



# Numerical approximation of the 3d hydrostatic Navier-Stokes system with free surface

Sébastien Allgeyer, Marie-Odile Bristeau, D. Froger, R. Hamouda, Anne Mangeney, J. Sainte-Marie, F. Souillé, M. Vallée

## ► To cite this version:

Sébastien Allgeyer, Marie-Odile Bristeau, D. Froger, R. Hamouda, Anne Mangeney, et al.. Numerical approximation of the 3d hydrostatic Navier-Stokes system with free surface. 2017. hal-01393147v3

**HAL Id: hal-01393147**

**<https://inria.hal.science/hal-01393147v3>**

Preprint submitted on 26 Sep 2017 (v3), last revised 19 Jul 2019 (v5)

**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.

# Numerical approximation of the 3d hydrostatic Navier-Stokes system with free surface.

S. Allgeyer<sup>\*</sup>, M.-O. Bristeau<sup>†,‡,§</sup>, D. Froger<sup>†,‡,§</sup>, R. Hamouda<sup>†,‡,§</sup>, A. Mangeney<sup>¶,†,‡,§</sup>,  
J. Sainte-Marie<sup>†,‡,§,||</sup>, F. Souill  <sup>†,‡,§</sup> and M. Vall  e<sup>¶</sup>

September 28, 2017

## Abstract

In this paper we propose a stable and robust strategy to approximate incompressible hydrostatic Euler and Navier-Stokes systems with free surface.

The idea is to use a Galerkin type approximation of the velocity field with piecewise constant basis functions in order to obtain an accurate description of the vertical profile of the horizontal velocity. Such a strategy has several advantages. It allows

- to rewrite the Navier-Stokes equations under the form of a system of conservation laws with source terms,
- the easy handling of the free surface, which does not require moving meshes,
- the possibility to take advantage of robust and accurate numerical techniques developed in extensive amount for Shallow Water type systems.

We show that the model admits a kinetic interpretation, and we use this result to formulate a robust finite volume scheme for its numerical approximation. All the aspects of the discrete scheme (fluxes, boundary conditions, . . .) are completely described and the stability properties of the proposed numerical scheme are discussed. We illustrate the capability of the model and of the discrete scheme with some numerical academic and realistic examples.

**Keywords** Free surface flows, Navier-Stokes equations, Euler system, Hydrostatic assumption, Kinetic description, Finite volumes.

## 1 Introduction

In this paper we present layer-averaged Euler and Navier-Stokes models for the numerical simulation of incompressible free surface flows over variable topographies. We are mainly interested in applications to geophysical water flows such as lakes, rivers, estuarine waters, hazardous flows in the context either of advection dominant flows or of wave propagation.

The simulation of these flows requires stable, accurate, conservative schemes able to sharply resolve stratified flows, to handle efficiently complex topographies and free surface deformations, and to capture robustly wet/dry fronts. In addition, the application to realistic three-dimensional problems demands efficient methods with respect to computational cost. The present work is aimed to build a simulation tool endowed with these properties.

---

<sup>\*</sup>Research School of Earth Sciences, Australian National University, Canberra, ACT, Australia

<sup>†</sup>Inria Paris, 2 rue Simone Iff, CS 42112, 75589 Paris Cedex 12, France

<sup>‡</sup>Sorbonne Universit  s UPMC Univ Paris 6, Lab. Jacques-Louis Lions, 4 Place Jussieu, 75252 Paris cedex 05, France

<sup>§</sup>CEREMA, 134 rue de Beauvais, F-60280 Margny-L  s-Compi  gne, France

<sup>¶</sup>Univ. Paris Diderot, Sorbonne Paris Cit  , Institut de Physique du Globe de Paris, Seismology Group, 1 rue Jussieu, Paris F-75005, France

<sup>||</sup>Corresponding author: Jacques.Sainte-Marie@inria.fr

Due to computational issues associated with the free surface Navier-Stokes or Euler equations, the simulations of geophysical flows are often carried out with shallow water type models of reduced complexity. Indeed, for vertically averaged models such as the Saint-Venant system [13], efficient and robust numerical techniques (relaxation schemes [20], kinetic schemes [41, 6], ...) are available and avoid to deal with moving meshes. In order to describe and simulate complex flows where the velocity field cannot be approximated by its vertical mean, multilayer models have been developed [4, 8, 9, 19, 28, 27, 40, 47, 21]. Unfortunately these models are physically relevant for non miscible fluids. In [31, 11, 10, 43, 24, 30], some authors have proposed a simpler and more general formulation for multilayer model with mass exchanges between the layers. The obtained model has the form of a conservation law with source terms and presents remarkable differences with respect to classical models for non miscible fluids. In our approach the layer partition is merely a discretization artefact, and it is not physical. Therefore, the internal layer boundaries do not necessarily correspond to isopycnic surfaces. A critical distinguishing feature of our model is that it allows fluid circulation between layers. This changes dramatically the properties of the model and its ability to describe flow configurations that are crucial for the foreseen applications, such as recirculation zones. Compared to previous works of some of the authors [11, 10, 24], this paper deals with the 3d case on unstructured meshes reinforcing the need of efficient numerical schemes.

The strong point of this paper is the layer-averaging, the kinetic interpretation and the numerical scheme for the Euler part of the 3d Navier-Stokes system. But in order to deal with real test cases, the rheology of the fluid has to be considered, therefore in paragraphs 2.2, 3.1, 3.2, 5.6 we propose energy-consistent, simplified and efficient – continuous and discrete – expressions for a Newtonian rheology in the Navier-Stokes system.

Most of the numerical models in the literature for environmental stratified flows use finite difference or finite element schemes solving the free surface Navier–Stokes equations. We refer in particular to [29, 36] and references therein for a partial review of these methods. Since the layer-averaged model has the form of a conservation law with source terms, we single out a finite volume scheme. Moreover, the kinetic interpretation of the continuous model leads to a kinetic solver endowed with strong stability properties (well-balancing, domain invariant, discrete entropy [12]). We give a complete description of all the ingredients of the numerical scheme (time scheme, fluxes, boundary conditions,...). Considering various analytical solutions and a real test case with measurements, we emphasize the accuracy of the discrete model.

The outline of the paper is as follows. In Section 2, we recall the incompressible Navier-Stokes equations, the boundary conditions and the hydrostatic approximation. The layer-averaged system obtained by a vertical discretization of the hydrostatic model is described in Section 3. The kinetic interpretation of the model is given in Section 4 allowing to derive a numerical scheme presented in Section 5. Numerical validations are shown in Section 6.

## 2 The Navier-Stokes system

We consider the classical incompressible Navier-Stokes system

$$\nabla \cdot \mathbf{U} = 0, \quad (1)$$

$$\frac{\partial \mathbf{U}}{\partial t} + \nabla \cdot (\mathbf{U} \otimes \mathbf{U}) = \nabla \cdot \boldsymbol{\Sigma} + \mathbf{g}, \quad (2)$$

where  $\mathbf{U}(t, x, y, z) = (u, v, w)^T$  is the velocity,  $\mathbf{u}(t, x, y, z) = (u, v)^T$  is the horizontal velocity,  $p$  is the fluid pressure,  $\mathbf{g} = (0, 0, -g)^T$  represents the gravity forces. The quantity  $\nabla$  denotes  $\nabla = \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)^T$ , and  $\nabla_{x,y}$  corresponds to the projection of  $\nabla$  on the horizontal plane i.e.  $\nabla_{x,y} = \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right)^T$ .

