) \quad \text{eq:mass_cons_divu}$$

With the thermodynamic relation (21) one can write  $ds = d\tilde{e}/\tilde{T} - (\tilde{p}/\tilde{T} \tilde{\rho}^2) d\tilde{\rho}$ , thus multiplying (40) by  $1/\tilde{T}$  and (41) by  $-\tilde{p}/\tilde{T} \tilde{\rho}^2$  we obtain

$$\tilde{\rho} \frac{Ds}{Dt} = \frac{1}{\tilde{T}} \sigma : D(\mathbf{U}) - \frac{1}{\tilde{T}} \nabla \cdot Q_{\tilde{T}}. \quad (42) \quad \text{eq:NSF_specific-entropy-1}$$

This can be written also

$$\frac{\partial(\tilde{\rho}s)}{\partial t} + \nabla \cdot (\tilde{\rho}s\mathbf{U}) = \frac{1}{\tilde{T}} \sigma : D(\mathbf{U}) - \nabla \cdot \frac{Q_{\tilde{T}}}{\tilde{T}} - Q_{\tilde{T}} \cdot \frac{\nabla \tilde{T}}{\tilde{T}^2}, \quad (43) \quad \text{?eq:NSF_specific-entropy-2?}$$

which gives the increase with time of  $\int \tilde{\rho}s$ , the second principle of thermodynamics.

The state law (20) plays the role of a closure relation. When written under the form (30), it allows to write

$$d\tilde{p} = \left( \frac{\partial \tilde{p}}{\partial \tilde{\rho}} \right)_{\tilde{T}} d\tilde{\rho} + \left( \frac{\partial \tilde{p}}{\partial \tilde{T}} \right)_{\tilde{\rho}} d\tilde{T}. \quad (44) \quad \text{eq:dp}$$

Therefore, from Eqs. (44),(42) and using Eqs. (41) we get

$$\begin{aligned} \tilde{\rho} \frac{D\tilde{p}}{Dt} + \tilde{\rho}^2 \left( \left( \frac{\partial \tilde{p}}{\partial \tilde{\rho}} \right)_{\tilde{T}} + \left( \frac{\partial \tilde{p}}{\partial \tilde{T}} \right)_{\tilde{\rho}} \left( \frac{\partial \tilde{T}}{\partial \tilde{\rho}} \right)_s \right) \nabla \cdot \mathbf{U} = \\ \frac{1}{\tilde{T}} \left( \frac{\partial \tilde{p}}{\partial \tilde{T}} \right)_{\tilde{\rho}} \left( \frac{\partial \tilde{T}}{\partial s} \right)_{\tilde{\rho}} (\sigma : D(\mathbf{U}) - \nabla \cdot Q_{\tilde{T}}). \end{aligned} \quad (45) \quad \text{?eq:NSF_pressure_1?}$$

Using the chain rule this can be written

$$\tilde{\rho} \frac{D\tilde{p}}{Dt} + \tilde{\rho}^2 \tilde{c}^2 \nabla \cdot \mathbf{U} = \frac{1}{\tilde{T}} \left( \frac{\partial \tilde{p}}{\partial s} \right)_{\tilde{\rho}} (\sigma : D(\mathbf{U}) - \nabla \cdot Q_{\tilde{T}}), \quad (46) \quad \text{eq:NSF_pressure_2}$$

with the sound speed  $\tilde{c}$  given by

$$\tilde{c}^2 = \left( \frac{\partial \tilde{p}}{\partial \tilde{\rho}} \right)_{\tilde{T}} + \left( \frac{\partial \tilde{p}}{\partial \tilde{T}} \right)_{\tilde{\rho}} \left( \frac{\partial \tilde{T}}{\partial \tilde{\rho}} \right)_s = \left( \frac{\partial \tilde{p}}{\partial \tilde{\rho}} \right)_s. \quad (47) \text{eq:sound_velocity}$$

And coupled with (15), Eq. (46) writes

$$\frac{\partial(\tilde{\rho}\tilde{p})}{\partial t} + \nabla \cdot (\tilde{\rho}\mathbf{U}\tilde{p}) + \tilde{\rho}^2 \tilde{c}^2 \nabla \cdot \mathbf{U} = \frac{1}{\tilde{T}} \left( \frac{\partial \tilde{p}}{\partial s} \right)_{\tilde{\rho}} (\sigma : D(\mathbf{U}) - \nabla \cdot \mathbf{Q}_{\tilde{T}}), \quad (48) \text{eq:NSF_pressure_3}$$

Equations (15), (16) and (48) with  $\lambda = \xi = \mu = 0$  give Eqs. (31)-(33).

Using (36), taking the scalar product of (32) with  $\mathbf{U}$  and after simple computations, we obtain the energy balance (35).  $\square$

*Remark 3* Whereas, Eq. (31) expresses the local mass conservation, the volume variations can be related to the temperature variations. Indeed, since

$$d\tilde{T} = \left( \frac{\partial \tilde{T}}{\partial \tilde{\rho}} \right)_s d\tilde{\rho} + \left( \frac{\partial \tilde{T}}{\partial s} \right)_{\tilde{\rho}} ds,$$

using relations (42) and (41) we get the following equation governing the temperature

$$\tilde{\rho} \frac{D\tilde{T}}{Dt} + \tilde{\rho}^2 \left( \frac{\partial \tilde{T}}{\partial \tilde{\rho}} \right)_s \nabla \cdot \mathbf{U} = \frac{1}{\tilde{T}} \left( \frac{\partial \tilde{T}}{\partial s} \right)_{\tilde{\rho}} (\sigma : D(\mathbf{U}) - \nabla \cdot \mathbf{Q}_{\tilde{T}}). \quad (49) \text{eq:NSF_T}$$

### 2.3 Acoustic waves and water waves

The system (31)-(33) – completed with the boundary conditions (22),(24) and (34) – is a compressible model with a free surface and hence acoustic and water waves can propagate.

Let us define  $\hat{p}$  by

$$\tilde{p} = p^a + \int_z^\eta \tilde{\rho} g dz + \hat{p},$$

with  $p^a = cst$  thus  $\hat{p}$  denotes the non-gravitational part of the pressure. Then the system (31)-(33) with (22) and (24) also writes

$$\frac{\partial \tilde{\rho}}{\partial t} + \nabla \cdot (\tilde{\rho}\mathbf{U}) = 0, \quad (50) \text{eq:NSF_vp1}$$

$$\frac{\partial \mathbf{U}}{\partial t} + (\mathbf{U} \cdot \nabla)\mathbf{U} + \frac{1}{\tilde{\rho}} \nabla \hat{p} + \frac{1}{\tilde{\rho}} \nabla \int_z^\eta \tilde{\rho} g dz_1 = \mathbf{g}, \quad (51) \text{?eq:NSF_vp2?}$$

$$\frac{\partial}{\partial t} \left( \int_z^\eta \tilde{\rho} g dz + \hat{p} \right) + \mathbf{U} \cdot \nabla \left( \int_z^\eta \tilde{\rho} g dz + \hat{p} \right) + \tilde{\rho} \tilde{c}^2 \nabla \cdot \mathbf{U} = 0, \quad (52) \text{?eq:NSF_vp3?}$$

$$\frac{\partial h}{\partial t} + \mathbf{v}_s \cdot \nabla_{x,y} h = u_{3,s}, \quad (53) \text{eq:NSF_vp4}$$

where the subscript corresponds to the value of the free surface  $z = \eta$ .

Assuming a flat bottom and in a two dimensional setting  $(x, z)$ , the system (50)-(53) has the following compact formulation

$$M \frac{\partial Y}{\partial t} + A_x \frac{\partial Y}{\partial x} + A_z \frac{\partial Y}{\partial z} = S, \quad (54) \quad \boxed{\text{eq:sys\_quasi\_lin}}$$

with

$$Y = \begin{pmatrix} \tilde{\rho} \\ u_1 \\ u_3 \\ \hat{p} \\ h \end{pmatrix}, \quad A_x = \begin{pmatrix} u_1 & \tilde{\rho} & 0 & 0 & 0 \\ \frac{g(h-z)}{\tilde{\rho}} & u_1 & 0 & \frac{1}{\tilde{\rho}} & g \\ 0 & 0 & u_1 & 0 & 0 \\ g(h-z)u_1 & \tilde{\rho}\tilde{c}^2 & 0 & u_1 & g\tilde{\rho}u_1 \\ 0 & 0 & 0 & 0 & u_{1,s} \end{pmatrix},$$

$$A_z = \begin{pmatrix} u_3 & 0 & \tilde{\rho} & 0 & 0 \\ 0 & u_3 & 0 & 0 & 0 \\ 0 & 0 & u_3 & \frac{1}{\tilde{\rho}} & 0 \\ 0 & 0 & \tilde{\rho}\tilde{c}^2 & u_3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad M = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ g(h-z) & 0 & 0 & 1 & \tilde{\rho}g \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

and

$$S = \begin{pmatrix} 0 \\ \frac{g}{\tilde{\rho}} \frac{\partial}{\partial x} \int_z^h (\tilde{\rho}(t, x, z) - \tilde{\rho}(t, x, z_1)) dz_1 \\ 0 \\ g \frac{\partial}{\partial t} \int_z^h (\tilde{\rho}(t, x, z) - \tilde{\rho}(t, x, z_1)) dz_1 - g u_1 \frac{\partial}{\partial x} \int_z^h (\tilde{\rho}(t, x, z) - \tilde{\rho}(t, x, z_1)) dz_1 + g \tilde{\rho} u_3 \\ u_{3,s} \end{pmatrix}.$$

Considering we are in a shallow water context, we can further assume

$$\frac{\partial \tilde{\rho}}{\partial z} = \frac{\partial u_1}{\partial z} = 0, \quad (55) \quad \text{?eq:sw\_regime?}$$

then the system (54) reduces to

$$M^{sw} \frac{\partial Y}{\partial t} + A_x^{sw} \frac{\partial Y}{\partial x} + A_z \frac{\partial Y}{\partial z} = B, \quad (56) \quad \text{?eq:sys\_quasi\_lin1?}$$

with

$$A_x^{sw} = \begin{pmatrix} u_1 & \tilde{\rho} & 0 & 0 & 0 \\ \frac{g(h-z)}{\tilde{\rho}} & u_1 & 0 & \frac{1}{\tilde{\rho}} & g \\ 0 & 0 & u_1 & 0 & 0 \\ gh u_1 & \tilde{\rho}(\tilde{c}^2 + gz) & 0 & u_1 & g\tilde{\rho}u_1 \\ 0 & 0 & 0 & 0 & u_1 \end{pmatrix}, \quad M^{sw} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ gh & 0 & 0 & 1 & \tilde{\rho}g \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

The eigenvalues of the matrix  $aA_x^{sw} + bA_z^{sw}$  for  $(a, b) \in \mathbb{R}^2$  cannot be easily computed explicitly but the following result holds.

**Proposition 2** *The eigenvalues of the matrix  $(M^{sw})^{-1}(aA_x^{sw} + bA_z^{sw})$  with  $u_1 = u_3 = 0$ , are given by*

$$0, \pm \frac{1}{2} \sqrt{2C_1 \pm 2\sqrt{C_2 + C_3}}, \quad (57) \text{eq:eig}$$

with

$$\begin{aligned} C_1 &= \tilde{c}^2(a^2 + b^2) - b^2gh, & C_2 &= (a^2 + b^2)^2\tilde{c}^4 + b^2g^2(8a^2hz - 4a^2z^2 + b^2h^2), \\ C_3 &= -2b^2(3a^2 + b^2)\tilde{c}^2gh. \end{aligned}$$

(prop:eig\_2d)

**Proof (Proof of Prop. 2)** The proof relies on simple computations that are not presented here.  $\square$

*Remark 4* Notice that the quantities  $C_2$  is non-negative whereas  $C_3$  is non-positive. In the situation where  $\tilde{c}^2 \geq gh$  that is encountered in practice, then  $C_1 \geq 0$ ,  $C_2 + C_3 \geq 0$  and  $C_1 - \sqrt{C_2 + C_3} \geq 0$  therefore the system has 4 real eigenvalues. But when  $\tilde{c}^2 \leq gh$  – corresponding to a less realistic situation – then complex eigenvalues could appear.

Notice also that we are considering situations where  $\tilde{c} \gg 1$  (see Eq. (68) below), hence the eigenvalues defined by (57) satisfy the estimates

$$0, \pm \tilde{c}\sqrt{a^2 + b^2} + O\left(\frac{1}{\tilde{c}}\right), \pm \sqrt{gh} \frac{ab}{\sqrt{a^2 + b^2}} + O\left(\frac{1}{\tilde{c}^2}\right), \quad (58) \text{eq:estim_vp}$$

where the second ones correspond to acoustic waves and the third one to surface waves. The first eigenvalue is zero because of the linearization of the velocity field.

## 2.4 Sound speed for sea water

(subsec:sound\_speed)

In this paragraph, we are going to precise the expression of the sound speed  $\tilde{c}$  defined by (47) in the particular case of sea water.

We start from an expression of the state law (30) given by [30] under the form

$$\tilde{\rho} = \tilde{\rho}(\tilde{p}, \tilde{T}) = \frac{\tilde{\rho}_0(\tilde{T})}{1 - \frac{\varepsilon}{\rho_0}\tilde{p}} = \frac{\rho_0\tilde{\rho}_0(\tilde{T})}{\rho_0 - \varepsilon\tilde{p}}, \quad (59) \text{eq:state\_law}$$

$\rho_0$  and  $\varepsilon$  being two constants with  $\varepsilon \ll 1$  and where

$$\tilde{\rho}_0(\tilde{T}) = \rho_0 + a(\tilde{T} - T_0)^2, \quad (60) \text{eq:rho\_def}$$

with  $T_0 = 4^\circ\text{C}$ ,  $\rho_0 = 9999.7 \text{ kg}\cdot\text{m}^{-3}$ ,  $a = -6.63 \cdot 10^{-3} \text{ kg}\cdot\text{m}^{-3}\cdot\text{K}^{-2}$ . Notice that from the state law (59) we have

$$\varepsilon \tilde{p} = o(1). \quad (61) \text{eq:scaling_p}$$

When multiplied by  $\varepsilon$ , the relation (21) is compatible with the scaling (61) if

$$\varepsilon \tilde{e} = o(1), \quad \text{and} \quad \varepsilon s = o(1). \quad (62) \text{eq:scaling_es}$$

Rewriting (21) under the form

$$d(\tilde{e} - \tilde{T}s) = -s d\tilde{T} + \frac{\tilde{p}}{\tilde{\rho}^2} d\tilde{\rho}, \quad (63) \text{eq:thermo_idbis}$$

the Schwarz theorem applied to (63) gives the equality

$$-\left(\frac{\partial s}{\partial \tilde{\rho}}\right)_{\tilde{T}} = \left(\frac{\partial(\tilde{p}/\tilde{\rho}^2)}{\partial \tilde{T}}\right)_{\tilde{\rho}},$$

and using the expression of  $\tilde{p}$  given by (59) we obtain

$$\left(\frac{\partial s}{\partial \tilde{\rho}}\right)_{\tilde{T}} = \frac{\rho_0 \tilde{\rho}'_0(\tilde{T})}{\varepsilon \tilde{\rho}^3}.$$

An integration of the previous relation gives

$$s = s_0(\tilde{T}) - \frac{\rho_0 \tilde{\rho}'_0(\tilde{T})}{2\varepsilon \tilde{\rho}^2}, \quad (64) \text{eq:s}$$

and from (64) we obtain

$$ds = \left( s'_0(\tilde{T}) - \frac{\rho_0 \tilde{\rho}''_0(\tilde{T})}{2\varepsilon \tilde{\rho}^2} \right) d\tilde{T} + \frac{\rho_0 \tilde{\rho}'_0(\tilde{T})}{\varepsilon \tilde{\rho}^3} d\tilde{\rho},$$

leading to

$$\left(\frac{\partial \tilde{T}}{\partial \tilde{\rho}}\right)_s = -\frac{2}{\tilde{\rho}} \frac{\rho_0 \tilde{\rho}'_0(\tilde{T})}{2\varepsilon \tilde{\rho}^2 s'_0(\tilde{T}) - \rho_0 \tilde{\rho}''_0(\tilde{T})}. \quad (65) \text{eq:ds}$$

Using (59),(65) we obtain the expression for the sound speed  $\tilde{c}$  defined by (47) under the form

$$\tilde{c}^2 = \frac{\rho_0 \tilde{\rho}_0(\tilde{T})}{\varepsilon \tilde{\rho}^2} + \frac{\rho_0 \tilde{\rho}'_0(\tilde{T})}{\varepsilon \tilde{\rho}} \frac{2}{\tilde{\rho}} \frac{\rho_0 \tilde{\rho}'_0(\tilde{T})}{2\varepsilon \tilde{\rho}^2 s'_0(\tilde{T}) - \rho_0 \tilde{\rho}''_0(\tilde{T})}. \quad (66) \text{?eq:sound_expr?}$$

Because of the estimate (62) concerning the entropy, a possible choice is  $s'_0(\tilde{T}) = 0$  leading to the expression

$$\tilde{\rho}^2 \tilde{c}^2 = \frac{\rho_0}{\varepsilon} \left( \tilde{\rho}_0(\tilde{T}) - 2 \frac{(\tilde{\rho}'_0(\tilde{T}))^2}{\tilde{\rho}''_0(\tilde{T})} \right). \quad (67) \text{?eq:sound_res?}$$

Moreover, since  $\varepsilon \ll 1$  and  $a \ll 1$ , we have  $\tilde{\rho}(\tilde{T}) \approx \rho_0$  leading to an expression for the sound speed under the form

$$\tilde{\rho}^2 \tilde{c}^2 \approx \frac{\rho_0^2}{\varepsilon} \left( 1 - 4a \frac{(\tilde{T} - T_0)^2}{\rho_0} \right) \approx \frac{\rho_0^2}{\varepsilon} = \rho_0^2 c^2. \quad (68) \text{eq:sound_res1}$$

Using the assumption (68), it is possible to combine Eq. (33) with Eq. (31) to obtain

$$\frac{\partial}{\partial t} \left( \tilde{\rho} \frac{\tilde{p}^2}{2\rho_0^2 c^2} \right) + \nabla \cdot \left( \tilde{\rho} \frac{\tilde{p}^2}{2\rho_0^2 c^2} \mathbf{U} \right) + \tilde{p} \nabla \cdot \mathbf{U} = 0, \quad (69) \text{?eq:NSF_ac3bis?}$$

giving a formulation of the energy balance (35) under the form

$$\frac{\partial}{\partial t} \left( \tilde{\rho} \frac{|\mathbf{U}|^2}{2} + \tilde{\rho} \frac{\tilde{p}^2}{2\rho_0^2 c^2} \right) + \nabla \cdot \left( \left( \tilde{\rho} \frac{|\mathbf{U}|^2}{2} + \tilde{\rho} \frac{\tilde{p}^2}{2\rho_0^2 c^2} + \tilde{p} \right) \mathbf{U} \right) = \tilde{\rho} \mathbf{g} \cdot \mathbf{U}. \quad (70) \text{?eq:NSF_ac5?}$$

Hence, when  $\tilde{\rho}^2 \tilde{c}^2 = \rho_0^2 c^2 = cst$ , the internal energy corresponds to  $\frac{\tilde{p}^2}{2\rho_0^2 c^2}$ .

## 2.5 A shallow water approximation of the compressible Euler system

For free surface flows, the vertical direction plays a particular role since it corresponds to the direction of the gravity. Moreover the fluid domain, in our case, is thin in this direction and it is natural to perform a depth averaging of system (31)-(33) together with some approximations. Then, the two following propositions hold.

*Remark 5* Notice that whereas the models described in the sequel differ from the model (9)-(11), for the sake of simplicity, we have kept the same notations for some of the variables of each model.

**Proposition 3** Assuming  $\tilde{\rho} \tilde{c}$  is constant, see (68), a shallow water approximation of the compressible Euler system (31)-(33) completed with (22), (24) leads to the model

$$\frac{\partial \eta}{\partial t} + \mathbf{u} \cdot \nabla_0 (\eta + z_b) = \gamma w, \quad (71) \text{eq:mod_sw0}$$

$$\frac{\partial (\rho h)}{\partial t} + \nabla_0 \cdot (\rho h \mathbf{u}) = 0, \quad (72) \text{eq:mod_sw1}$$

$$\frac{\partial (\rho hu)}{\partial t} + \frac{\partial}{\partial x} \left( \rho hu^2 + \frac{\rho g}{2} h^2 + hP \right) + \frac{\partial (\rho huv)}{\partial y} = -(\rho gh + \frac{\gamma^2}{2} P) \frac{\partial z_b}{\partial x}, \quad (73) \text{eq:mod_sw2}$$

$$\frac{\partial (\rho hv)}{\partial t} + \frac{\partial (\rho huv)}{\partial x} + \frac{\partial}{\partial y} \left( \rho hv^2 + \frac{\rho g}{2} h^2 + hP \right) = -(\rho gh + \frac{\gamma^2}{2} P) \frac{\partial z_b}{\partial y}, \quad (74) \text{?eq:mod_sw2bis?}$$

$$\frac{\partial (\rho hw)}{\partial t} + \nabla_0 \cdot (\rho hw \mathbf{u}) = \gamma P, \quad (75) \text{eq:mod_sw3}$$

$$\frac{\partial}{\partial t} \left( \rho hP + \frac{\rho^2 g h^2}{2} \right) + \nabla_0 \cdot \left( \left( \rho hP + \frac{\rho^2 g h^2}{2} \right) \mathbf{u} \right) + \rho_0^2 c^2 \operatorname{div}_{sw}^\gamma (\mathbf{u}) = 0, \quad (76) \text{eq:mod_sw4}$$

where  $\rho, \mathbf{u} = (u, v, w)^T, P$  represent respectively a density, a velocity vector, and a pressure term, all functions of  $(t, x, y)$ , and  $\gamma$  a parameter.

(prop:model0\_deriv)

The model described in Prop. 3 satisfies an energy balance described in the following proposition.

**Proposition 4** Smooth solutions of the system (71)-(76) satisfy the energy balance

$$\begin{aligned} \frac{\partial}{\partial t} \left( \frac{\rho h}{2} |\mathbf{u}|^2 + \frac{1}{2} \left( 2 - \frac{\gamma^2}{2} \right) \rho g h z_b + \rho h e \right) + \nabla_0 \cdot \left( \mathbf{u} \left( \frac{\rho h}{2} |\mathbf{u}|^2 \right. \right. \\ \left. \left. + \frac{1}{2} \left( 2 - \frac{\gamma^2}{2} \right) \rho g h z_b + \rho h e + \frac{\rho g}{2} h^2 + hP \right) \right) = -\frac{\gamma}{2} \rho g h w, \end{aligned} \quad (77) \quad \boxed{\text{eq:mod\_sw6\_0}}$$

where  $e$  is defined by

$$e = \frac{1}{2\rho_0^2 c^2} \left( P + \frac{\rho g h}{2} \right)^2. \quad (78) \quad \boxed{\text{eq:e\_rho\_c}}$$

And Eq. (77) can be written in a conservative form

$$\begin{aligned} \frac{\partial}{\partial t} \left( \frac{\rho h}{2} |\mathbf{u}|^2 + \rho g \frac{h(\eta + z_b)}{2} + \frac{1}{2} \left( 2 - \frac{\gamma^2}{2} \right) \rho g h z_b + \rho h e \right) + \nabla_0 \cdot \left( \mathbf{u} \left( \frac{\rho h}{2} |\mathbf{u}|^2 \right. \right. \\ \left. \left. + \rho g \frac{h(\eta + z_b)}{2} + \frac{1}{2} \left( 2 - \frac{\gamma^2}{2} \right) \rho g h z_b + \rho h e + \frac{\rho g}{2} h^2 + hP \right) \right) = 0. \end{aligned} \quad (79) \quad \boxed{\text{eq:mod\_sw6\_1}}$$

(prop:energy0\_deriv)

In other words, (78) gives a shallow water version of the internal energy defined after (36).

The two operators  $\text{div}_{sw}^\gamma$  and  $\nabla_{sw}^\gamma$  defined by (12),(13) appear in Eqs. (71)-(76) so that we can rewrite this system in a more compact form

$$\frac{\partial \eta}{\partial t} + \mathbf{u} \cdot \nabla_0 (\eta + z_b) = \gamma w, \quad (80) \quad \boxed{\text{eq:mod\_sw00}}$$

$$\frac{\partial(\rho h)}{\partial t} + \nabla_0 \cdot (\rho h \mathbf{u}) = 0, \quad (81) \quad \boxed{\text{eq:mod\_sw11}}$$

$$\frac{\partial(\rho h \mathbf{u})}{\partial t} + \nabla_0 \cdot (\rho h \mathbf{u} \otimes \mathbf{u}) + \nabla_0 \left( \frac{\rho g}{2} h^2 \right) + \nabla_{sw}^\gamma P = -g h \nabla_0 z_b, \quad (82) \quad \boxed{\text{eq:mod\_sw22?}}$$

$$\frac{\partial}{\partial t} \left( \rho h P + \frac{\rho^2 g h^2}{2} \right) + \nabla_0 \cdot \left( \left( \rho h P + \frac{\rho^2 g h^2}{2} \right) \mathbf{u} \right) + \rho_0^2 c^2 \text{div}_{sw}^\gamma (\mathbf{u}) = 0, \quad (83) \quad \boxed{\text{eq:mod\_sw33}}$$

An important point is that whatever the value of  $\gamma$ , these operators satisfy the duality relation

$$\int_{\Omega} \nabla_{sw}^\gamma (f) \cdot \mathbf{w} d\mathbf{x} = - \int_{\Omega} \text{div}_{sw}^\gamma (\mathbf{w}) f d\mathbf{x} + \int_{\Gamma} h f \mathbf{w} \cdot \mathbf{n} ds, \quad (84) \quad \boxed{\text{todeq:duality}}$$

where the vector  $\mathbf{n} = (n_x, n_y, 0)^T$  is the outward unit normal vector to the boundary  $\Gamma$ , see Fig. 1. In Eq. (84),  $f$  and  $\mathbf{w}$  belong to suitable function spaces that will be



precised later. Notice that, we also have a local form of (84) under the form

$$\nabla_{sw}^\gamma(f) \cdot \mathbf{w} = \nabla_0 \cdot (hf\mathbf{w}) - \text{div}_{sw}^\gamma(\mathbf{w})f. \quad (85) \quad \boxed{\text{todeq:duality_loc}}$$

**Proof (Proof of Prop. 3)** It is easy to see (cf. [22, Lemma 2.1]) that a depth averaging of the compressible Euler system with gravity and free surface (31)-(33) – with, according to paragraph 2.4,  $\tilde{\rho}\tilde{c} \approx \rho_0c$  – completed with the boundary conditions (22),(24),(34) leads to

$$\frac{\partial}{\partial t} \int_{z_b}^{\eta} \tilde{\rho} dz + \nabla_{x,y} \cdot \int_{z_b}^{\eta} \tilde{\rho} \mathbf{v} dz = 0, \quad (86) \quad \boxed{\text{eq:euler_3d11}}$$

$$\frac{\partial}{\partial t} \int_{z_b}^{\eta} \tilde{\rho} \mathbf{v} dz + \nabla_{x,y} \cdot \int_{z_b}^{\eta} \tilde{\rho} \mathbf{v} \otimes \mathbf{v} dz + \nabla_{x,y} \int_{z_b}^{\eta} \tilde{p} dz = \tilde{p}(t, \mathbf{x}, z_b(\mathbf{x})) \nabla_{x,y} z_b \quad (87) \quad \boxed{\text{eq:euler_3d22?}}$$

$$\frac{\partial}{\partial t} \int_{z_b}^{\eta} \tilde{\rho} u_3 dz + \nabla_{x,y} \cdot \int_{z_b}^{\eta} \tilde{\rho} u_3 \mathbf{v} dz = \tilde{p}(t, \mathbf{x}, z_b(\mathbf{x})) - \int_{z_b}^{\eta} \tilde{\rho} g dz, \quad (88) \quad \boxed{\text{eq:euler_3d33?}}$$

$$\frac{\partial}{\partial t} \int_{z_b}^{\eta} \tilde{\rho} \tilde{p} dz + \nabla_{x,y} \cdot \int_{z_b}^{\eta} \tilde{\rho} \tilde{p} \mathbf{v} dz + \rho_0^2 c^2 \int_{z_b}^{\eta} \nabla \cdot \mathbf{U} dz = 0. \quad (89) \quad \boxed{\text{eq:euler_3d44}}$$

As in [2] we are now going to make some assumptions concerning the variations along the vertical axis of the velocity field  $\mathbf{U}$ , the density  $\tilde{\rho}$  and of the pressure  $\tilde{p}$ . In order to be consistent with the shallow water assumption, the choice

$$u_1(t, \mathbf{x}, z) = u(t, \mathbf{x}), \quad u_2(t, \mathbf{x}, z) = v(t, \mathbf{x}), \quad \tilde{\rho}(t, \mathbf{x}, z) = \rho(t, \mathbf{x}), \quad (90) \quad \boxed{\text{eq:u12_gal}}$$

is natural since it consists in assimilating the horizontal velocity field and the density with their vertical means. For the velocity  $u_3$  and the pressure  $\tilde{p}$ , we choose

$$u_3(t, \mathbf{x}, z) = \varphi_\delta \left( \frac{\eta - z}{h} \right) w(t, \mathbf{x}), \quad (91) \quad \boxed{\text{eq:w_gal?}}$$

$$\tilde{p}(t, \mathbf{x}, z) = \rho g (\eta - z) + \psi_\delta \left( \frac{\eta - z}{h} \right) P(t, \mathbf{x}), \quad (92) \quad \boxed{\text{eq:p_gal?}}$$

and the two families of functions  $\psi_\delta = \psi_\delta(z)$  and  $\varphi_\delta = \varphi_\delta(z)$  satisfy

$$\begin{cases} \int_0^1 \varphi_\delta(z) dz = \int_0^1 \psi_\delta(z) dz = \frac{1}{2} \int_0^1 \varphi_\delta(z) \psi_\delta'(z) dz = 1, \\ \psi_\delta(1) = \delta, \psi_\delta(0) = 0, \varphi_\delta(1) = 1. \end{cases} \quad (93) \quad \boxed{\text{eq:psi_3}}$$

Notice that these choices are similar to those in [2]. Figure 2 in [2, paragraph 2.3.2] illustrates the shape of the functions  $\psi_\delta$  and  $\varphi_\delta$  for two typical values of  $\delta$  namely  $\delta = 2$  and  $\delta = 3/2$  (corresponding to  $\gamma = 2$  and  $\gamma = \sqrt{3}$ ). It appears that the functions  $\psi_\delta$  and  $\varphi_\delta$  do not significantly differ when  $\delta = 2$  or when  $\delta = 3/2$ , the choice  $\delta = 2$  corresponding to a linear profile.