We assume a Newtonian fluid and hence the stress tensor  $\boldsymbol{\Sigma}$  is given by

$$\boldsymbol{\Sigma} = -p\mathbf{1} + \mu[\nabla \mathbf{U} + (\nabla \mathbf{U})^T]. \quad (3)$$

where  $\mu$  is the viscosity coefficient. We consider a free surface flow (see Fig. 1-(a)), therefore we assume

$$z_b(x, y) \leq z \leq \eta(t, x, y) := h(t, x, y) + z_b(x, y)$$

with  $z_b(x, y)$  the bottom elevation and  $h(t, x, y)$  the water depth.

## 2.1 Boundary conditions

### 2.1.1 Bottom and free surface

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 1-(a))

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

On the bottom we prescribe an impermeability condition

$$\mathbf{U} \cdot \mathbf{n}_b = 0, \tag{4}$$

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

$$(\boldsymbol{\Sigma} \cdot \mathbf{n}_b) \cdot \mathbf{t}_i = -\kappa \mathbf{U} \cdot \mathbf{t}_i, \quad i = 1, 2 \tag{5}$$

with  $\kappa$  a Navier coefficient and  $(\mathbf{t}_i, i = 1, 2)$  two tangential vectors. For some applications, we rather use more specific friction laws and the equation (5) is then replaced by

$$(\boldsymbol{\Sigma} \cdot \mathbf{n}_b) \cdot \mathbf{t}_i = -\kappa(h, \mathbf{U}) \cdot \mathbf{t}_i, \quad i = 1, 2 \tag{6}$$

with  $\kappa(h, \mathbf{U}) \cdot \mathbf{U} \geq 0$ . On the free surface, the kinematic boundary condition is satisfied

$$\frac{\partial \eta}{\partial t} + \mathbf{u}(t, x, y, \eta) \cdot \nabla_{x,y} \eta - w(t, x, y, \eta) = 0, \tag{7}$$

and the no stress condition

$$\boldsymbol{\Sigma} \cdot \mathbf{n}_s = -p^a(t, x, y) \mathbf{n}_s + W(t, x, y) \mathbf{t}_s, \tag{8}$$

where  $p^a(t, x, y), W(t, x, y)$  are two given quantities,  $p^a$  (resp.  $W$ ) mimicks the effects of the atmospheric pressure (resp. the wind blowing at the free surface). Throughout the paper  $p^a = 0$ ,  $W = 0$  except in paragraph 6.3 where the effects of the atmospheric pressure and the wind are considered.

### 2.1.2 Fluid boundaries and solid walls

On solid walls, we prescribe a slip condition

$$\mathbf{U} \cdot \mathbf{n} = 0, \tag{9}$$

$\mathbf{n}$  being the outward normal to the considered wall.

In this paper we consider fluid boundaries on which we prescribe zero, one or two of the following conditions depending on the type of the flow (fluvial or torrential) : Water level  $h + z_b(x, y)$  given, flux  $h\mathbf{U}$  given.

The system is completed with some initial conditions

$$h(0, x, y) = h^0(x, y), \quad \mathbf{U}(0, x, y, z) = \mathbf{U}^0(x, y, z),$$

with  $\mathbf{U}^0$  satisfying the divergence free condition (1).

## 2.2 The hydrostatic Navier-Stokes system

The vorticity is defined by

$$\mathbf{\Omega} = \left( \frac{\partial w}{\partial y} - \frac{\partial v}{\partial z}, \frac{\partial u}{\partial z} - \frac{\partial w}{\partial x}, \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right)^T,$$

and we assume the two first components of  $\mathbf{\Omega}$  vanish i.e. there is no vorticity along the horizontal axis. We will see hereafter that this assumption simplifies the expression of the viscous terms that reduce to a laplacian operator whose numerical treatment is easier than the complete expression of the rheology terms. Some of the authors have proposed a discretization of the Navier-Stokes system with a general rheology law [24], but as already mentioned we focus here on the numerical approximation of the Euler part of the Navier-Stokes system. This is the reason why we adopt a simplified expression for the viscous terms.

If the two first components of  $\mathbf{\Omega}$  vanish then the viscous terms in the momentum equation along  $z$  of the Navier-Stokes system namely,

$$\frac{\partial}{\partial x} \left( \mu \frac{\partial u}{\partial z} + \mu \frac{\partial w}{\partial x} \right) + \frac{\partial}{\partial y} \left( \mu \frac{\partial v}{\partial z} + \mu \frac{\partial w}{\partial y} \right) + 2\mu \frac{\partial^2 w}{\partial z^2},$$

also vanish using the divergence free condition (1). Hence the momentum equation along  $z$  of the Navier-Stokes reduces to

$$\frac{\partial w}{\partial t} + \frac{\partial uw}{\partial x} + \frac{\partial vw}{\partial y} + \frac{\partial w^2}{\partial z} + \frac{\partial p}{\partial z} = -g.$$

The hydrostatic assumption consists in neglecting the vertical acceleration of the fluid

$$\frac{\partial w}{\partial t} + \frac{\partial uw}{\partial x} + \frac{\partial vw}{\partial y} + \frac{\partial w^2}{\partial z} \approx 0,$$

see [23, 34, 39] for the analysis of hydrostatic models and for their asymptotic derivation [4, 32, 38]. This implies that the pressure writes

$$p = g(\eta - z). \quad (10)$$

Therefore, the hydrostatic approximation of the system (1)-(9) consists in the model

$$\nabla \cdot \mathbf{U} = 0, \quad (11)$$

$$\frac{\partial \mathbf{u}}{\partial t} + \nabla_{x,y} \cdot (\mathbf{u} \otimes \mathbf{u}) + \frac{\partial \mathbf{u}w}{\partial z} + g \nabla_{x,y} \eta = \nabla \cdot \hat{\mathbf{T}}, \quad (12)$$

with

$$\hat{\mathbf{T}} = \mu \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \end{pmatrix},$$

with the kinematic boundary conditions (4),(7) and the dynamical boundary conditions

$$(\hat{T} \cdot \mathbf{n}_b) \cdot \mathbf{t}_i = -\kappa \mathbf{U} \cdot \mathbf{t}_i \quad (13)$$

$$\hat{T} \cdot \mathbf{n}_s = \mathbf{0}. \quad (14)$$

Notice that  $\nabla \cdot \hat{\mathbf{T}} = \mu \Delta \mathbf{u}$  where  $\Delta$  denotes the classical laplacian operator.