The duality relation (37) is a guideline for the definition of the shallow water version of the divergence operator. Therefore, from (90)-(93) and using an integration by parts, we obtain

$$\begin{aligned}
\int_{z_b}^{\eta} \psi_{\delta} \left( \frac{\eta - z}{h} \right) \nabla \cdot \mathbf{U} dz &= \int_{z_b}^{\eta} \psi_{\delta} \left( \frac{\eta - z}{h} \right) \nabla_0 \cdot \mathbf{u} dz + \left[ u_3 \psi_{\delta} \left( \frac{\eta - z}{h} \right) \right]_{z_b}^{\eta} \\
&\quad - \int_{z_b}^{\eta} w \varphi_{\delta} \left( \frac{\eta - z}{h} \right) \frac{\partial}{\partial z} \psi_{\delta} \left( \frac{\eta - z}{h} \right) dz \\
&= 2w + h \nabla_0 \cdot \mathbf{u} - \delta \mathbf{u} \cdot \nabla_0 z_b,
\end{aligned} \tag{94} \text{eq:psi_div_u}$$

where (22) has been used and allowing us to write

$$\begin{aligned}
\int_{z_b}^{\eta} \bar{p} \nabla \cdot \mathbf{U} dz &\approx \int_{z_b}^{\eta} \left( \frac{1}{h} \int_{z_b}^{\eta} g(\eta - z_1) dz_1 + P \right) \psi_{\delta} \left( \frac{\eta - z}{h} \right) \nabla \cdot \mathbf{U} dz \\
&= \left( \frac{\rho g h}{2} + P \right) \int_{z_b}^{\eta} \psi_{\delta} \left( \frac{\eta - z}{h} \right) \nabla \cdot \mathbf{U} dz \\
&= \left( \frac{\rho g h}{2} + P \right) (2w + h \nabla_0 \cdot \mathbf{u} - \delta \mathbf{u} \cdot \nabla_0 z_b).
\end{aligned} \tag{95} \text{eq:pdiv_u}$$

The computations (94),(95) are used to approximate the last term in Eq. (89) under the form

$$\int_{z_b}^{\eta} \nabla \cdot \mathbf{U} dz \approx 2w + h \nabla_0 \cdot \mathbf{u} - \delta \mathbf{u} \cdot \nabla_0 z_b.$$

And with the choices (90)-(93), the system (86)-(89) writes

$$\frac{\partial(\rho h)}{\partial t} + \nabla_0 \cdot (\rho h \mathbf{u}) = 0, \tag{96} \text{eq:euler_3d111}$$

$$\frac{\partial(\rho h u)}{\partial t} + \frac{\partial}{\partial x} \left( \rho h u^2 + \frac{\rho g}{2} h^2 + h P \right) + \frac{\partial(\rho h u v)}{\partial y} = -(\rho g h + \delta P) \frac{\partial z_b}{\partial x}, \tag{97} \text{?eq:euler_3d222?}$$

$$\frac{\partial(\rho h v)}{\partial t} + \frac{\partial(\rho h u v)}{\partial x} + \frac{\partial}{\partial y} \left( \rho h v^2 + \frac{\rho g}{2} h^2 + h P \right) = -(\rho g h + \delta P) \frac{\partial z_b}{\partial y}, \tag{98} \text{?eq:euler_3d222bis?}$$

$$\frac{\partial(\rho h w)}{\partial t} + \nabla_0 \cdot (\rho h w \mathbf{u}) = \delta P, \tag{99} \text{?eq:euler_3d333?}$$

$$\begin{aligned}
\frac{\partial}{\partial t} \left( \rho h P + \frac{\rho^2 g h^2}{2} \right) + \nabla_0 \cdot \left( \left( \rho h P + \frac{\rho^2 g h^2}{2} \right) \mathbf{u} \right) \\
+ \rho_0^2 c^2 (2w + h \nabla_0 \cdot \mathbf{u} - \delta \mathbf{u} \cdot \nabla_0 z_b) = 0.
\end{aligned} \tag{100} \text{eq:euler_3d444}$$

Using the choices (90)-(93), Eq. (19) gives

$$\frac{\partial}{\partial t} \left( \frac{\rho h (\eta + z_b)}{2} \right) + \nabla_0 \cdot \left( \frac{\rho h (\eta + z_b)}{2} \mathbf{u} \right) = \rho h w, \tag{101} \text{eq:def_w_sw_0}$$

and combining (101) with (96) gives

$$\frac{\partial \eta}{\partial t} + \mathbf{u} \cdot \nabla_0 (\eta + z_b) = 2w. \tag{102} \text{eq:def_w_sw_1}$$

Finally, a simple change of variables, namely  $w = \gamma \hat{w}/2$  with  $\gamma^2 = 2\delta$  in the system (96)-(100),(102) leads to Eqs. (71)-(76) where the symbol  $\hat{\cdot}$  has been dropped.  $\square$

**Proof (Proof of Prop. 4)** After simple computations and using Eq. (72), Eq. (76) multiplied by  $(P + \frac{\rho gh}{2})/(\rho_0^2 c^2)$  gives

$$\frac{\partial}{\partial t} \left( \frac{\rho h}{2\rho_0^2 c^2} \left( P + \frac{\rho gh}{2} \right)^2 \right) + \nabla_0 \cdot \left( \frac{\rho h}{2\rho_0^2 c^2} \left( P + \frac{\rho gh}{2} \right)^2 \mathbf{u} \right) + \left( P + \frac{\rho gh}{2} \right) \operatorname{div}_{sw}^\gamma(\mathbf{u}) = 0, \quad (103) \quad \boxed{\text{eq:e\_sw\_rho}}$$

Now, taking the scalar product of Eqs. (73)-(75) with  $\mathbf{u}$ , using the duality relation (85) and adding the obtained relation with (103) gives (77).

And the sum of Eq. (77) with (101) multiplied by  $g$  – in which the change of variable  $w = \gamma \hat{w}/2$  is done and the symbol  $\hat{\cdot}$  has been dropped – gives (79).

## 2.6 When the density is almost constant

On the one hand, the propagation of acoustic waves requires a compressible medium, on the other hand the variations of the fluid density are often neglected e.g. when considering a linearized version of the Euler system (31)-(33).

In this paragraph, we assume that the variations of the fluid density have little influence over the hydrodynamic regime and the waves propagation, that is not a strong assumption for water, see (60). Nevertheless, it is not possible to simply consider that the density is constant in the considered models. Indeed, the assumption  $\rho = cst$  in the 3d case – Eq. (15) – or in the shallow water context – Eq. (101) – leads to a divergence free condition that is not compatible with the equations governing the pressure variations namely Eq. (33) or Eq. (76) in the shallow water regime.

Hence when the variations of the fluid density can be neglected, the proposition 3 can be reformulated as follows.

**Proposition 5** *Assuming the setting of Prop. 3 and neglecting the variation of the fluid density, a shallow water approximation of the compressible Euler system (31)-(33) is given by*

$$\frac{\partial h}{\partial t} + \nabla_0 \cdot (h\mathbf{u}) = 0, \quad (104) \quad \boxed{\text{eq:eq1\_mod}}$$

$$\frac{\partial (h\mathbf{u})}{\partial t} + \nabla_0 \cdot (h\mathbf{u} \otimes \mathbf{u}) + \nabla_0 \left( \frac{g}{2} h^2 \right) + \nabla_{sw}^\gamma p = -gh \nabla_0 z_b, \quad (105) \quad \boxed{\text{eq:eq2\_mod}}$$

$$\frac{\partial}{\partial t} \left( hp + \frac{gh^2}{2} \right) + \nabla_0 \cdot \left( \left( hp + \frac{gh^2}{2} \right) \mathbf{u} \right) + c^2 \operatorname{div}_{sw}^\gamma(\mathbf{u}) = 0, \quad (106) \quad \boxed{\text{eq:eq3\_mod}}$$

(prop:model\_deriv) where the operators  $\nabla_{sw}^\gamma$  and  $\operatorname{div}_{sw}^\gamma$  are defined by (12),(13).

**Proof (Proof of Prop. 5)** The model (81)-(83) is nothing else than a rewriting of Eqs. (72)-(76) where the variations of the fluid density are neglected i.e.  $\rho \equiv \rho_0$  and we have introduced  $p = P/\rho_0$ .  $\square$

For the model obtained in Prop. 3, Eq. (101) is crucial to obtain an energy balance. In order to obtain an energy balance for the model given in Prop. 5, we introduce a function  $\tilde{\zeta} = \tilde{\zeta}(t, \mathbf{x})$  solution of the transport equation

$$\frac{\partial \tilde{\zeta}}{\partial t} + \mathbf{u} \cdot \nabla_0 \tilde{\zeta} = \gamma w, \quad (107) \quad \text{eq:def\_xi}$$

or equivalently using (104)

$$\frac{\partial(h\tilde{\zeta})}{\partial t} + \nabla_0 \cdot (h\tilde{\zeta}\mathbf{u}) = h\gamma w. \quad (108) \quad \text{eq:def\_xibis}$$

From the definition (13) and using (104), we can write

$$h \operatorname{div}_{sw}^\gamma(\mathbf{u}) = h\gamma w + h\nabla_0 \cdot (h\mathbf{u}) - h\mathbf{u} \cdot \nabla_0 \zeta = h\gamma w - \frac{\partial(h\zeta)}{\partial t} - \nabla_0 \cdot (h\zeta\mathbf{u}),$$

and hence, Eq. (108) can be written under the form

$$\frac{\partial(h\hat{\zeta})}{\partial t} + \nabla_0 \cdot (h\hat{\zeta}\mathbf{u}) = h \operatorname{div}_{sw}^\gamma(\mathbf{u}), \quad (109) \quad \text{eq:def\_xi1}$$

with  $\hat{\zeta} = \tilde{\zeta} - h - \gamma^2 z_b/2 = \tilde{\zeta} - \zeta$ . Notice that  $\tilde{\zeta}$  is an approximation in the constant density case of the variable  $\eta + z_b$  governed by Eq. (80). And from Eq. (106),  $\hat{\zeta} = O(1/c^2)$ . The existence of solution for Eq. (107) has been widely studied, let us mention the contributions of Di Perna and Lions by the means of renormalized solutions [17] and two extensions [16, 27], see also [11]. We assume here that the variables  $h$ ,  $\mathbf{u}$  and the quantity  $z_b$  are regular enough so that these existence results are valid.

As already mentioned, the assumption  $\rho = cst$  implies that mass and volume are conserved that could be seen as contradictory with the capability of acoustic waves to propagate – with finite speed – since it requires a compressibility in the considered media. But the quantity  $\hat{\zeta}$  can also be related to the temperature effects and allows to circumvent this difficulty. More precisely, when  $\lambda = \mu = 0$  and assuming

$$\left( \frac{\partial \tilde{T}}{\partial \tilde{\rho}} \right)_s = \tau = cst,$$

Eq. (109) looks like a shallow water version of Eq. (49), the variable  $\hat{\zeta}$  corresponding to a shallow water approximation of the quantity  $\frac{\tilde{T}}{\tau \rho_0}$ . Hence the variations of  $\hat{\zeta}$  correspond to volume variations and can be assimilated to dilatation effects generated by temperature variations.

**Proposition 6** *Smooth solutions of the system (104)-(106) satisfy the energy balance*

$$\begin{aligned} \frac{\partial}{\partial t} \left( \frac{h}{2} |\mathbf{u}|^2 + \frac{g}{2} \left( 2 - \frac{\gamma^2}{2} \right) h z_b + \frac{h}{2c^2} \left( p + \frac{gh}{2} \right)^2 \right) + \nabla_0 \cdot \left( \mathbf{u} \left( \frac{h}{2} |\mathbf{u}|^2 \right. \right. \\ \left. \left. + \frac{g}{2} \left( 2 - \frac{\gamma^2}{2} \right) h z_b + \frac{h}{2c^2} \left( p + \frac{gh}{2} \right)^2 + \frac{g}{2} h^2 + hp \right) \right) = -\frac{\gamma h}{2} g w, \end{aligned} \quad (110) \quad \boxed{\text{eq: eq8\_mod}}$$

that is a shallow water version of Eq. (35). Equation (110) can be rewritten under a conservative form given by

$$\begin{aligned} \frac{\partial}{\partial t} \left( \frac{h}{2} |\mathbf{u}|^2 + g \frac{h(\eta + z_b)}{2} + g \frac{h\hat{\zeta}}{2} + \frac{h}{2c^2} \left( p + \frac{gh}{2} \right)^2 \right) + \nabla_0 \cdot \left( \mathbf{u} \left( \frac{h}{2} |\mathbf{u}|^2 \right. \right. \\ \left. \left. + g \frac{h(\eta + z_b)}{2} + \frac{h}{2c^2} \left( p + \frac{gh}{2} \right)^2 + \frac{g}{2} h^2 + g \frac{h\hat{\zeta}}{2} + hp \right) \right) = 0. \end{aligned} \quad (111) \quad \boxed{\text{eq: eq9\_mod}}$$

(prop:energy\_deriv)

Multiplying Eq. (106) by  $p + gh/2$  and after simple computations, we obtain the relation

$$\frac{\partial}{\partial t} \left( \frac{h}{2c^2} \left( p + \frac{gh}{2} \right)^2 \right) + \nabla_0 \cdot \left( \frac{h}{2c^2} \left( \left( p + \frac{gh}{2} \right)^2 \mathbf{u} \right) \right) + \left( p + \frac{gh}{2} \right) \text{div}_{sw}^\gamma(\mathbf{u}) = 0. \quad (112) \quad \boxed{\text{eq: momu3}}$$

And comparing (112) with (36) we obtain that when the density is kept constant, the internal energy is given by  $(p + gh/2)^2/(2c^2)$ , see also (78).

**Proof (Proof of Prop. 6)** Taking the scalar product of Eq. (105) by  $\mathbf{u}$  gives

$$\frac{\partial}{\partial t} \left( \frac{h}{2} |\mathbf{u}|^2 \right) + \nabla_0 \cdot \left( \mathbf{u} \left( \frac{h}{2} |\mathbf{u}|^2 + \frac{g}{2} h^2 \right) \right) + \mathbf{u} \cdot \nabla_{sw}^\gamma p - \frac{g}{2} h^2 \nabla_0 \cdot \mathbf{u} + g h \mathbf{u} \cdot \nabla_0 z_b = 0. \quad (113) \quad \boxed{\text{eq: momu2}}$$

Using the duality relation (85) in (113) and adding Eq. (112) gives (110).

Besides, using Eq. (109) multiplied by  $g/2$  and added to (110) gives (111).  $\square$

## 2.7 The boundary conditions

The set of equations (104)-(106) is completed with the following boundary conditions. We are considering a channel with an inlet  $\Gamma_{in}$  and an outlet  $\Gamma_{out}$  and we impose specific conditions on each of them, see Fig. 1. The inflow is imposed by a given discharge  $\mathbf{q}_g$  on  $\Gamma_{in}$ , and a water depth  $h_g$  is imposed on  $\Gamma_{out}$ . Finally, we prescribe slip boundary conditions for the velocity at the walls of the channel  $\Gamma_s$ . Hence we have

$$h \mathbf{u}(t, \mathbf{x}) = \mathbf{q}_g(t, \mathbf{x}), \quad \text{on } \Gamma_{in}, \quad (114) \quad \boxed{\text{todeq:bc\_inflow?}}$$

$$h(t, \mathbf{x}) = h_g(t, \mathbf{x}), \quad \text{on } \Gamma_{out}, \quad (115) \quad \boxed{\text{todeq:bc\_outflow?}}$$

$$\mathbf{u}(t, \mathbf{x}) \cdot \mathbf{n} = 0, \quad \text{on } \Gamma_s. \quad (116) \quad \boxed{\text{todeq:bc\_slip?}}$$

Notice that we can replace the prescribed water depth at the outflow by a free outflow consisting in imposing a Neumann boundary condition over the elevation

$$\nabla_0 h \cdot \mathbf{n} = 0, \quad \text{on } \Gamma_{out}.$$

## 2.8 Dispersion relation

The model (104)-(106) is a shallow water type model with compressible effects coming from the acoustic wave propagation. A fundamental question is to know what are the velocities of the waves propagating in such a model and typically the influence of the sound speed  $c$  over these velocities.

Let us consider the system (104)-(106) in the one-dimensional case, with flat topography and where the temperature variations are neglected. It has the form of an advection-reaction system, namely

$$\frac{\partial Y}{\partial t} + A \frac{\partial Y}{\partial x} + BY = 0, \quad (117) \text{eq:advection\_reaction}$$

with

$$Y = \begin{pmatrix} h \\ u \\ w \\ p \end{pmatrix}, \quad A = \begin{pmatrix} u & h & 0 & 0 \\ g + \frac{p}{h} & u & 0 & 1 \\ 0 & 0 & u & 0 \\ 0 & c^2 - \frac{gh}{2} & 0 & u \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{\gamma}{h} \\ 0 & 0 & \frac{\gamma c^2}{h} & 0 \end{pmatrix}.$$

Let us introduce  $Y_0 \in \mathbb{R}^4$  and  $k, \omega$  being two constants, namely the wave number and the frequency. A necessary condition so that the system (117) admits a solution having the form  $Y = Y_0 e^{i(kx - \omega t)}$  is that

$$\det(i\omega I_4 - ikA + B) = 0,$$

$I_4$  being the identity matrix of dimension 4. This leads to the four roots

$$\frac{\omega}{k} = u \pm \frac{\sqrt{2}}{2} \sqrt{C_{sw,1} \pm \sqrt{(C_{sw,1})^2 - C_{sw,3}}}, \quad (118) \text{eq:four\_roots?}$$

with  $C_{sw,1} = c^2 + gh + p + c^2 \frac{\gamma^2}{(hk)^2}$  and  $C_{sw,2} = 4c^2 \gamma^2 (gh + p)$ . As in (58), we can assume that  $c \gg 1$  leading to the four approximated roots

$$u \pm \gamma \sqrt{\frac{gh + p}{\gamma^2 - (hk)^2}} + \mathcal{O}\left(\frac{1}{c^2}\right), u \pm c \sqrt{1 + \frac{\gamma^2}{(hk)^2}} + \mathcal{O}\left(\frac{1}{c}\right). \quad (119) \text{eq:estim\_vp\_sw}$$

*Remark 6* From the estimates (119), it appears that the model (104)-(106) is able to propagate both water waves and acoustic waves. But since we are in a shallow water

context, we have  $hk \ll 1$  and for the acoustic waves we do not exactly recover the expected velocities  $u \pm c$ .

*Remark 7* In the context of wave propagation i.e. with flat bottom and assuming the water depth has the form  $h = h_0 + f(kx - \omega t)$  with  $h_0 = cst$  and  $|f(\cdot)| \ll h_0$ , it is easy to see that the first term in (119) becomes

$$\gamma \sqrt{\frac{gh_0}{\gamma^2 - (hk)^2}} + \mathcal{O}\left(\frac{1}{c^2}\right),$$

corresponding for  $\gamma = \sqrt{3}$  in the context of large wavelength ( $kh_0 \ll 1$ ) and up to  $\mathcal{O}((kh_0)^2)$  terms, to the classical Airy wave dispersion relation [1].

## 2.9 A pseudo-compressible model

As we have seen in the previous paragraph, if  $c$  is chosen corresponding to the sound speed in water, then the model (104)-(106) is able to propagate, in a shallow water context, both water and acoustic waves. But since  $c \gg 1$ , we introduce

$$\varepsilon = \frac{1}{c^2},$$

then the model (104)-(106) can be seen as a pseudo-compressible version of the model (9)-(11) allowing to derive an explicit in time numerical scheme that will be studied in the following section. More precisely Eq. (106) writes

$$\varepsilon \left( \frac{\partial}{\partial t} \left( hp + \frac{gh^2}{2} \right) + \nabla_0 \cdot \left( \left( hp + \frac{gh^2}{2} \right) \mathbf{u} \right) \right) + \operatorname{div}_{s_w}^\gamma(\mathbf{u}) = 0,$$

and corresponds to Eq. (11) when  $\varepsilon$  goes to 0. Following the results obtained in Prop. 6, it is important to notice that in the formulation (104)-(106), the limit  $\varepsilon \rightarrow 0$  is not singular. Unlike the incompressible limit of compressible models, the limit when  $\varepsilon \rightarrow 0$  of the model (104)-(106) is the model (9)-(11).

Hence the model (104)-(106) can be seen as

- a dispersive shallow water type model propagating water and acoustic waves,
- a pseudo-compressible dispersive model whose numerical resolution is easier to implement compared to a fully compressible model. This second aspect is studied in the two next sections.

Notice that several authors have proposed approximated versions of the divergence free constraint for dispersive models [19, 25], for which the origin of relaxation is not related to acoustics. The model formulation (104)-(106) is similar to the one studied in [20] but the derivation process – based on the so-called hyperbolic divergence cleaning [15] – differs. The numerical strategy proposed in [20] based on high order discontinuous Galerkin schemes is also different from the one presented hereafter.

### 3 The numerical scheme (explicit in time)

`?(sec:scheme)?` In this section, we propose and study a numerical scheme for the system (104)-(106) with  $\varepsilon = 1/c^2$ .

Let us introduce the notations

$$X = \begin{pmatrix} h \\ hu \\ hv \\ hw \end{pmatrix}, \quad F(X) = \begin{pmatrix} hu & hv \\ hu^2 + \frac{g}{2}h^2 & huv \\ huv & hv^2 + \frac{g}{2}h^2 \\ huw & hvw \end{pmatrix},$$

and  $S(X) = (0, -gh\nabla_0(z_b))^T$ ,  $R = (0, \nabla_{sw}^\gamma p)^T$  where  $\nabla_{sw}^\gamma p$  is defined by (12). Then, the system (104)-(106) can be written under the form

$$\frac{\partial X}{\partial t} + \nabla_{x,y} \cdot F(X) + R = S(X), \quad (120) \quad \boxed{\text{eq: eqX1}}$$

$$\varepsilon \left( \frac{\partial}{\partial t} (hp + \frac{gh^2}{2}) + \nabla_0 \cdot \left( (hp + \frac{gh^2}{2}) \mathbf{u} \right) \right) + \text{div}_{sw}^\gamma(\mathbf{u}) = 0. \quad (121) \quad \boxed{\text{eq: eqX2}}$$

#### 3.1 Time discretisation

`(subsec:time_scheme)` Let us sketch the main steps of the procedure. We set  $t^0$  the initial time and  $t^{n+1} = t^n + \Delta t^n$  where  $\Delta t^n$  satisfies a stability condition (CFL) precised later – at the fully discrete level – and the state  $X^n$  denotes an approximation of  $X(t^n)$ . For each time step, we consider an intermediate state which will be denoted with the superscript  $n+1/2$ . The first step consists in solving the Saint-Venant part of the system (120) with the topography source term and completed with the hyperbolic part of (121) in order to obtain the state  $X^{n+1/2} = (h^{n+1/2}, (h\mathbf{u})^{n+1/2})^T$  and  $(hp)^{n+1/2}$ . Then the state  $X^{n+1}$  is computed taking into account the contribution of the non-hydrostatic pressure terms.

More precisely, in the system (120),(121) water waves generally propagate at a lower velocity than acoustic waves. Therefore, we propose an explicit time scheme – constrained by an associated CFL condition that will be precised in the fully discrete case, see (159) – for the Saint-Venant part of Eq. (120). For the dispersive terms, we adopt an iterative resolution scheme explicit in time and constrained by a generally more restrictive CFL condition associated with the sound speed. Hence, the proposed semi-discretization in time consists in the following time-splitting strategy

$$\begin{cases} X^{n+1/2} = X^n - \Delta t^n \nabla_{x,y} \cdot F(X^n) + \Delta t^n S(X^n), \\ (hp)^{n+1/2} = (hp)^n - \frac{g}{2} (h^{n+1/2})^2 - h^n)^2 \\ \quad - \Delta t^n \nabla_0 \cdot \left( h^n (p^n + \frac{g}{2} (h^n)^2) \mathbf{u}^n \right), \end{cases} \quad (122) \quad \boxed{\text{eq: eq1\_sd}}$$

$$\begin{cases} p^{n+1/2,k+1} = p^{n+1/2,k} - \frac{\Delta t^n}{\varepsilon K h^{n+1}} \text{div}_{sw}^\gamma \mathbf{u}^{n+1/2,k}, \\ \mathbf{u}^{n+1/2,k+1} = \mathbf{u}^{n+1/2,k} - \frac{\Delta t^n}{K h^{n+1}} \nabla_{sw}^\gamma p^{n+1/2,k+1}, \end{cases} \quad k = 1, \dots, K \quad (123) \quad \boxed{\text{eq: eq2\_sd}}$$



with  $p^{n+1/2,1} = p^{n+1/2}$ ,  $\mathbf{u}^{n+1/2,1} = \mathbf{u}^{n+1/2}$ ,  $p^{n+1} = p^{n+1/2,K+1}$ ,  $\mathbf{u}^{n+1} = \mathbf{u}^{n+1/2,K+1}$  and where for the first component of  $X$  we have  $h^{n+1} = h^{n+1/2}$  since the first component of  $R$  is zero. Notice that the two operators  $\text{div}_{sw}^\gamma$  and  $\nabla_{sw}^\gamma$  are defined by Eqs. (12),(13) using  $h^{n+1}$ .  $K$  is an integer that is defined in order to ensure a stability condition for the acoustic-like wave propagation. The value of  $K$  is precised in the fully discrete case, see (162).

### 3.2 Influence of the pseudo-compressibility over the computational costs

(subsec:influence)

In [2], the authors have studied the model (1)-(5) – that is exactly the model (104)-(106) with  $\varepsilon = 0$  – and proposed the following semi-discretization in time

$$X^{n+1/2} = X^n - \Delta t^n \nabla_{x,y} \cdot F(X^n) + \Delta t^n S(X^n), \quad (124) \text{ \codeq:pred}$$

$$(\mathbf{h}\mathbf{u})^{n+1} = (\mathbf{h}\mathbf{u})^{n+1/2} - \Delta t^n \nabla_{sw}^\gamma p^{n+1}, \quad (125) \text{ \codeq:corr}$$

$$\text{div}_{sw}^\gamma \mathbf{u}^{n+1} = 0, \quad (126) \text{ \codeq:divu}$$

in which Eq. (125) allows to correct the predicted value  $X^{n+1/2}$  in order to obtain a state which satisfies the divergence free condition (126). The equation satisfied by the pressure is then an elliptic equation which is obtained by applying the shallow water divergence operator  $\text{div}_{sw}^\gamma$  to Eq. (125) and reads

$$\text{div}_{sw}^\gamma \left( \frac{1}{h^{n+1}} \nabla_{sw}^\gamma p^{n+1} \right) = \frac{1}{\Delta t^n} \text{div}_{sw}^\gamma \left( \frac{(\mathbf{h}\mathbf{u})^{n+1/2}}{h^{n+1}} \right). \quad (127) \text{ \codeq:lapsew}$$

Once the pressure has been determined by the elliptic equation (127), the correction step (125) gives the final step  $X^{n+1}$ .