The smooth solutions of the system (11)-(14) satisfy the energy balance

$$\frac{\partial}{\partial t} \int_{z_b}^{\eta} E \, dz + \nabla_{x,y} \cdot \int_{z_b}^{\eta} \mathbf{u} \left( (E + g(\eta - z)) - \mu \nabla_{x,y} \mathbf{u} \right) dz = -\frac{1}{\mu} \int_{z_b}^{\eta} \hat{\mathbf{T}}^2 dz - \kappa |\mathbf{u}|_b^2, \quad (15)$$

with  $\hat{\mathbf{T}}^2 = \sum_{i,j} T_{i,j}^2$  and

$$E = E(z, u, v) = \frac{u^2 + v^2}{2} + gz. \quad (16)$$

**Remark 2.1** The viscous terms in Eq. (12) result from a rewriting of the quantity  $\nabla \cdot \Sigma$  appearing in Eq. (2) using the divergence free condition (1). Therefore, the dynamic boundary conditions (13)-(14) are slightly different from (5),(8).

In the case of an inviscid fluid, the system (11)-(14) consists in the incompressible and hydrostatic Euler equations with free surface and reads

$$\nabla \cdot \mathbf{U} = 0, \quad (17)$$

$$\frac{\partial \mathbf{u}}{\partial t} + \nabla_{x,y} \cdot (\mathbf{u} \otimes \mathbf{u}) + \frac{\partial \mathbf{u} w}{\partial z} + g \nabla_{x,y} \eta = 0, \quad (18)$$

coupled with the two kinematic boundary conditions (4),(7) and  $p(t, x, y, h(t(x, y))) = 0$ .

We recall the fundamental stability property related to the fact that the hydrostatic Euler system admits, for smooth solutions, an energy conservation that can be written under the form

$$\frac{\partial}{\partial t} \int_{z_b}^{\eta} E \, dz + \nabla_{x,y} \cdot \int_{z_b}^{\eta} (\mathbf{u} (E + g(\eta - z))) \, dz = 0, \quad (19)$$

with  $E$  defined by (16).

**Remark 2.2** When considering hydrostatic models for incompressible fluids, the pressure  $p$  is no more defined as the Lagrange multiplier of the incompressibility condition but has an expression as a function of the other variables, see (10).

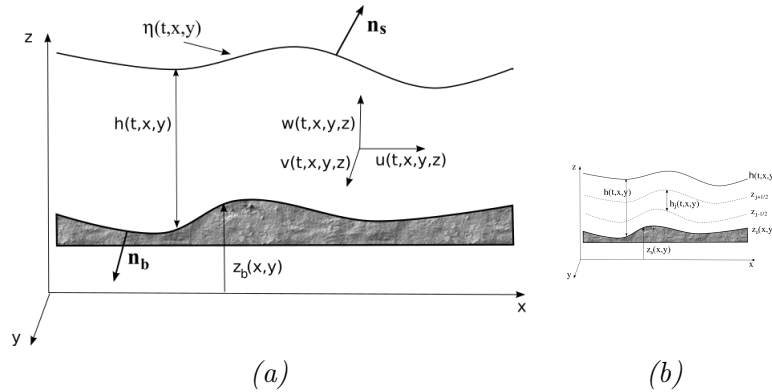


Figure 1: Fluid domain, notations and layerwise discretization.

### 3 The layer-averaged model

We consider a discretization of the fluid domain by layers (see Fig. 1-(b)) where the layer  $\alpha$  contains the points of coordinates  $(x, y, z)$  with  $z \in L_\alpha(t, x, y) = (z_{\alpha-1/2}, z_{\alpha+1/2})$  and  $\{z_{\alpha+1/2}\}_{\alpha=1,\dots,N}$  is defined by

$$\begin{cases} z_{\alpha+1/2}(t, x, y) = z_b(x, y) + \sum_{j=1}^{\alpha} h_j(t, x, y), & \alpha \in [0, \dots, N], \\ h_\alpha(t, x, y) = z_{\alpha+1/2}(t, x, y) - z_{\alpha-1/2}(t, x, y) = l_\alpha h(t, x, y), \end{cases} \quad (20)$$

and  $\sum_{\alpha=1}^N l_\alpha = 1$ .

### 3.1 The layer-averaged Euler system

The layer-averaging process for the 2d hydrostatic Euler and Navier-Stokes systems is precisely described in the paper [24] with a general rheology, the reader can refer to it. In the following, we present a Galerkin type approximation of the Euler system also leading to a layer-averaged version of the Euler system. The obtained model and the one described in [24] are identical.

Using the notations (20), let us consider the space  $\mathbb{P}_{0,h}^{N,t}$  of piecewise constant functions defined by

$$\mathbb{P}_{0,h}^{N,t} = \{ \mathbf{1}_{z \in L_\alpha(t,x,y)}(z), \quad \alpha \in \{1, \dots, N\} \}, \quad (21)$$

where  $\mathbf{1}_{z \in L_\alpha(t,x,y)}(z)$  is the characteristic function of the layer  $L_\alpha(t,x,y)$ . Using this formalism, the projection of  $u$ ,  $v$  and  $w$  on  $\mathbb{P}_{0,h}^{N,t}$  is a piecewise constant function defined by

$$X^N(t, x, y, z, \{z_\alpha\}) = \sum_{\alpha=1}^N \mathbf{1}_{[z_{\alpha-1/2}, z_{\alpha+1/2}]}(z) X_\alpha(t, x, y), \quad (22)$$

for  $X \in (u, v, w)$ .

The three following propositions hold.

**Proposition 3.1** *Using the space  $\mathbb{P}_{0,h}^{N,t}$  defined by (21) and the decomposition (22), the Galerkin approximation of the incompressible and hydrostatic Euler equations (17)-(18),(4),(7) leads to the system*

$$\sum_{\alpha=1}^N \frac{\partial h_\alpha}{\partial t} + \sum_{\alpha=1}^N \nabla_{x,y} \cdot (h_\alpha \mathbf{u}_\alpha) = 0, \quad (23)$$

$$\begin{aligned} \frac{\partial h_\alpha \mathbf{u}_\alpha}{\partial t} + \nabla_{x,y} \cdot (h_\alpha \mathbf{u}_\alpha \otimes \mathbf{u}_\alpha) + \nabla_{x,y} \left( \frac{g}{2} h h_\alpha \right) &= -g h_\alpha \nabla_{x,y} z_b \\ &+ \mathbf{u}_{\alpha+1/2} G_{\alpha+1/2} - \mathbf{u}_{\alpha-1/2} G_{\alpha-1/2}, \quad \alpha = 1, \dots, N. \end{aligned} \quad (24)$$

The quantity  $G_{\alpha+1/2}$  (resp.  $G_{\alpha-1/2}$ ) corresponds to mass exchange accross the interface  $z_{\alpha+1/2}$  (resp.  $z_{\alpha-1/2}$ ) and  $G_{\alpha+1/2}$  is defined by

$$G_{\alpha+1/2} = \sum_{j=1}^{\alpha} \left( \frac{\partial h_j}{\partial t} + \nabla_{x,y} \cdot (h_j \mathbf{u}_j) \right) = - \sum_{j=1}^N \left( \sum_{p=1}^{\alpha} l_p - \mathbf{1}_{j \leq \alpha} \right) \nabla_{x,y} \cdot (h_j \mathbf{u}_j), \quad (25)$$

for  $\alpha = 1, \dots, N$ . The velocities at the interfaces  $\mathbf{u}_{\alpha+1/2}$  are defined by

$$\mathbf{u}_{\alpha+1/2} = \begin{cases} \mathbf{u}_\alpha & \text{if } G_{\alpha+1/2} \leq 0 \\ \mathbf{u}_{\alpha+1} & \text{if } G_{\alpha+1/2} > 0 \end{cases} \quad (26)$$

The smooth solutions of (23),(24) satisfy an energy balance and we have the following proposition.