The main drawback of the time scheme (124)-(126) is the numerical cost of the resolution of Eq. (127). And Eq. (106) can be seen as a relaxed version of Eq. (126) allowing to replace the step (125)-(126) by the iterative method (123) applied to the model (120)-(121). More precisely, inserting the second equation of (123) (at iteration  $k - 1$ ) into the first one gives the relation

$$\begin{aligned} p^{k+1} &= p^k - \frac{\Delta t^n}{\varepsilon K} \text{div}_{sw}^\gamma \mathbf{u}^{n+1/2,k-1} + \frac{(\Delta t^n)^2}{\varepsilon K^2 h^{n+1}} \text{div}_{sw}^\gamma \left( \frac{1}{h^{n+1}} \nabla_{sw}^\gamma p^k \right), \\ &= 2p^k - p^{k-1} + \frac{(\Delta t^n)^2}{\varepsilon K^2 h^{n+1}} \text{div}_{sw}^\gamma \left( \frac{1}{h^{n+1}} \nabla_{sw}^\gamma p^k \right), \end{aligned} \quad (128) \text{ \codeq:wave_epsilon_p}$$

where the superscripts  $n+1/2$  have been dropped. Equation (128) appears as an explicit in time discretization of a wave equation. As expected, when  $\varepsilon$  tends to 0, Eq. (128) reduces to Eq. (127). Likewise, inserting the first equation of (123) into the second one gives the relation

$$\mathbf{u}^{k+1} = 2\mathbf{u}^k - \mathbf{u}^{k-1} + \frac{(\Delta t^n)^2}{\varepsilon K^2 h^{n+1}} \nabla_{sw}^\gamma \left( \frac{1}{h^{n+1}} \operatorname{div}_{sw}^\gamma \mathbf{u}^k \right). \quad (129) \quad \boxed{\text{eq:wave_epsilon_u}}$$

The stability of the two discretizations (128),(129) will be examined in paragraph 4.4.

As already mentioned, if  $N$  is the number of cells in the considered mesh, the computational cost of the resolution of (127) is  $\mathcal{O}(N^{3/2})$  whereas the resolution of (123) is  $\mathcal{O}(KN) = \mathcal{O}(N/\sqrt{\varepsilon})$ . And hence, an estimation of  $\varepsilon$  is required to compare the costs of the explicit and implicit resolutions.

### 3.3 Choice of $\varepsilon$

?(subsec:choice)?

If one is interested in the simulation of both water and acoustic waves,  $\varepsilon$  is chosen so that  $\varepsilon = 1/c^2$ ,  $c$  being the sound speed. But if the objective is to approximate a relaxed version of the system (1)-(5) then  $\varepsilon$  is no more a physical parameter and has to be chosen so that the system (104)-(106) is a good approximation of the system (9)-(11). Hence, at each time step,  $\varepsilon$  can be chosen according to the computed values of the velocities and of the water depth.

And we can proceed as follows.

The energy of the model (104)-(106) behaves as

$$h \frac{u^2 + v^2 + w^2}{2} + \frac{g}{2} h \zeta + g \left( 2 - \frac{\gamma^2}{2} \right) h z_b + \frac{\varepsilon h}{2} \left( p + \frac{g}{2} h \right)^2.$$

Hence, we have to choose  $\varepsilon$  such that

$$\varepsilon \left( p + \frac{g}{2} h \right)^2 \ll u^2 + v^2 + w^2 + g(\eta + z_b) + g \left( 4 - \gamma^2 \right) z_b. \quad (130) \quad \boxed{\text{eq:epsilon_estim1}}$$

Another possibility is to recall that  $\varepsilon$  is related to the sound speed with  $\varepsilon = 1/c^2$  and hence  $\varepsilon$  has to satisfy

$$\frac{1}{\sqrt{\varepsilon}} = c \gg |u| + |v| + \sqrt{gh},$$

i.e.

$$\varepsilon \ll \frac{1}{(|u| + |v| + \sqrt{gh})^2}. \quad (131) \quad \boxed{\text{eq:epsilon_estim2}}$$

The two conditions (130) and (131) are easy to implement and similar when  $|u| + |v| + |w| \ll \sqrt{gh}$ . But, in the context of dispersive flows (130) is more appropriate since the vertical velocity  $w$  is taken into account.

We have seen that the choice  $\varepsilon = 1/c^2$  corresponds to the propagation of acoustic waves. Smaller values of  $\varepsilon$  increase the computational costs of the scheme since it enlarges the value of the number of iterations  $K$  for the resolution of (123), see paragraphs 3.1 and 3.2

#### 4 Detailed numerical scheme in 1d

A numerical scheme for the model (9)-(11) has been proposed and studied in [2]. Here we focus on the one dimensional case in order to prove the capability of the pseudo-compressible formulation.

In the one dimensional case, the model (104)-(106) writes

$$\frac{\partial h}{\partial t} + \frac{\partial(hu)}{\partial x} = 0, \quad (132) \quad \boxed{\text{eq: eq1\_1d\_mod}}$$

$$\frac{\partial(hu)}{\partial t} + \frac{\partial}{\partial x} \left( hu^2 + \frac{g}{2}h^2 + hp \right) = -(gh + \frac{\gamma^2}{2}p) \frac{\partial z_b}{\partial x}, \quad (133) \quad \boxed{\text{eq: eq2\_2d\_mod?}}$$

$$\frac{\partial(hw)}{\partial t} + \frac{\partial(huw)}{\partial x} = \gamma p, \quad (134) \quad \boxed{\text{eq: eq3\_1d\_mod?}}$$

$$\varepsilon \left( \frac{\partial}{\partial t} \left( hp + \frac{gh^2}{2} \right) + \frac{\partial}{\partial x} \left( (hp + \frac{gh^2}{2})u \right) \right) + \gamma w + h \frac{\partial u}{\partial x} - \frac{\gamma^2}{2} u \frac{\partial z_b}{\partial x} = 0 \quad (135) \quad \boxed{\text{eq: eq4\_1d\_mod}}$$

In a more compact form and with obvious notations, the system (132)-(135) becomes

$$\frac{\partial X}{\partial t} + \frac{\partial F(X)}{\partial x} + R = S(X), \quad (136) \quad \boxed{\text{eq: eqX1\_1d\_mod}}$$

$$\varepsilon \left( \frac{\partial(h\hat{p})}{\partial t} + \frac{\partial(hu\hat{p})}{\partial x} \right) + \text{div}_{sw}^\gamma(\mathbf{u}) = 0, \quad (137) \quad \boxed{\text{eq: eqX2\_1d\_mod}}$$

with  $\mathbf{u} = (u, w)^T$ ,  $\hat{p} = p + gh/2$  and

$$X = \begin{pmatrix} h \\ hu \end{pmatrix}, \quad (138) \quad \boxed{\text{eq: X}}$$

$$\nabla_{sw}^\gamma f = \begin{pmatrix} h \frac{\partial f}{\partial x} + \frac{\partial \zeta}{\partial x} f \\ -\gamma f \end{pmatrix}, \quad \text{div}_{sw}^\gamma \mathbf{u} = \frac{\partial(hu)}{\partial x} - u \frac{\partial \zeta}{\partial x} + \gamma w. \quad (139) \quad \boxed{\text{eq: def\_gradsw}}$$

Notice that the fundamental duality relation

$$\int_C p \text{div}_{sw}^\gamma \mathbf{u} \, dx = [hup]_{\partial C} - \int_C \nabla_{sw}^\gamma p \cdot \mathbf{u} \, dx, \quad (140) \quad \boxed{\text{eq: duality\_discrete}}$$

holds for any interval  $C$ .

The smooth solutions of Eqs. (132)-(135) satisfy the energy equality

$$\begin{aligned} \frac{\partial}{\partial t} \left( \frac{h}{2}(u^2 + w^2) + \frac{g}{2} \left( 2 - \frac{\gamma^2}{2} \right) + \frac{\varepsilon h}{2} \left( p + \frac{gh}{2} \right)^2 \right) + \frac{\partial}{\partial x} \left( u \left( \frac{h}{2}(u^2 + w^2) \right) \right. \\ \left. + \frac{g}{2} \left( 2 - \frac{\gamma^2}{2} \right) + \frac{\varepsilon h}{2} \left( p + \frac{gh}{2} \right)^2 + \frac{g}{2} h^2 + hp \right) = -\frac{\gamma gh}{2} w. \end{aligned} \quad (141) \quad \boxed{\text{eq: eq7\_1d\_mod}}$$

Introducing the 1d version of Eq. (109) given by

$$\frac{\partial(h\hat{\zeta})}{\partial t} + \frac{\partial}{\partial x}(h\hat{\zeta}u) = h\text{div}_{sw}^{\gamma}(\mathbf{u}), \quad (142) \quad \boxed{\text{eq: def\_xi1\_1d}}$$

allows to have a conservative form of Eq. (141) under the form

$$\frac{\partial}{\partial t}\left(\bar{E} + \frac{\varepsilon h}{2}\left(p + \frac{gh}{2}\right)^2\right) + \frac{\partial}{\partial x}\left(u\left(\bar{E} + \frac{\varepsilon h}{2}\left(p + \frac{gh}{2}\right)^2 + \frac{g}{2}h^2 + hp\right)\right) = 0, \quad (143) \quad \boxed{\text{eq: eq8\_1d\_mod?}}$$

with  $\bar{E} = h(u^2 + w^2)/2 + gh(\eta + z_b)/2 + g\frac{h\hat{\zeta}}{2}$ , see Prop. 6.

#### 4.1 Semi-discrete (in time) scheme

`(subsec: semi_d_time)?` The 1d version of the time discretization (122)-(123) writes

$$\begin{cases} X^{n+1/2} = X^n - \Delta t^n \frac{\partial F(X^n)}{\partial x} + \Delta t^n S(X^n) \\ (h\hat{p})^{n+1/2} = (h\hat{p})^n - \Delta t^n \frac{\partial(h^n \hat{p}^n u^n)}{\partial x} \end{cases} \quad (144) \quad \boxed{\text{eq: eq1\_1d\_sd}}$$

$$\begin{cases} p^{n+1/2,k+1} = p^{n+1/2,k} - \frac{\Delta t^n}{\varepsilon K h^{n+1}} \text{div}_{sw}^{\gamma} \mathbf{u}^{n+1/2,k} \\ \mathbf{u}^{n+1/2,k+1} = \mathbf{u}^{n+1/2,k} - \frac{\Delta t^n}{K h^{n+1}} \nabla_{sw}^{\gamma} p^{n+1/2,k+1} \end{cases} \quad (145) \quad \boxed{\text{eq: eq2\_1d\_sd}}$$

with  $p^{n+1/2,1} = p^{n+1/2}$ ,  $\mathbf{u}^{n+1/2,1} = \mathbf{u}^{n+1/2}$  and  $p^{n+1} = p^{n+1/2,K+1}$ ,  $\mathbf{u}^{n+1} = \mathbf{u}^{n+1/2,K+1}$  where for the first component of  $X$  we have  $h^{n+1} = h^{n+1/2}$ .

The scheme (144)-(145) is explicit in time so it is important to examine its stability w.r.t. the discretisation step  $\Delta t^n$ , this will be done in paragraph 4.4.

#### 4.2 The semi-discrete (in space) scheme

`(subsec: semi_d_space)` To approximate the solution  $X = (h, hu, hw)^T$ ,  $hp$  of the system (132)-(135), we use a combined finite volume/finite difference framework. We assume that the computational domain is discretized with  $I$  nodes  $x_i$ ,  $i = 1, \dots, I$ . We denote  $C_i$  the cell  $(x_{i-1/2}, x_{i+1/2})$  of length  $\Delta x_i = x_{i+1/2} - x_{i-1/2}$  with  $x_{i+1/2} = (x_i + x_{i+1})/2$ . We denote  $X_i = (h_i, q_{x,i}, q_{z,i})^T$  with

$$X_i \approx \frac{1}{\Delta x_i} \int_{C_i} X(t, x) dx,$$

the approximate solution at time  $t$  on the cell  $C_i$  with  $q_{x,i} = h_i u_i$ ,  $q_{z,i} = h_i w_i$ . Likewise, for the topography, we define

$$z_{b,i} = \frac{1}{\Delta x_i} \int_{C_i} z_b(x) dx.$$

The non-hydrostatic part of the pressure is discretized on a staggered grid

$$p_{i+1/2} \approx \frac{1}{\Delta x_{i+1/2}} \int_{x_i}^{x_{i+1}} p(t, x) dx,$$

with  $\Delta x_{i+1/2} = x_{i+1} - x_i$  and we set  $\hat{p}_{i+1/2} = p_{i+1/2} + gh_{i+1/2}/2$  where  $h_{i+1/2}$  is defined by  $\Delta x_{i+1/2} h_{i+1/2} = (\Delta x_i h_i + \Delta x_{i+1} h_{i+1})/2$ .

Now we propose and study the semi-discrete (in space) scheme approximating the model (136)-(137). The semi-discrete scheme writes

$$\Delta x_i \frac{\partial X_i}{\partial t} + (F_{i+1/2-} - F_{i-1/2+}) + R_i = 0, \quad (146) \text{eq:mom_d}$$

$$\Delta x_{i+1/2} \varepsilon \frac{\partial}{\partial t} (h_{i+1/2} \hat{p}_{i+1/2}) + \varepsilon (F_{\hat{p},i+1} - F_{\hat{p},i}) + \text{div}_{sw,i+1/2}^\gamma (\{\mathbf{u}_j\}) = 0, \quad (147) \text{eq:div_d}$$

with the numerical fluxes

$$F_{i+1/2+} = \mathcal{F}(X_i, X_{i+1}, z_{b,i}, z_{b,i+1}) + \mathcal{S}_{i+1/2+} \quad (148) \text{eq:f+}$$

$$F_{i+1/2-} = \mathcal{F}(X_i, X_{i+1}, z_{b,i}, z_{b,i+1}) + \mathcal{S}_{i+1/2-}. \quad (149) \text{eq:f-}$$

$\mathcal{F}$  is a numerical flux for the conservative part of the system,  $\mathcal{S}$  is a convenient discretization of the topography source term.