**Proposition 3.2** *The system (23),(24) admits, for smooth solutions, the energy balance*

$$\begin{aligned} \frac{\partial}{\partial t} E_\alpha + \nabla_{x,y} \cdot \left( \mathbf{u}_\alpha \left( E_\alpha + \frac{g}{2} h_\alpha h \right) \right) \\ = \left( \mathbf{u}_{\alpha+1/2} \cdot \mathbf{u}_\alpha - \frac{\mathbf{u}_\alpha^2 - \mathbf{u}_{\alpha+1}^2}{2} \right) G_{\alpha+1/2} \\ - \left( \mathbf{u}_{\alpha-1/2} \cdot \mathbf{u}_\alpha - \frac{\mathbf{u}_\alpha^2 - \mathbf{u}_{\alpha-1}^2}{2} \right) G_{\alpha-1/2}, \end{aligned} \quad (27)$$

with

$$E_\alpha = \frac{h_\alpha u_\alpha^2 + h_\alpha v_\alpha^2}{2} + \frac{g}{2} h_\alpha h + g h_\alpha z_b. \quad (28)$$

The sum of Eqs. (27) for  $\alpha = 1, \dots, N$  gives the energy balance

$$\begin{aligned} \frac{\partial}{\partial t} \sum_{\alpha=1}^N E_{\alpha} + \sum_{\alpha=1}^N \nabla_{x,y} \cdot \mathbf{u}_{\alpha} \left( E_{\alpha} + \frac{g}{2} h_{\alpha} h \right) \\ = - \sum_{\alpha=1}^N \frac{(u_{\alpha+1/2} - u_{\alpha})^2 + (v_{\alpha+1/2} - v_{\alpha})^2}{2} |G_{\alpha+1/2}|. \end{aligned} \quad (29)$$

**Remark 3.1** Equation (29) is a layer discretization of the energy balance (16). The definition of  $\mathbf{u}_{\alpha+1/2}$  given in (26) ensures the right hand side in Eq. (29) is nonpositive. Notice that a centered definition for  $\mathbf{u}_{\alpha+1/2}$  i.e.

$$\mathbf{u}_{\alpha+1/2} = \frac{\mathbf{u}_{\alpha} + \mathbf{u}_{\alpha+1}}{2}, \quad (30)$$

instead of (26) leads to a vanishing right hand side in Eq. (29). But the centered choice (30) does not allow to obtain an energy balance in the variable density case and does not give a maximum principle, at the discrete level, see [10]. Simple calculations show that any other choice than (26) or (30) leads to a non negative r.h.s. in (29).

It is noticeable that, thanks to the kinematic boundary condition at each interface, the vertical velocity is no more a variable of the system (24). This is an advantage of this formulation over the hydrostatic model where the vertical velocity is needed in the momentum equation (12) and is deduced from the incompressibility condition (11). Even if the vertical velocity  $w$  no more appears in the model (23)-(24), it can be obtained as follows.

**Proposition 3.3** The piecewise constant approximation of the vertical velocity  $w$  satisfying Eq. (22) is given by

$$w_{\alpha} = k_{\alpha} - z_{\alpha} \nabla_{x,y} \cdot \mathbf{u}_{\alpha} \quad (31)$$

with

$$k_1 = \nabla_{x,y} \cdot (z_b \mathbf{u}_1), \quad k_{\alpha+1} = k_{\alpha} + \nabla_{x,y} \cdot (z_{\alpha+1/2} (\mathbf{u}_{\alpha+1} - \mathbf{u}_{\alpha})).$$

The quantities  $\{w_{\alpha}\}_{\alpha=1}^N$  are obtained only using a post-processing of the variables governing the system (23)-(24).

**Proof of prop. 3.1** Considering the divergence free condition (17), using the decomposition (22) and the space of test functions (21), we consider the quantity

$$\int_{\mathbb{R}} \mathbf{1}_{z \in L_{\alpha}(t,x,y)} \nabla \cdot \mathbf{U}^N dz = 0,$$

with  $\mathbf{U}^N = (u^N, v^N, w^N)^T$ . Simple computations give

$$0 = \int_{\mathbb{R}} \mathbf{1}_{z \in L_{\alpha}(t,x,y)} \nabla \cdot \mathbf{U}^N dz = \frac{\partial h_{\alpha}}{\partial t} + \frac{\partial}{\partial x} \int_{z_{\alpha-1/2}}^{z_{\alpha+1/2}} u dz + \frac{\partial}{\partial y} \int_{z_{\alpha-1/2}}^{z_{\alpha+1/2}} v dz - G_{\alpha+1/2} + G_{\alpha-1/2},$$

leading to

$$\frac{\partial h_{\alpha}}{\partial t} + \nabla_{x,y} \cdot (h_{\alpha} \mathbf{u}_{\alpha}) = G_{\alpha+1/2} - G_{\alpha-1/2}, \quad (32)$$

with  $G_{\alpha \pm 1/2}$  defined by

$$G_{\alpha+1/2} = \frac{\partial z_{\alpha+1/2}}{\partial t} + \mathbf{u}_{\alpha+1/2} \cdot \nabla_{x,y} z_{\alpha+1/2} - w_{\alpha+1/2}.$$



The sum for  $\alpha = 1, \dots, N$  of the above relations gives Eq. (23) where the kinematic boundary conditions (4),(7) corresponding to

$$G_{1/2} = G_{N+1/2} = 0, \quad (33)$$

have been used. Similarly, the sum for  $j = 1, \dots, \alpha$  of the relations (32) with (33) gives the expression (25) for  $G_{\alpha+1/2}$ .

Now we consider the Galerkin approximation of Eq. (18) i.e. the quantity

$$\int_{\mathbb{R}} \mathbf{1}_{z \in L_{\alpha}(t,x,y)} \left( \frac{\partial \mathbf{u}^N}{\partial t} + \nabla_{x,y} \cdot (\mathbf{u}^N \otimes \mathbf{u}^N) + \frac{\partial \mathbf{u}^N w^N}{\partial z} + g \nabla_{x,y} \eta \right) dz = 0,$$

leading, after simple computations, to Eq. (24). ■

**Proof of prop. 3.2** In order to obtain (27) we multiply (24) by  $\mathbf{u}_{\alpha}$  and we perform simple manipulations. More precisely, the momentum equation along the  $x$  axis multiplied by  $u_{\alpha}$  gives

$$\begin{aligned} \left( \frac{\partial}{\partial t} (h_{\alpha} u_{\alpha}) + \frac{\partial}{\partial x} \left( h_{\alpha} u_{\alpha}^2 + \frac{g}{2} h h_{\alpha} \right) + \frac{\partial}{\partial y} (h_{\alpha} u_{\alpha} v_{\alpha}) \right) u_{\alpha} = \\ \left( -g h_{\alpha} \frac{\partial z_b}{\partial x} + u_{\alpha+1/2} G_{\alpha+1/2} - u_{\alpha-1/2} G_{\alpha-1/2} \right) u_{\alpha}. \end{aligned}$$

Considering first the left hand side of the preceding equation excluding the pressure terms, we denote

$$I_{u,\alpha} = \left( \frac{\partial}{\partial t} (h_{\alpha} u_{\alpha}) + \frac{\partial}{\partial x} (h_{\alpha} u_{\alpha}^2) + \frac{\partial}{\partial y} (h_{\alpha} u_{\alpha} v_{\alpha}) \right) u_{\alpha},$$

and using (32) we have

$$I_{u,\alpha} = \frac{\partial}{\partial t} \left( \frac{h_{\alpha} u_{\alpha}^2}{2} \right) + \frac{\partial}{\partial x} \left( u_{\alpha} \frac{h_{\alpha} u_{\alpha}^2}{2} \right) + \frac{\partial}{\partial y} \left( v_{\alpha} \frac{h_{\alpha} u_{\alpha}^2}{2} \right) + \frac{u_{\alpha}^2}{2} (G_{\alpha+1/2} - G_{\alpha-1/2}).$$

Now we consider the contribution of the pressure terms over the energy balance i.e.

$$I_{p,u,\alpha} = \left( \frac{\partial}{\partial x} \left( \frac{g}{2} h h_{\alpha} \right) + g h_{\alpha} \frac{\partial z_b}{\partial x} \right) u_{\alpha},$$

and it comes

$$\begin{aligned} I_{p,u,\alpha} &= g h_{\alpha} \frac{\partial}{\partial x} (h + z_b) u_{\alpha} \\ &= g \frac{\partial}{\partial x} (h_{\alpha} (h + z_b) u_{\alpha}) - g (h + z_b) \frac{\partial}{\partial x} (h_{\alpha} u_{\alpha}) \\ &= \frac{\partial}{\partial x} \left( \left( \frac{g}{2} h_{\alpha} h + \frac{g}{2} h_{\alpha} (h + 2z_b) \right) u_{\alpha} \right) - g (h + z_b) \frac{\partial}{\partial x} (h_{\alpha} u_{\alpha}). \end{aligned}$$

Performing the same manipulations over the momentum equation along  $y$  and adding the terms  $I_{u,\alpha}, I_{v,\alpha}, I_{p,u,\alpha}$  and  $I_{p,v,\alpha}$  gives the result. ■

**Proof of prop. 3.3** Using the boundary condition (4), an integration from  $z_b$  to  $z$  of the divergence free condition (11) easily gives

$$w = -\nabla_{x,y} \cdot \int_{z_b}^z \mathbf{u} \, dz.$$

Replacing formally in the above equation  $\mathbf{u}$  (resp.  $w$ ) by  $\mathbf{u}^N$  (resp.  $w^N$ ) defined by (22) and performing an integration over the layer  $L_1$  of the obtained relation yields

$$h_1 w_1 = - \int_{z_b}^{z_{3/2}} \nabla_{x,y} \cdot \int_{z_b}^z \mathbf{u}_1 \, dz dz_1 = h_1 \nabla_{x,y} \cdot (z_b \mathbf{u}_1) - \frac{z_{3/2}^2 - z_b^2}{2} \nabla_{x,y} \cdot \mathbf{u}_1,$$

i.e.

$$w_1 = \nabla_{x,y} \cdot (z_b \mathbf{u}_1) - z_1 \nabla_{x,y} \cdot \mathbf{u}_1,$$

corresponding to (31) for  $\alpha = 1$ . A similar computation for the layers  $L_2, \dots, L_N$  proves the result (31) for  $\alpha = 2, \dots, N$ .

A more detailed version of this proof is given in [24]. ■

### 3.2 The layer-averaged Navier-Stokes system

In paragraph 3.1, we have applied the layer-averaging to the Euler system, we now use the same process for the hydrostatic Navier-Stokes system. First, we consider the Navier-Stokes system (11)-(14) for a Newtonian fluid and then with a simplified rheology.

#### 3.2.1 Complete model

The layer-averaging process applied to the Navier-Stokes system (11)-(14) leads to the following proposition.

**Proposition 3.4** *The layer-averaged hydrostatic Navier-Stokes system (11)-(14) is given by*

$$\sum_{\alpha=1}^N \frac{\partial h_\alpha}{\partial t} + \sum_{\alpha=1}^N \nabla_{x,y} \cdot (h_\alpha \mathbf{u}_\alpha) = 0. \quad (34)$$

$$\begin{aligned} & \frac{\partial h_\alpha \mathbf{u}_\alpha}{\partial t} + \nabla_{x,y} \cdot (h_\alpha \mathbf{u}_\alpha \otimes \mathbf{u}_\alpha) + \nabla_{x,y} \left( \frac{g}{2} h h_\alpha \right) = -g h_\alpha \nabla_{x,y} z_b \\ & + \mathbf{u}_{\alpha+1/2} G_{\alpha+1/2} - \mathbf{u}_{\alpha-1/2} G_{\alpha-1/2} + \nabla_{x,y} \cdot (h_\alpha \mathbf{T}_\alpha) \\ & - \mathbf{T}_{\alpha+1/2} \nabla_{x,y} z_{\alpha+1/2} + \mathbf{T}_{\alpha-1/2} \nabla_{x,y} z_{\alpha-1/2} \\ & + 2\mu_{\alpha+1/2} \frac{\mathbf{u}_{\alpha+1} - \mathbf{u}_\alpha}{h_{\alpha+1} + h_\alpha} - 2\mu_{\alpha-1/2} \frac{\mathbf{u}_\alpha - \mathbf{u}_{\alpha-1}}{h_\alpha + h_{\alpha-1}} - \kappa_\alpha \mathbf{u}_\alpha, \quad \alpha = 1, \dots, N, \end{aligned} \quad (35)$$

with

$$\mathbf{T}_{\alpha+1/2} = \begin{pmatrix} T_{xx,\alpha+1/2} & T_{xy,\alpha+1/2} \\ T_{yx,\alpha+1/2} & T_{yy,\alpha+1/2} \end{pmatrix}, \quad (36)$$

$$T_{xx,\alpha+1/2} = \frac{\mu_{\alpha+1/2}}{h_{\alpha+1} + h_\alpha} \left( h_\alpha \frac{\partial u_\alpha}{\partial x} + h_{\alpha+1} \frac{\partial u_{\alpha+1}}{\partial x} \right) - 2\mu_{\alpha+1/2} \frac{\partial z_{\alpha+1/2}}{\partial x} \frac{u_{\alpha+1} - u_\alpha}{h_{\alpha+1} + h_\alpha}, \quad (37)$$

$$T_{xy,\alpha+1/2} = \frac{\mu_{\alpha+1/2}}{h_{\alpha+1} + h_\alpha} \left( h_\alpha \frac{\partial u_\alpha}{\partial y} + h_{\alpha+1} \frac{\partial u_{\alpha+1}}{\partial y} \right) - 2\mu_{\alpha+1/2} \frac{\partial z_{\alpha+1/2}}{\partial y} \frac{u_{\alpha+1} - u_\alpha}{h_{\alpha+1} + h_\alpha}, \quad (38)$$

$$\mathbf{T}_\alpha = \begin{pmatrix} T_{xx,\alpha} & T_{xy,\alpha} \\ T_{yx,\alpha} & T_{yy,\alpha} \end{pmatrix} = \frac{\mathbf{T}_{\alpha+1/2} + \mathbf{T}_{\alpha-1/2}}{2}, \quad (39)$$

and

$$\kappa_\alpha = \begin{cases} \kappa & \text{if } \alpha = 1 \\ 0 & \text{if } \alpha \neq 1 \end{cases} \quad \mu_{\alpha+1/2} = \begin{cases} 0 & \text{if } \alpha = 0 \\ \mu & \text{if } \alpha = 1, \dots, N-1 \\ 0 & \text{if } \alpha = N \end{cases} \quad (40)$$

The vertical velocities  $\{w_\alpha\}_{\alpha=1}^N$  are defined by (31).