Since the first two lines of (136) correspond to the classical Saint-Venant system, the numerical fluxes

$$F_{i+1/2\pm} = \begin{pmatrix} F_{h,i+1/2} \\ F_{q_x,i+1/2\pm} \\ F_{q_z,i+1/2} \end{pmatrix}, \quad (150) \text{eq:flux_hyp}$$

can be constructed using any numerical solver for the Saint-Venant system. More precisely for  $F_{h,i+1/2}, F_{q_x,i+1/2\pm}$  we adopt numerical fluxes suitable for the Saint-Venant system with topography [23, 10, 26]. Notice that from the definition (138), since only the second component of  $S(X)$  is non zero, only  $F_{q_x}$  has two interface values under the form  $F_{q_x,i+1/2\pm}$ . For the definition of  $F_{q_z,i+1/2}$ , the formula (see [5])

$$F_{q_z,i+1/2} = F_{h,i+1/2} w_{i+1/2}, \quad (151) \text{eq:fluxHw}$$

with

$$w_{i+1/2} = \begin{cases} w_i & \text{if } F_{h,i+1/2} \geq 0 \\ w_{i+1} & \text{if } F_{h,i+1/2} < 0 \end{cases} \quad (152) \text{eq:def_upwind_w}$$

can be used. The fluxes  $F_{\hat{p},i}$  are defined similarly to (151),(152) but on the staggered grid by the following formula

$$F_{\hat{p},i} = \frac{F_{h,i+1/2} + F_{h,i-1/2}}{2} \hat{p}_i, \quad (153) \text{eq:flux_p}$$

with

$$\hat{p}_i = \begin{cases} \hat{p}_{i-1/2} & \text{if } \frac{F_{h,i+1/2} + F_{h,i-1/2}}{2} \geq 0 \\ \hat{p}_{i+1/2} & \text{if } \frac{F_{h,i+1/2} + F_{h,i-1/2}}{2} < 0. \end{cases}$$

Combining the finite volume approach for the hyperbolic part with a finite difference strategy for the parabolic part, the non-hydrostatic part  $R_i$  is defined by

$$R_i = \begin{pmatrix} 0 \\ \nabla_{sw,i}^\gamma p \end{pmatrix},$$

where the two components of  $\nabla_{sw,i}^\gamma p$  are defined (see (139)) by

$$\begin{aligned} \Delta x_i \nabla_{sw,i}^\gamma p|_1 &= h_i(p_{i+1/2} - p_{i-1/2}) + \frac{p_{i+1/2}}{2}(\zeta_{i+1} - \zeta_i) \\ &\quad + \frac{p_{i-1/2}}{2}(\zeta_i - \zeta_{i-1}), \end{aligned} \quad (154) \quad \boxed{\text{eq:pnh\_sd1}}$$

$$\Delta x_i \nabla_{sw,i}^\gamma p|_2 = -\frac{\gamma}{2}(\Delta x_{i+1/2} p_{i+1/2} + \Delta x_{i-1/2} p_{i-1/2}), \quad (155) \quad \boxed{\text{eq:pnh\_sd2}}$$

with  $\zeta_i = h_i + \frac{\gamma^2}{2} z_{b,i}$ . And in (147),  $\text{div}_{sw,i+1/2}^\gamma(\mathbf{u})$  is defined by

$$\begin{aligned} \Delta x_{i+1/2} \text{div}_{sw,i+1/2}^\gamma(\mathbf{u}) &= \frac{h_{i+1} + h_i}{2}(u_{i+1} - u_i) - \frac{u_i + u_{i+1}}{2}(z_{b,i+1} - z_{b,i}) \\ &\quad + \frac{\gamma \Delta x_{i+1/2}}{2}(w_{i+1} + w_i) \\ &= (hu)_{i+1} - (hu)_i - \frac{u_i + u_{i+1}}{2}(\zeta_{i+1} - \zeta_i) \\ &\quad + \frac{\gamma \Delta x_{i+1/2}}{2}(w_{i+1} + w_i). \end{aligned} \quad (156) \quad \boxed{\text{eq:div\_ip12}}$$

Notice that in the definitions (154)-(155) and in the sequel, the quantity  $p$  means  $\{p_j\}$ . Likewise in Eq. (156) and in the sequel,  $\mathbf{u}$  means  $\{\mathbf{u}_j\}$  for  $1 \leq j \leq I$ .

### 4.3 Wet-dry interface

?{sec:wet\\_dry}? The method presented above supposes that the water depth does not vanish since the resolution of (145) requires dividing the shallow water gradient and divergence operators by  $h$ . We use a strategy similar to [2, paragraph 5.2] that can be viewed as a Dirichlet condition on the dry zone of the domain, such that the non-hydrostatic pressure is solved only on the wet domain.

In practice, we introduce a small parameter  $\theta \ll 1$  and the definitions (154)-(155) become

$$\begin{aligned}
\Delta x_i \nabla_{sw,i}^\gamma p \Big|_1 &= \mathbf{1}_{h_i \geq h_\theta} \left( h_i (p_{i+1/2} - p_{i-1/2}) + \frac{p_{i+1/2}}{2} (\zeta_{i+1} - \zeta_i) + \frac{p_{i-1/2}}{2} (\zeta_i - \zeta_{i-1}) \right), \\
\Delta x_i \nabla_{sw,i}^\gamma p \Big|_2 &= -\frac{\gamma \mathbf{1}_{h_i \geq h_\theta}}{2} \left( \Delta x_{i+1/2} p_{i+1/2} + \Delta x_{i-1/2} p_{i-1/2} \right), \\
\Delta x_{i+1/2} \operatorname{div}_{sw,i+1/2}^\gamma (\mathbf{u}) &= \mathbf{1}_{h_i \geq h_\theta} \left( (hu)_{i+1} - (hu)_i - \frac{u_i + u_{i+1}}{2} (\zeta_{i+1} - \zeta_i) \right. \\
&\quad \left. + \frac{\gamma \Delta x_{i+1/2}}{2} (w_{i+1} + w_i) \right),
\end{aligned}$$

with  $h_\theta = \max(h, \theta)$ .

#### 4.4 Stability of the scheme

(subsec:stability) Using the definitions (144),(145),(146),(147),(154) and (155), the fully discrete scheme for the system (136)-(137) writes

$$\begin{cases} X_i^{n+1/2} = X_i^n - \frac{\Delta t^n}{\Delta x_i} \left( F_{i+1/2-}^n - F_{i-1/2+}^n \right), \\ (h\hat{p})_{i+1/2}^{n+1/2} = (h\hat{p})_{i+1/2}^n - \frac{\Delta t^n}{\Delta x_{i+1/2}} \left( F_{\hat{p},i+1}^n - F_{\hat{p},i}^n \right), \end{cases} \quad (157) \quad \boxed{\text{eq:eq1_1d_fd}}$$

$$\begin{cases} p_{i+1/2}^{n+1/2,k+1} = p_{i+1/2}^{n+1/2,k} - \frac{\Delta t^n}{\varepsilon K h^{n+1}} \operatorname{div}_{sw,i+1/2}^\gamma \mathbf{u}^{n+1/2,k}, \\ \mathbf{u}_i^{n+1/2,k+1} = \mathbf{u}_i^{n+1/2,k} - \frac{\Delta t^n}{K h^{n+1}} \nabla_{sw,i}^\gamma p^{n+1/2,k+1}. \end{cases} \quad (158) \quad \boxed{\text{eq:eq2_1d_fd}}$$

The first equation of (157) gives a finite volume scheme for the Saint-Venant system. The choice of numerical fluxes  $F_{i+1/2\pm}$  (see [10]) coupled with a numerical treatment of the topography source term e.g. using the hydrostatic reconstruction [3] gives a numerical resolution of the Saint-Venant system endowed with strong stability properties [4] that are recalled in Propositions 7 and 8. In Eqs. (157)-(158),  $\Delta t^n$  satisfies a CFL condition having the form

$$\Delta t^n = \max_{i \in I} \frac{\Delta x_i}{|V_i^n|}, \quad (159) \quad \boxed{\text{eq:cfl_sv}}$$

where  $V_i^n$  is related to the eigenvalues of the Saint-Venant system, see [10]. Since the expression of the numerical fluxes (Rusanov, HLL, kinetic solver. . .) is not precised we are not able to give the exact expression of the CFL condition. In order to study the discrete energy balance induced by the numerical scheme (157)-(158), we define a discrete version of (142) under the form

$$\begin{aligned}
\frac{\Delta t^n}{\Delta x_i} (h\hat{\zeta})_i^{n+1/2,k+1} &= \frac{\Delta x_i}{\Delta t^n} (h\hat{\zeta})_i^{n+1/2,k} - \frac{1}{K} \left( \hat{\zeta}_{i+1/2}^{n+1/2,k} F_{h,i+1/2} - \hat{\zeta}_{i-1/2}^{n+1/2,k} F_{h,i-1/2} \right) \\
&\quad + \frac{\Delta x_{i+1/2} h_{i+1/2}^{n+1}}{2K} \operatorname{div}_{sw,i+1/2}^\gamma (\mathbf{u}^{n+1/2,k}) + \frac{\Delta x_{i-1/2} h_{i-1/2}^{n+1}}{2K} \operatorname{div}_{sw,i-1/2}^\gamma (\mathbf{u}^{n+1/2,k}),
\end{aligned} \quad (160) \quad \boxed{\text{eq:tilde_zeta_1d_d}}$$

where  $\hat{\zeta}_{i+1/2}$  is defined by

$$\hat{\zeta}_{i+1/2} = \begin{cases} \hat{\zeta}_i & \text{if } F_{h,i+1/2} \geq 0 \\ \hat{\zeta}_{i+1} & \text{if } F_{h,i+1/2} < 0. \end{cases}$$

Now we focus on the stability condition for the resolution of (145) or equivalently (128). Using the definitions (154) and (155), we obtain the discrete version of the operator

$$\Delta_{sw}^\gamma p = \text{div}_{sw}^\gamma \left( \frac{1}{h} \nabla_{sw}^\gamma p \right),$$

with  $D_{i+1/2} p = -\Delta x_{i+1/2} \Delta_{sw,i+1/2}^\gamma p$  and

$$\begin{aligned} D_{i+1/2} p = & -\frac{h_{i+1}}{\Delta x_{i+1}} (p_{i+3/2} - p_{i+1/2}) + \frac{h_i}{\Delta x_i} (p_{i+1/2} - p_{i-1/2}) \\ & - \frac{p_{i+3/2}}{2\Delta x_{i+1}} (\zeta_{i+2} - 2\zeta_{i+1} + \zeta_i) - \frac{\Delta x_i - \Delta x_{i+1}}{\Delta x_{i+1} \Delta x_i} p_{i+1/2} (\zeta_{i+1} - \zeta_i) \\ & - \frac{p_{i-1/2}}{2\Delta x_i} (\zeta_{i+1} - 2\zeta_i + \zeta_{i-1}) \\ & + \frac{p_{i+3/2}}{4h_{i+1}\Delta x_{i+1}} (\zeta_{i+2} - \zeta_{i+1}) (\zeta_{i+1} - \zeta_i) \\ & + \frac{p_{i+1/2}}{4} \left( \frac{1}{h_{i+1}\Delta x_{i+1}} + \frac{1}{h_i\Delta x_i} \right) (\zeta_{i+1} - \zeta_i)^2 \\ & + \frac{p_{i-1/2}}{4h_i\Delta x_i} (\zeta_{i+1} - \zeta_i) (\zeta_i - \zeta_{i-1}) \\ & + \frac{\gamma^2 \Delta x_{i+1/2}}{4} \left( \frac{\Delta x_{i+3/2} p_{i+3/2} + \Delta x_{i+1/2} p_{i+1/2}}{\Delta x_{i+1} h_{i+1}} \right. \\ & \left. + \frac{\Delta x_{i+1/2} p_{i+1/2} + \Delta x_{i-1/2} p_{i-1/2}}{\Delta x_i h_i} \right). \end{aligned} \quad (161) \quad \boxed{\text{eq:laplac_d}}$$

Using the expression (161), we are now able to precise the CFL type stability condition for the discretized version of Eq. (128) that writes

$$\begin{aligned} 2 - \frac{(\Delta t^n)^2}{\varepsilon K^2 h_{i+1/2} \Delta x_{i+1/2}} \left( \frac{h_{i+1}}{\Delta x_{i+1}} + \frac{h_i}{\Delta x_i} + \frac{1}{4} \left( \frac{1}{h_{i+1}\Delta x_{i+1}} + \frac{1}{h_i\Delta x_i} \right) (\zeta_{i+1} - \zeta_i)^2 \right. \\ \left. - \frac{\Delta x_i - \Delta x_{i+1}}{\Delta x_{i+1} \Delta x_i} (\zeta_{i+1} - \zeta_i) + \frac{\gamma^2 \Delta x_{i+1/2}^2}{4} \left( \frac{1}{h_{i+1}\Delta x_{i+1}} + \frac{1}{h_i\Delta x_i} \right) \right) \geq 0, \end{aligned}$$

that is fulfilled for

$$\begin{aligned} K^2 \geq \frac{(\Delta t^n)^2}{2\varepsilon h_{i+1/2} \Delta x_{i+1/2}} \left( \frac{h_{i+1}}{\Delta x_{i+1}} + \frac{h_i}{\Delta x_i} + \frac{1}{4} \left( \frac{1}{h_{i+1}\Delta x_{i+1}} + \frac{1}{h_i\Delta x_i} \right) (\zeta_{i+1} - \zeta_i)^2 \right. \\ \left. + \frac{|\Delta x_i - \Delta x_{i+1}|}{\Delta x_{i+1} \Delta x_i} |\zeta_{i+1} - \zeta_i| + \frac{\gamma \Delta x_{i+1/2}^2}{4} \left( \frac{1}{h_{i+1}\Delta x_{i+1}} + \frac{1}{h_i\Delta x_i} \right) \right). \end{aligned} \quad (162) \quad \boxed{\text{eq:CFL_iter}}$$



And the condition (162) is satisfied when

$$K^2 \geq \frac{(\Delta t^n)^2}{2\varepsilon h_{\min} \Delta x_{\min}^2} \left( 2h_{\max} + \frac{2}{h_{\min}} (\delta\zeta)_{\max}^2 + (\delta\zeta)_{\max} + \frac{\gamma^2 \Delta x_{\max}^2}{2h_{\min}} \right),$$

with  $r_{\min} = \min_{1 \leq i \leq I} r_i$ ,  $r_{\max} = \max_{1 \leq i \leq I} r_i$  for  $r = h, \Delta x$  and  $\delta\zeta_{\max} = \max_{1 \leq i \leq I-1} |\zeta_{i+1} - \zeta_i|$ .

The fully discrete scheme (157),(158) satisfies the following stability properties.

**Proposition 7** *Assuming a suitable CFL condition (159) adapted to the chosen numerical fluxes (150) for the hyperbolic part, the scheme obtained by coupling the semi-discretizations (144),(145) and (146),(147)*

- (i) *preserves the nonnegativity of the water depth i.e.  $h_i^n \geq 0, \forall i, \forall n$ ,*
- (ii) *preserves the steady state of the lake at rest,*
- (iii) *is consistent with the model (136)-(137).*

`<prop:stab_prop>`

Let us consider that, under a suitable CFL condition associated with the time discretization (144) and the chosen numerical fluxes  $F_{h,i\pm 1/2}$  and  $F_{q_x,i\pm 1/2}$  in (150), the numerical approximation of the Saint-Venant part of Eq. (136) allows to obtain a discrete entropy equality under the form

$$\Delta x_i (E_i^{sv})^{n+1/2} = \Delta x_i (E_i^{sv})^n - \Delta t^n (\mathcal{G}_{i+1/2}^n - \mathcal{G}_{i-1/2}^n) + \mathcal{D}_i^n, \quad (163) \quad \boxed{\text{eq:entro\_SV}}$$

with  $E_i^{sv} = \frac{h_i}{2} u_i^2 + \frac{g}{2} (\eta_i^2 - z_{b,i}^2)$  and where  $\mathcal{G}_{i\pm 1/2}^n$  are numerical fluxes.  $\mathcal{D}_i^n$  is a nonpositive term and contains typically two different contributions: the numerical dissipation coming from the upwinding in the space discretization and the error due to the explicit time scheme.

Then assuming (163), we now prove that the numerical scheme (157),(158) satisfies a discrete entropy equality.