For smooth solutions, the system (34)-(35) admits the energy balance

$$\begin{aligned} & \frac{\partial}{\partial t} \sum_{\alpha=1}^N E_\alpha + \nabla_{x,y} \cdot \sum_{\alpha=1}^N \mathbf{u}_\alpha \left( E_\alpha + \frac{g}{2} h_\alpha h - h_\alpha \mathbf{T}_\alpha \right) \\ & - \mu_{\alpha+1/2} (\mathbf{u}_{\alpha+1} + \mathbf{u}_\alpha) \cdot \frac{\mathbf{u}_{\alpha+1} - \mathbf{u}_\alpha}{h_{\alpha+1} + h_\alpha} - \mu_{\alpha-1/2} (\mathbf{u}_\alpha + \mathbf{u}_{\alpha-1}) \cdot \frac{\mathbf{u}_\alpha - \mathbf{u}_{\alpha-1}}{h_\alpha + h_{\alpha-1}} \end{aligned}$$

$$\begin{aligned}
&= - \sum_{\alpha=1}^{N-1} \frac{(u_{\alpha+1/2} - u_{\alpha})^2 + (v_{\alpha+1/2} - v_{\alpha})^2}{2} |G_{\alpha+1/2}| \\
&\quad - \sum_{\alpha=1}^{N-1} \frac{h_{\alpha+1} + h_{\alpha}}{2\mu} \mathbf{T}_{\alpha+1/2}^2 - \sum_{\alpha=1}^{N-1} 2\mu \frac{(\mathbf{u}_{\alpha+1} - \mathbf{u}_{\alpha})^2}{h_{\alpha+1} + h_{\alpha}} - \kappa(u_1^2 + v_1^2), \tag{41}
\end{aligned}$$

with  $E_{\alpha}$  defined by (28) and  $\mathbf{T}_{\alpha+1/2}^2 = \sum_{i,j} T_{i,j,\alpha+1/2}^2$ . Relation (41) is consistent with a layer-averaged discretisation of the equation (15).

**Proof of prop. 3.4** The proof is given in appendix A. ■

**Remark 3.2** Notice in the definition (39), since we consider viscous terms we use a centered approximation.

**Remark 3.3** Using simple manipulations, the viscous terms in the layer-averaged Navier-Stokes can be rewritten and Eq. (35) becomes

$$\begin{aligned}
&\frac{\partial h_{\alpha} \mathbf{u}_{\alpha}}{\partial t} + \nabla_{x,y} \cdot (h_{\alpha} \mathbf{u}_{\alpha} \otimes \mathbf{u}_{\alpha}) + \nabla_{x,y} \left( \frac{g}{2} h h_{\alpha} \right) = -g h_{\alpha} \nabla_{x,y} z_b \\
&\quad + \mathbf{u}_{\alpha+1/2} G_{\alpha+1/2} - \mathbf{u}_{\alpha-1/2} G_{\alpha-1/2} \\
&\quad + h_{\alpha} \nabla_{x,y} \cdot \mathbf{T}_{\alpha} - (\mathbf{T}_{\alpha+1/2} - \mathbf{T}_{\alpha-1/2}) \nabla_{x,y} z_{\alpha} \\
&\quad + 2\mu_{\alpha+1/2} \frac{\mathbf{u}_{\alpha+1} - \mathbf{u}_{\alpha}}{h_{\alpha+1} + h_{\alpha}} - 2\mu_{\alpha-1/2} \frac{\mathbf{u}_{\alpha} - \mathbf{u}_{\alpha-1}}{h_{\alpha} + h_{\alpha-1}} - \kappa_{\alpha} \mathbf{u}_{\alpha}, \quad \alpha = 1, \dots, N,
\end{aligned}$$

### 3.2.2 Simplified rheology

The viscous terms in the layer-averaged Navier-Stokes system are difficult to discretize and in this paragraph, we propose a simplified version of the model given in prop. 3.4.

First, using simple manipulations, the viscous terms in the layer-averaged Navier-Stokes can be rewritten and Eq. (35) becomes

$$\begin{aligned}
&\frac{\partial h_{\alpha} \mathbf{u}_{\alpha}}{\partial t} + \nabla_{x,y} \cdot (h_{\alpha} \mathbf{u}_{\alpha} \otimes \mathbf{u}_{\alpha}) + \nabla_{x,y} \left( \frac{g}{2} h h_{\alpha} \right) = -g h_{\alpha} \nabla_{x,y} z_b \\
&\quad + \mathbf{u}_{\alpha+1/2} G_{\alpha+1/2} - \mathbf{u}_{\alpha-1/2} G_{\alpha-1/2} + \nabla_{x,y} \cdot (h_{\alpha} \mathbf{T}_{\alpha}^0) + \mathbf{T}_{\alpha}^1 \\
&\quad + \Lambda_{\alpha+1/2} (\mathbf{u}_{\alpha+1} - \mathbf{u}_{\alpha}) - \Lambda_{\alpha-1/2} (\mathbf{u}_{\alpha} - \mathbf{u}_{\alpha-1}) - \kappa_{\alpha} \mathbf{u}_{\alpha}, \quad \alpha = 1, \dots, N,
\end{aligned}$$

with

$$T_{xx,\alpha+1/2}^0 = \frac{\mu_{\alpha+1/2}}{h_{\alpha+1} + h_{\alpha}} \left( h_{\alpha} \frac{\partial u_{\alpha}}{\partial x} + h_{\alpha+1} \frac{\partial u_{\alpha+1}}{\partial x} \right), \tag{42}$$

$$T_{xy,\alpha+1/2}^0 = \frac{\mu_{\alpha+1/2}}{h_{\alpha+1} + h_{\alpha}} \left( h_{\alpha} \frac{\partial u_{\alpha}}{\partial y} + h_{\alpha+1} \frac{\partial u_{\alpha+1}}{\partial y} \right), \tag{43}$$

$$T_{xx,\alpha+1/2}^1 = 2\mu_{\alpha+1/2} \frac{\partial z_{\alpha+1/2}}{\partial x} \frac{\partial u_{\alpha+1}}{\partial x}, \tag{44}$$

$$T_{xy,\alpha+1/2}^1 = 2\mu_{\alpha+1/2} \frac{\partial z_{\alpha+1/2}}{\partial y} \frac{\partial u_{\alpha+1}}{\partial y}, \tag{45}$$

$$\Lambda_{\alpha+1/2} = 2\mu_{\alpha+1/2} \frac{1 + |\nabla_{x,y} z_{\alpha+1/2}|^2}{h_{\alpha+1} + h_{\alpha}} - \mu_{\alpha+1/2} \nabla_{x,y} \cdot \left( \frac{h_{\alpha} \nabla_{x,y} z_{\alpha+1/2}}{h_{\alpha+1} + h_{\alpha}} \right), \tag{46}$$

$$\mathbf{T}_{\alpha+1/2}^i = \begin{pmatrix} T_{xx,\alpha+1/2}^i & T_{xy,\alpha+1/2}^i \\ T_{yx,\alpha+1/2}^i & T_{yy,\alpha+1/2}^i \end{pmatrix}, \tag{47}$$

$$\mathbf{T}_{\alpha}^i = \begin{pmatrix} T_{xx,\alpha}^i & T_{xy,\alpha}^i \\ T_{yx,\alpha}^i & T_{yy,\alpha}^i \end{pmatrix} = \frac{\mathbf{T}_{\alpha+1/2}^i + \mathbf{T}_{\alpha-1/2}^i}{2}, \quad \text{for } i = 0, 1. \tag{48}$$

Then, following the same arguments as in [37],[11, paragraph 3.3] (we consider a shallow water type regime where the viscous effects are less important than the advection and pressure terms) it is possible to neglect the quantities  $\mathbf{T}_\alpha^1$  and the second term in the expression (46), we get the following result.

**Proposition 3.5** *With a simplified expression of the rheology terms, the layer-averaged hydrostatic Navier-Stokes system given in prop.3.4 becomes*

$$\sum_{\alpha=1}^N \frac{\partial h_\alpha}{\partial t} + \sum_{\alpha=1}^N \nabla_{x,y} \cdot (h_\alpha \mathbf{u}_\alpha) = 0, \quad (49)$$

$$\begin{aligned} \frac{\partial h_\alpha \mathbf{u}_\alpha}{\partial t} + \nabla_{x,y} \cdot (h_\alpha \mathbf{u}_\alpha \otimes \mathbf{u}_\alpha) + \nabla_{x,y} \left( \frac{g}{2} h h_\alpha \right) &= -g h_\alpha \nabla_{x,y} z_b \\ &+ \mathbf{u}_{\alpha+1/2} G_{\alpha+1/2} - \mathbf{u}_{\alpha-1/2} G_{\alpha-1/2} + \nabla_{x,y} \cdot (h_\alpha \mathbf{T}_\alpha^0) \\ &+ \Gamma_{\alpha+1/2} (\mathbf{u}_{\alpha+1} - \mathbf{u}_\alpha) - \Gamma_{\alpha-1/2} (\mathbf{u}_\alpha - \mathbf{u}_{\alpha-1}) - \kappa_\alpha \mathbf{u}_\alpha, \end{aligned} \quad (50)$$

with  $T^0$  defined by (42),(43),(47), (48) and

$$\Gamma_{\alpha+1/2} = 2\mu_{\alpha+1/2} \frac{1 + |\nabla_{x,y} z_{\alpha+1/2}|^2}{h_{\alpha+1} + h_\alpha}. \quad (51)$$

For smooth solutions, the system (49)-(50) admits the energy balance

$$\begin{aligned} \frac{\partial}{\partial t} \sum_{\alpha=1}^N E_\alpha + \nabla_{x,y} \cdot \sum_{\alpha=1}^N \mathbf{u}_\alpha \left( E_\alpha + \frac{g}{2} h_\alpha h - h_\alpha \mathbf{T}_\alpha^0 \right) \\ - \Gamma_{\alpha+1/2} \frac{\mathbf{u}_{\alpha+1} + \mathbf{u}_\alpha}{2} \cdot (\mathbf{u}_{\alpha+1} - \mathbf{u}_\alpha) - \Gamma_{\alpha-1/2} \frac{\mathbf{u}_\alpha + \mathbf{u}_{\alpha-1}}{2} \cdot (\mathbf{u}_\alpha - \mathbf{u}_{\alpha-1}) \\ = - \sum_{\alpha=1}^N \frac{(u_{\alpha+1/2} - u_\alpha)^2 + (v_{\alpha+1/2} - v_\alpha)^2}{2} |G_{\alpha+1/2}| \\ - \sum_{\alpha=1}^{N-1} \frac{h_{\alpha+1} + h_\alpha}{2\mu} (\mathbf{T}_{\alpha+1/2}^0)^2 - \sum_{\alpha=1}^{N-1} \Gamma_{\alpha+1/2} (\mathbf{u}_{\alpha+1} - \mathbf{u}_\alpha)^2 - \kappa(u_1^2 + v_1^2), \end{aligned} \quad (52)$$

**Remark 3.4** *The layer-averaged Navier-Stokes system obtained in Prop. 3.5 has the form*

$$\frac{\partial U}{\partial t} + \frac{\partial F_x(U)}{\partial x} + \frac{\partial F_y(U)}{\partial y} = \mathcal{S}_b(U) + S_e(U, \partial_t U, \partial_x U) + S_{v,f}(U), \quad (53)$$

where

$$U = (h, q_{x,1}, \dots, q_{x,N}, q_{y,1}, \dots, q_{y,N})^T,$$

and

$$\mathcal{S}_b(U) = \left( 0, gh_1 \frac{\partial z_b}{\partial x}, \dots, gh_N \frac{\partial z_b}{\partial x}, gh_1 \frac{\partial z_b}{\partial y}, \dots, gh_N \frac{\partial z_b}{\partial y} \right)^T,$$

with  $q_{x,\alpha} = h_\alpha u_\alpha$ ,  $q_{y,\alpha} = h_\alpha v_\alpha$ . We denote with  $F_x(U)$ ,  $F_y(U)$  the fluxes of the conservative part, and with  $S_e(U, \partial_t U, \partial_x U)$  and  $S_{v,f}(U)$  the source terms, representing respectively the momentum exchanges and the viscous and friction effects.

## 4 Kinetic description for the Euler system

In this section we give a kinetic interpretation of the system (23)-(27). The numerical scheme for the system (23)-(24),(31) will be deduced from the kinetic description.

## 4.1 Preliminaries

We begin this section by recalling the classical kinetic approach – used in [42] for example – for the 1d Saint-Venant system

$$\begin{aligned}\partial_t h + \partial_x(hu) &= 0, \\ \partial_t(hu) + \partial_x(hu^2 + g\frac{h^2}{2}) + gh\partial_x z_b &= 0,\end{aligned}\tag{54}$$

with the water depth  $h(t, x) \geq 0$ , the water velocity  $u(t, x) \in \mathbb{R}$  and a slowly varying topography  $z_b(x)$ .