**Proposition 8** *Assuming (163), the scheme (144),(145),(146),(147) satisfies the following discrete entropy equality*

$$\Delta x_i \bar{E}_i^{n+1} = \Delta x_i \bar{E}_i^n - \Delta t^n (\bar{\mathcal{G}}_{i+1/2}^n - \bar{\mathcal{G}}_{i-1/2}^n) + \bar{\mathcal{D}}_i^n, \quad (164) \quad \boxed{\text{eq:entro\_discrete}}$$

where  $\bar{E}_i = E_i^{sv} + \frac{h_i}{2} w_i^2 + g(h\hat{\zeta})_i + \frac{\varepsilon}{2} h_i \hat{p}_i^2$  and

$$\begin{aligned} \bar{\mathcal{G}}_{i+1/2}^n &= \mathcal{G}_{i+1/2}^n + F_{h,i+1/2}^n \frac{(w_{i+1/2}^n)^2}{2} + \varepsilon F_{h,i+1/2}^n \frac{(\hat{p}_{i+1/2}^n)^2}{2} \\ &\quad + \frac{1}{K} \sum_{k=1}^K (h_{i+1/2}^{n+1} u_{i+1/2}^{n+1/2,k} p_{i+1/2}^{n+1/2,k+1} + \hat{\zeta}_{i+1/2}^{n+1/2,k} F_{h,i+1/2}), \\ \Delta x_i (\bar{h}_i \hat{p}_i^2)^{n+1} &= \frac{1}{2K} \sum_{k=1}^K \left( \Delta x_{i+1/2} \frac{h_{i+1/2}^{n+1}}{2} (\hat{p}_{i+1/2}^{n+1/2,k+1})^2 + \Delta x_{i-1/2} \frac{h_{i-1/2}^{n+1}}{2} (\hat{p}_{i-1/2}^{n+1/2,k+1})^2 \right), \end{aligned}$$

with

$$\begin{aligned}
\Delta x_i \bar{\mathcal{D}}_i^n &= \Delta x_i \mathcal{D}_i^n + \Delta t^n \left( [F_{h,i+1/2}^n]_- (w_{i+1}^n - w_i^n)^2 - [F_{h,i-1/2}^n]_+ (w_i^n - w_{i-1}^n)^2 \right) \\
&+ \frac{\Delta t^n}{2} \left( \frac{F_{h,i+1}^n}{2} (\hat{p}_{i+1}^n - \hat{p}_{i+1/2}^n)^2 - \frac{F_{h,i}^n}{2} (\hat{p}_i^n - \hat{p}_{i+1/2}^n)^2 \right) \\
&+ \frac{\Delta t^n}{2} \left( \frac{F_{h,i}^n}{2} (\hat{p}_i^n - \hat{p}_{i-1/2}^n)^2 - \frac{F_{h,i-1}^n}{2} (\hat{p}_{i-1}^n - \hat{p}_{i-1/2}^n)^2 \right) \\
&+ \frac{\Delta x_i h_i^{n+1}}{2} (w_i^{n+1/2} - w_i^n)^2 + \frac{\Delta x_{i+1/2} h_{i+1/2}^{n+1}}{2} (\hat{p}_{i+1/2}^{n+1/2} - \hat{p}_{i+1/2}^n)^2 \\
&+ \frac{\Delta x_{i-1/2} h_{i-1/2}^{n+1}}{2} (\hat{p}_{i-1/2}^{n+1/2} - \hat{p}_{i-1/2}^n)^2 \\
&- \sum_{k=1}^K \frac{\Delta x_{i+1/2} h_{i+1/2}^{n+1}}{2} \left( \varepsilon (\hat{p}_{i+1/2}^{n+1/2,k+1} - \hat{p}_{i+1/2}^{n+1/2,k})^2 - |\mathbf{u}_{i+1/2}^{n+1/2,k+1} - \mathbf{u}_{i+1/2}^{n+1/2,k}|^2 \right) \\
&- \sum_{k=1}^K \frac{\Delta x_{i-1/2} h_{i-1/2}^{n+1}}{2} \left( \varepsilon (\hat{p}_{i-1/2}^{n+1/2,k+1} - \hat{p}_{i-1/2}^{n+1/2,k})^2 - |\mathbf{u}_{i-1/2}^{n+1/2,k+1} - \mathbf{u}_{i-1/2}^{n+1/2,k}|^2 \right).
\end{aligned}$$

(prop:entro\_prop)

*Remark 8* When considering the semi-discrete in space scheme detailed in Section 4.2, a semi-discrete in space version of (164) holds where all the non-negative terms in the expression of  $\bar{\mathcal{D}}_i^n$  corresponding to time discretisation errors vanish.

**Proof (Proof of Prop. 7)** (i) The statement that  $\mathcal{F}$  preserves the nonnegativity of the water depth means exactly that

$$F_h(h_i = 0, u_i, h_{i+1}, u_{i+1}) - F_h(h_{i-1}, u_{i-1}, h_i = 0, u_i) \leq 0,$$

for all choices of the other arguments. From (144),(146),(148) and (149), we need to check that, with obvious notations

$$F_h(X_{i+1/2-}^n, X_{i+1/2+}^n) - F_h(X_{i-1/2-}^n, X_{i-1/2+}^n) \leq 0,$$

whenever  $h_i^n = 0$ . And this property holds typically when the hydrostatic reconstruction (HR) is used to approximate the topography source term since for the HR technique  $h_i = 0$  implies  $h_{i+1/2-} = h_{i-1/2+} = 0$ , see [3].

(ii) When  $u_i^n = 0$  for all  $i$ , the properties of the hydrostatic reconstruction technique ensure  $F_{i+1/2-}^n = F_{i-1/2+}^n$  in (144),(146) and  $F_{p,i+1-}^n = F_{p,i+}^n$  in (147). Moreover since  $u_i^n = 0 \forall i$  we have  $R_i = 0$  in (146) and  $\text{div}_{sw,i+1/2}^y(\{\mathbf{u}\}) = 0$  in (147). Therefore  $\forall i$

$$X_i^{n+1} = X_i^n, \quad \text{and} \quad p_{i+1/2}^{n+1} = p_{i+1/2}^n,$$

proving that the scheme is well-balanced.

(iii) The discretization (144),(145) is an explicit first order time scheme. The numerical fluxes defined by (148),(149) and (153) are a consistent discretization of the

hyperbolic part of the system (136),(137) without topography. Likewise, the hydrostatic reconstruction applied to the fluxes (150),(153) gives a consistent discretization of the system (136),(137) with topography and the discretizations (154),(155) being obviously consistent with the dispersive part, this proves the result.  $\square$

**Proof (Proof of Prop. 8)** Since we have assumed that the kinetic energy of the Saint-Venant part of Eq. (136) satisfies (163), this means that the first two components of the first equation of (157) multiplied respectively by  $gh_i^n - (u_i^n)^2/2$  and  $u_i^n$  give Eq. (163). It remains to consider the contributions to the energy balance of the last two components of (157) and of Eq. (158).

First let us multiply the third component of the first equation of (157) by  $w_i^n$ , then we get

$$\begin{aligned} & \frac{h_i^{n+1}}{2} (w_i^{n+1/2})^2 - \frac{h_i^n}{2} (w_i^n)^2 + \frac{\Delta t^n}{\Delta x_i} \left( F_{h,i+1/2}^n \frac{(w_{i+1/2}^n)^2}{2} - F_{h,i-1/2}^n \frac{(w_{i-1/2}^n)^2}{2} \right) = \\ & \frac{\Delta t^n}{\Delta x_i} \left( [F_{h,i+1/2}^n]_- (w_{i+1}^n - w_i^n)^2 - [F_{h,i-1/2}^n]_+ (w_i^n - w_{i-1}^n)^2 \right) + \frac{h_i^{n+1}}{2} (w_i^{n+1/2} - w_i^n)^2, \end{aligned}$$

with the notations  $[a]_+ = \max(a, 0)$ ,  $[a]_- = \min(a, 0)$ ,  $a = [a]_+ + [a]_-$  and  $w_{i+1/2}^n$  is defined by (152). Then we multiply the last component of Eq. (157) by  $p_{i+1/2}^n + \frac{g}{2} h_{i+1/2}^n$  leading to

$$\begin{aligned} & \frac{h_{i+1/2}^{n+1}}{2} (\hat{p}_{i+1/2}^{n+1/2})^2 - \frac{h_{i+1/2}^n}{2} (\hat{p}_{i+1/2}^n)^2 + \frac{\Delta t^n}{\Delta x_{i+1/2}} \left( F_{h,i+1}^n \frac{(\hat{p}_{i+1}^n)^2}{2} - F_{h,i}^n \frac{(\hat{p}_i^n)^2}{2} \right) \\ & = \frac{\Delta t^n}{\Delta x_{i+1/2}} \left( \frac{F_{h,i+1}^n}{2} (\hat{p}_{i+1}^n - \hat{p}_{i+1/2}^n)^2 - \frac{F_{h,i}^n}{2} (\hat{p}_i^n - \hat{p}_{i+1/2}^n)^2 \right) \\ & \quad + \frac{h_{i+1/2}^{n+1}}{2} (\hat{p}_{i+1/2}^{n+1/2} - \hat{p}_{i+1/2}^n)^2, \end{aligned}$$

with  $F_{h,i+1} = (F_{h,i+3/2} + F_{h,i+1/2})/2$ . Thanks to the definition (4.2), the two quantities

$$\frac{F_{h,i+1}^n}{2} (\hat{p}_{i+1}^n - \hat{p}_{i+1/2}^n)^2, \quad \text{and} \quad -\frac{F_{h,i}^n}{2} (\hat{p}_i^n - \hat{p}_{i+1/2}^n)^2,$$

are always non-positive.

Second, we multiply the equations (158) respectively by  $\hat{p}_{i+1/2}^{n+1/2,k+1}$  and  $\mathbf{u}_i^{n+1/2,k}$  and sum the obtained relations for  $k = 1, \dots, K$ . Precisely, starting from the definitions (154),(155), we rewrite  $\nabla_{sw,i}^\gamma p$  under the form

$$\nabla_{sw,i}^\gamma p = \nabla_{sw,i+1/2}^\gamma p + \nabla_{sw,i-1/2}^\gamma p,$$

with

$$\nabla_{sw,i+1/2}^\gamma p = \left| \begin{array}{l} h_i(p_{i+1/2} - p_i) + \frac{p_{i+1/2}}{2} (\zeta_{i+1} - \zeta_i) \\ -\frac{\gamma}{2} \Delta x_{i+1/2} p_{i+1/2} \end{array} \right.$$

and we obtain a discrete version of the duality relation (140) under the form

$$\begin{aligned} \Delta x_i \nabla_{sw,i+1/2-}^\gamma P^{n+1/2,k+1} \cdot \mathbf{u}_i^{n+1/2,k} &= e_{i+1/2-}^{n+1/2,k+1/2} \\ &\quad - \frac{\Delta x_{i+1/2}}{2} \operatorname{div}_{sw,i+1/2}^\gamma (\mathbf{u}^{n+1/2,k}) p_{i+1/2}^{n+1/2,k+1}, \end{aligned} \quad (165) \quad \boxed{\text{eq:duality\_d}}$$

with  $e_{i+1/2-}^{n+1/2,k+1/2}$  defined by

$$\begin{aligned} e_{i+1/2-}^{n+1/2,k+1/2} &= h_{i+1/2}^{n+1} u_{i+1/2}^{n+1/2,k} p_{i+1/2}^{n+1/2,k+1} - h_i^{n+1} u_i^{n+1/2,k} p_i^{n+1/2,k+1} \\ &\quad + p_{i+1/2}^{n+1/2,k+1} (\zeta_{i+1} - \zeta_i) \frac{u_i^{n+1/2,k} - u_{i+1}^{n+1/2,k}}{2} \\ &\quad - \frac{\gamma}{4} \Delta x_{i+1/2} P_{i+1/2}^{n+1/2,k+1} (w_i^{n+1/2,k} - w_{i+1}^{n+1/2,k}). \end{aligned}$$

Notice that in the expression of  $e_{i+1/2-}^{n+1/2,k+1/2}$  the last two terms are second order and  $e_{i+1/2+}^{n+1/2,k+1/2} + e_{i+1/2-}^{n+1/2,k+1/2} = h_{i+1}^{n+1} u_{i+1}^{n+1/2,k} p_{i+1}^{n+1/2,k+1} - h_i^{n+1} u_i^{n+1/2,k} p_i^{n+1/2,k+1}$ .

The duality relation (165) has been written for the variable  $p_{i\pm 1/2}^{n+1/2,k+1}$  but the last two terms in (165) should be a discrete version of the r.h.s. of Eq. (112) i.e. of the quantity  $\hat{p} \operatorname{div}_{sw}^\gamma (\mathbf{u})$ . And since  $\hat{p} = p + gh/2$  the reminder is (for interfaces  $i \pm 1/2$ )

$$\frac{g}{2} h_{i+1/2}^{n+1} \operatorname{div}_{sw,i+1/2}^\gamma (\mathbf{u}^{n+1/2,k}) + \frac{g}{2} h_{i-1/2}^{n+1} \operatorname{div}_{sw,i-1/2}^\gamma (\mathbf{u}^{n+1/2,k}),$$

corresponding to the right hand side of (160) multiplied by  $g$ .

For the errors coming from the time discretization of Eqs. (158), we have

$$\begin{aligned} ((hp)_{i+1/2}^{n+1/2,k+1} - (hp)_{i+1/2}^{n+1/2,k}) \hat{p}_{i+1/2}^{n+1/2,k+1} &= ((h\hat{p})_{i+1/2}^{n+1/2,k+1} - (h\hat{p})_{i+1/2}^{n+1/2,k}) \hat{p}_{i+1/2}^{n+1/2,k+1} \\ &= \frac{1}{2} (h\hat{p}^2)_{i+1/2}^{n+1/2,k+1} - \frac{1}{2} (h\hat{p}^2)_{i+1/2}^{n+1/2,k} \\ &\quad + \frac{h_{i+1/2}^{n+1}}{2} (\hat{p}_{i+1/2}^{n+1/2,k+1} - \hat{p}_{i+1/2}^{n+1/2,k})^2, \\ ((h\mathbf{u})_i^{n+1/2,k+1} - (h\mathbf{u})_i^{n+1/2,k}) \cdot \mathbf{u}_i^{n+1/2,k} &= \frac{h_i^{n+1}}{2} |\mathbf{u}_i^{n+1/2,k+1}|^2 - \frac{h_i^{n+1}}{2} |\mathbf{u}_i^{n+1/2,k}|^2 \\ &\quad - \frac{h_{i+1/2}^{n+1}}{2} |\mathbf{u}_{i+1/2}^{n+1/2,k+1} - \mathbf{u}_{i+1/2}^{n+1/2,k}|^2. \end{aligned}$$

Summing the previous relations for  $k = 1, \dots, K$  and adding the result to the other contributions gives the corresponding expressions appearing in relation (164). This ends the proof.  $\square$

*Remark 9* For the discretization of the model (132)-(135), we have presented a first order scheme in space and time. Second order extensions (in space and time) can be proposed, following [2].

## 4.5 Simulation results

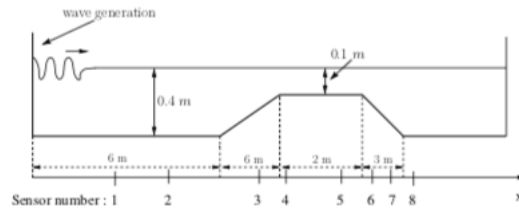
?(subsec:num)?

In this paragraph, only few numerical examples are presented. A more complete validation of the numerical procedure will be presented in a companion paper. Notice that in the 1d case, we mainly validate the numerical scheme but the reduction of the computation costs will be more significant in a two-dimensional setting.

### 4.5.1 Dingemans experiments

(subsubsec:dingemans)

The experiments carried out by Dingemans [18] at Delft Hydraulics deal with the wave propagation over uneven bottoms. A wave generator produces a small amplitude wave (0.02 m) at the left boundary of a basin with vertical shores. A vertical shore closes the basin at the right boundary, due to the considered time window, the measurements are not perturbed by the reflected wave on the right boundary. At rest, the water depth in the channel varies from 0.4 m to 0.1 m, see Fig. 3. Eight sensors recording the free surface elevation are located at abscissa 2m, 4m, 10.5m, 12.5m, 13.5m, 14.5m, 15.7m and 17.3m.

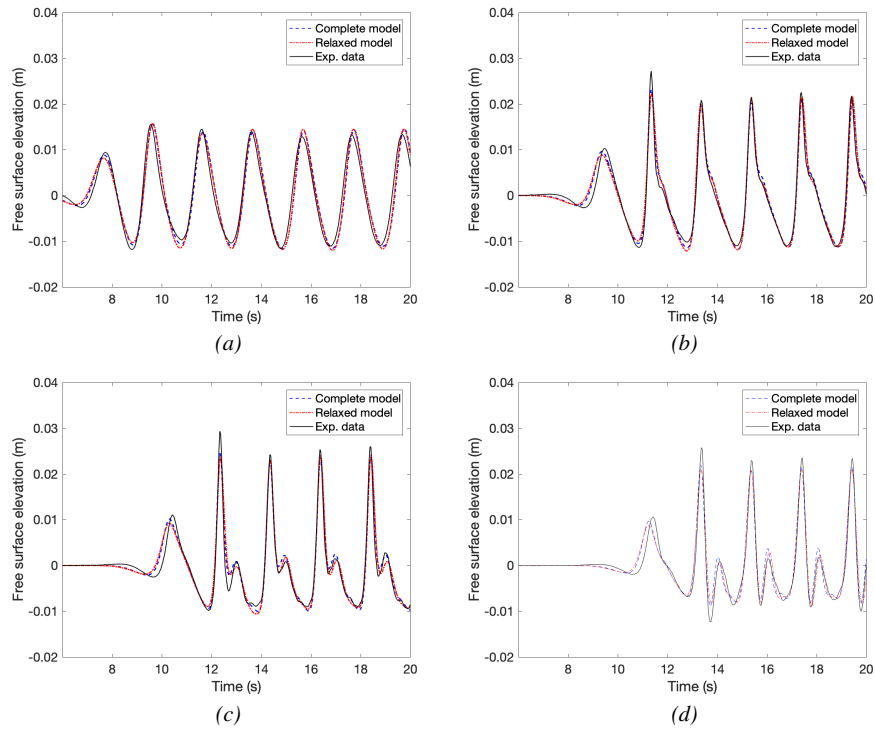


(fig:dingemans)

**Fig. 3** Channel profile for the experiments and location of the sensors.

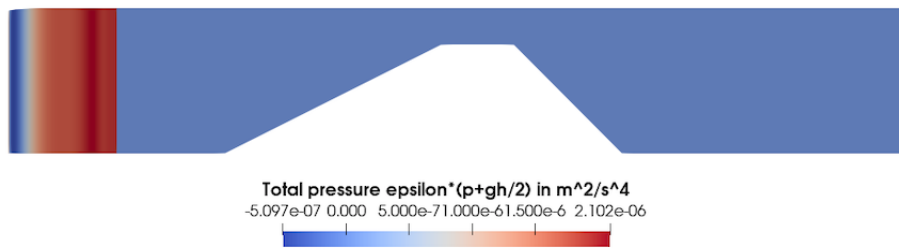
For  $\gamma = \sqrt{3}$ , we compare the simulation results obtained with the two numerical schemes (the one proposed in [2] and the one proposed in this paper with  $\varepsilon = 1/c^2 = 10^{-4} \text{ m}^{-2} \cdot \text{s}^2$ ). The results obtained with a uniform mesh of 1600 nodes are depicted over Fig. 4 where the computed and measured free surface elevations at four points are presented. Notice that for  $\varepsilon = 10^{-7} \text{ m}^{-2} \cdot \text{s}^2$  that is the most physical choice, the simulations of the complete and relaxed model cannot be distinguished.

For the test case depicted in paragraph 4.5.1, the basin is at rest at the initial instant and we give at time  $t = 0.01 \text{ s}$ , the value of the quantity  $\varepsilon(p + gh/2)$  representing the pseudo-compressible effects. It appears over Fig. 5 that at time  $t = 0.1 \text{ s}$ , whereas the free surface has just begun to deform at the boundary where the wave is generated, the acoustic-type waves have already propagated in the basin.



**Fig. 4** Comparisons between the experimental data (solid line), the simulations of the dispersive model with the model presented in [2] (blue dashed line) and the relaxed model presented in this paper (red dashed-dotted line) with  $\epsilon = 10^{-4} \text{ m}^{-2} \cdot \text{s}^2$ . Figs. (a), (b), (c) and (d) respectively correspond to the results for the sensors 3, 4, 5 and 6.

<fig:dingemans\_res>



**Fig. 5** Variations of the quantity  $x \mapsto \epsilon(p + gh/2)$  with  $\epsilon = 10^{-7} \text{ m}^{-2} \cdot \text{s}^2$  at time  $t = 0.01 \text{ s}$ .

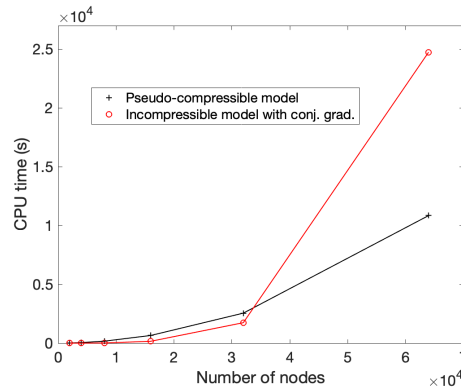
<fig:dingemans\_acc>

### 4.5.2 Comparison of the computational costs

For the simulation results given in paragraph 4.5.1, we compare the computational costs of the numerical schemes with and without pseudo-compressibility effects. More precisely, we compare the CPU time necessary to simulate the test case presented in paragraph 4.5.1 with the method proposed in [2] – corresponding to an incompressible model and requiring to solve the elliptic equation (127) – and the proposed explicit in time scheme (122)-(123) with the pseudo-compressible effects.

The advantages of the model and numerical strategy presented in this paper are significant for 2d problems with a large number of nodes but can hardly be highlighted in the 1d case where the elliptic operator to inverse is a symmetric tridagonal matrix. Hence, in order to illustrate the interest of the proposed scheme, we have used a conjugate gradient technique mimicking what would be done to solve (127) in 2d for an unstructured mesh.

Figure 6 presents the CPU time required to perform the simulations of the Dinguemans experiment with several meshes namely with 2000, 4000, 8000, 16000 and 32000 nodes. It appears that when the number of nodes increases, the proposed explicit in time scheme is more efficient than the conjugate gradient algorithm (used here without preconditioning). Notice that the authors have not performed an exhaustive comparison between the costs of the conjugate gradient technique – for which several optimizations are possible – and the iterative and explicit time resolution scheme (122)-(123).



**Fig. 6** Computational costs necessary to simulate the Dinguemans experiment with several meshes.   
 (fig:cpu)

### Acknowledgements

The authors thank François Bouchut for his helpful and constructive discussions that greatly contributed to improve the final version of the paper.

## References

- [airy] [1] G.B. Airy, *Tides and waves*, Encycl. Metropolitana **5** (1845), 291–369.
- [aissiouene:hal-01632522] [2] N. Aissiouene, M.-O. Bristeau, E. Godlewski, A. Mangeney, C. Parés, and J. Sainte-Marie, *A two-dimensional method for a family of dispersive shallow water model*, working paper or preprint, <https://hal.archives-ouvertes.fr/hal-01632522>, May 2020.
- [bristeau] [3] 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 J. Sci. Comput. **25** (2004), no. 6, 2050–2065.
- [JSM\_entro] [4] E. Audusse, F. Bouchut, M.-O. Bristeau, and J. Sainte-Marie, *Kinetic entropy inequality and hydrostatic reconstruction scheme for the Saint-Venant system*, Math. Comp. **85** (2016), no. 302, 2815–2837. MR 3522971
- [bristeau4] [5] E. Audusse and M.-O. Bristeau, *Transport of pollutant in shallow water flows : A two time steps kinetic method*, ESAIM: M2AN **37** (2003), no. 2, 389–416.
- [bristeau] [6] ———, *A well-balanced positivity preserving second-order scheme for Shallow Water flows on unstructured meshes*, J. Comput. Phys. **206** (2005), no. 1, 311–333.
- [saint-venant] [7] A.-J.-C. Barré de Saint-Venant, *Théorie du mouvement non permanent des eaux avec applications aux crues des rivières et à l'introduction des marées dans leur lit*, C. R. Acad. Sci. Paris **73** (1871), 147–154.
- [bbm] [8] J.-L. Bona, T.-B. Benjamin, and J.-J. Mahony, *Model equations for long waves in nonlinear dispersive systems*, Philos. Trans. Royal Soc. London Series A **272** (1972), 47–78.
- [saut1] [9] J.L. Bona, M. Chen, and J.-C. Saut, *Boussinesq equations and other systems for small-amplitude long waves in nonlinear dispersive media: Part I. Derivation and linear theory*, J. Nonlinear Sci. **12** (2002), 283–318.
- [bouchut\_book] [10] F. Bouchut, *Nonlinear stability of finite volume methods for hyperbolic conservation laws and well-balanced schemes for sources*, Birkhäuser, 2004.
- [boyer:hal-00004420] [11] F. Boyer, *Trace theorems and spatial continuity properties for the solutions of the transport equation*, Differential and integral equations **18 no 8** (2005), 891–934.
- [JSM\_nhyd] [12] M.-O. Bristeau, A. Mangeney, J. Sainte-Marie, and N. Seguin, *An energy-consistent depth-averaged Euler system: Derivation and properties*, Discrete and Continuous Dynamical Systems - Series B **20** (2015), no. 4, 961–988.
- [camassa] [13] R. Camassa, D. D. Holm, and C. D. Levermore, *Long-time effects of bottom topography in shallow water*, Phys. D **98** (1996), no. 2-4, 258–286, Nonlinear phenomena in ocean dynamics (Los Alamos, NM, 1995). MR 1422281 (98a:76005)
- [chorin] [14] A. J. Chorin, *Numerical solution of the Navier-Stokes equations*, Math. Comp. **22** (1968), 745–762. MR 0242392 (39 #3723)
- [dedner] [15] A. Dedner, F. Kemm, D. Kroner, C.-D. Munz, T. Schnitzer, and M. Wesenberg, *Hyperbolic divergence cleaning for the MHD equations*, J. Comp. Phys. **172** (2002), no. 2, 645–673.
- [desjardins] [16] B. Desjardins, *A few remarks on ordinary differential equations*, Comm. Part. Diff. Eq. **21** (1996), no. 11-12.
- [diperna\_lions] [17] R.J. Di Perna and P.L. Lions, *Ordinary differential equations, transport theory and Sobolev spaces*, Invent. Math. **98** (1989), 511–547.
- [dingemans] [18] M.-W. Dingemans, *Wave propagation over uneven bottoms*, Advanced Series on Ocean Engineering - World Scientific, 1997.
- [duchene\_2019] [19] V. Duchêne, *Rigorous justification of the Favrie-Gavrilyuk approximation to the Serre-Green-Naghdi model*, Nonlinearity **32** (2019), no. 10, 3772–3797.
- [ESCALANTE2019385] [20] C. Escalante, M. Dumbser, and M.J. Castro, *An efficient hyperbolic relaxation system for dispersive non-hydrostatic water waves and its solution with high order discontinuous Galerkin schemes*, Journal of Computational Physics **394** (2019), 385 – 416.
- [?filippini:breaking?] [21] A.G. Filippini, M. Kazolea, and M. Ricchiuto, *A flexible genuinely nonlinear approach for nonlinear wave propagation, breaking and run-up*, J. Comp. Phys. **310** (2016), 381 – 417.
- [gerbeau] [22] J.-F. Gerbeau and B. Perthame, *Derivation of Viscous Saint-Venant System for Laminar Shallow Water; Numerical Validation*, Discrete Contin. Dyn. Syst. Ser. B **1** (2001), no. 1, 89–102.



- `godlewski_book` [23] E. Godlewski and P.-A. Raviart, *Numerical approximation of hyperbolic systems of conservation laws*, Applied Mathematical Sciences, vol. 118, Springer, New York, 1996.
- `green` [24] A.E. Green and P.M. Naghdi, *A derivation of equations for wave propagation in water of variable depth*, J. Fluid Mech. **78** (1976), 237–246.
- `GUERMOND2019108917` [25] J.-L. Guermond, B. Popov, E. Tovar, and C. Kees, *Robust explicit relaxation technique for solving the Green-Naghdi equations*, J. Comp. Phys. **399** (2019), 108917.
- `leveque_book` [26] R.-J. LeVeque, *Finite volume methods for hyperbolic problems*, Cambridge University Press, 2002.
- `LIONS1998833` [27] P.-L. Lions, *Sur les équations différentielles ordinaires et les équations de transport*, Comptes Rendus de l'Académie des Sciences - Series I - Mathematics **326** (1998), no. 7, 833 – 838.
- `nwogu` [28] O. Nwogu, *Alternative form of Boussinesq equations for nearshore wave propagation*, Journal of Waterway, Port, Coastal and Ocean Engineering, ASCE **119** (1993), no. 6, 618–638.
- `peregrine` [29] D.H. Peregrine, *Long waves on a beach*, J. Fluid Mech. **27** (1967), 815–827.
- `unesco` [30] Unesco, *Thermodynamic equation of seawater: Calculation and use of thermodynamic properties*, <http://unesdoc.unesco.org/images/0018/001881/188170e.pdf>, 2010.