The kinetic Maxwellian is given by

$$M(U, \xi) = \frac{1}{g\pi} \left( 2gh - (\xi - u)^2 \right)_+^{1/2},\tag{55}$$

where  $U = (h, hu)^T$ ,  $\xi \in \mathbb{R}$  and  $x_+ \equiv \max(0, x)$  for any  $x \in \mathbb{R}$ . It satisfies the following moment relations,

$$\int_{\mathbb{R}} \begin{pmatrix} 1 \\ \xi \end{pmatrix} M(U, \xi) d\xi = U, \quad \int_{\mathbb{R}} \xi^2 M(U, \xi) d\xi = hu^2 + g\frac{h^2}{2}.\tag{56}$$

These definitions allow us to obtain a *kinetic representation* of the Saint-Venant system.

**Lemma 4.1** *If the topography  $z_b(x)$  is Lipschitz continuous, the pair of functions  $(h, hu)$  is a weak solution to the Saint-Venant system (54) if and only if  $M(U, \xi)$  satisfies the kinetic equation*

$$\partial_t M + \xi \partial_x M - g(\partial_x z_b) \partial_\xi M = Q,\tag{57}$$

for some “collision term”  $Q(t, x, \xi)$  that satisfies, for a.e.  $(t, x)$ ,

$$\int_{\mathbb{R}} Q d\xi = \int_{\mathbb{R}} \xi Q d\xi = 0.\tag{58}$$

**Proof** If (57) and (58) are satisfied, we can multiply (57) by  $(1, \xi)^T$ , and integrate with respect to  $\xi$ . Using (56) and (58) and integrating by parts the term in  $\partial_\xi M$ , we obtain (54). Conversely, if  $(h, hu)$  is a weak solution to (54), just define  $Q$  by (57); it will satisfy (58) according to the same computations. ■

The standard way to use Lemma 4.1 is to write a kinetic relaxation equation [6, 15, 16], like

$$\partial_t f + \xi \partial_x f - g(\partial_x z_b) \partial_\xi f = \frac{M - f}{\epsilon},\tag{59}$$

where  $f(t, x, \xi) \geq 0$ ,  $M = M(U, \xi)$  with  $U(t, x) = \int (1, \xi)^T f(t, x, \xi) d\xi$ , and  $\epsilon > 0$  is a relaxation time. In the limit  $\epsilon \rightarrow 0$  we recover formally the formulation (57), (58). We refer to [15] for general considerations on such kinetic relaxation models without topography, the case with topography being introduced in [42]. Note that the notion of *kinetic representation* as (57), (58) differs from the so called *kinetic formulations* where a large set of entropies is involved, see [41]. For systems of conservation laws, these kinetic formulations include non-advective terms that prevent from writing down simple approximations. In general, kinetic relaxation approximations can be compatible with just a single entropy. Nevertheless this is enough for proving the convergence as  $\epsilon \rightarrow 0$ , see [14].

## 4.2 Kinetic interpretation

In this paragraph, we give a kinetic interpretation of the model (23)-(24),(29).

To build the Gibbs equilibria, we choose the function

$$\chi_0(z_1, z_2) = \frac{1}{4\pi} \mathbf{1}_{z_1^2 + z_2^2 \leq 4}. \quad (60)$$

This choice corresponds to the 2d version of the kinetic maxwellian used in 1d (see remark 4.2) and we have

$$M_\alpha = M(U_\alpha, \xi, \gamma) = \frac{h_\alpha}{c^2} \chi_0 \left( \frac{\xi - u_\alpha}{c}, \frac{\gamma - v_\alpha}{c} \right), \quad (61)$$

with  $c = \sqrt{\frac{g}{2}h}$

$$U_\alpha = (h_\alpha, h_\alpha u_\alpha, h_\alpha v_\alpha)^T, \quad (62)$$

and where  $(\xi, \gamma) \in \mathbb{R}^2$ . In other words, we have  $M_\alpha = \frac{l_\alpha}{2g\pi} \mathbf{1}_{(\xi - u_\alpha)^2 + (\gamma - v_\alpha)^2 \leq 2gh}$ .

**Remark 4.2** *Starting from the 2d maxwellian in the single layer case i.e.*

$$M_{sv} = \frac{1}{2g\pi} \mathbf{1}_{(\xi - u)^2 + (\gamma - v)^2 \leq 2gh}, \quad (63)$$

and computing its integral w.r.t. the variable  $\gamma$  yields

$$\int_{\mathbb{R}} M_{sv} d\gamma = \int_{v - \sqrt{(2gh - (\xi - u)^2)_+}}^{v + \sqrt{(2gh - (\xi - u)^2)_+}} \frac{1}{2g\pi} d\gamma = \frac{1}{g\pi} \sqrt{(2gh - (\xi - u)^2)_+},$$

that is exactly the expression (55).

The quantity  $M_\alpha$  satisfies the following moment relations

$$\int_{\mathbb{R}^2} \begin{pmatrix} 1 \\ \xi \\ \gamma \end{pmatrix} M(U_\alpha, \xi, \gamma) d\xi d\gamma = \begin{pmatrix} h_\alpha \\ h_\alpha u_\alpha \\ h_\alpha v_\alpha \end{pmatrix}, \quad \int_{\mathbb{R}^2} \begin{pmatrix} \xi^2 \\ \xi\gamma \\ \gamma^2 \end{pmatrix} M(U_\alpha, \xi, \gamma) d\xi d\gamma = \begin{pmatrix} h_\alpha u_\alpha^2 + g \frac{h_\alpha h}{2} \\ h_\alpha u_\alpha v_\alpha \\ h_\alpha v_\alpha^2 + g \frac{h_\alpha h}{2} \end{pmatrix}. \quad (64)$$

The interest of the function  $\chi_0$  and hence the particular form (61) lies in its link with a kinetic entropy. Consider the kinetic entropy

$$H(f, \xi, \gamma, z_b) = \frac{\xi^2 + \gamma^2}{2} f + g z_b f, \quad (65)$$

where  $f \geq 0$ ,  $(\xi, \gamma) \in \mathbb{R}^2$ ,  $z_b \in \mathbb{R}$ . Then one can check the relations

$$\int_{\mathbb{R}^2} \begin{pmatrix} 1 \\ \xi \\ \gamma \end{pmatrix} H(M_\alpha, \xi, \gamma) d\xi d\gamma = \begin{pmatrix} E_\alpha = \frac{h_\alpha}{2} (u_\alpha^2 + v_\alpha^2) + \frac{g}{2} h_\alpha (h + 2z_b) \\ u_\alpha (E_\alpha + \frac{g}{2} h_\alpha h) \\ v_\alpha (E_\alpha + \frac{g}{2} h_\alpha h) \end{pmatrix} \quad (66)$$

Let us introduce the Gibbs equilibria  $N_{\alpha+1/2}$  defined by for  $\alpha = 0, \dots, N$  by

$$\begin{aligned} N_{\alpha+1/2} &= N(U_{\alpha+1/2}, \xi) = \frac{G_{\alpha+1/2}}{c^2} \chi_0 \left( \frac{\xi - u_{\alpha+1/2}}{c}, \frac{\gamma - v_{\alpha+1/2}}{c} \right) \\ &= \frac{G_{\alpha+1/2}}{g\pi h} \mathbf{1}_{(\xi - u_{\alpha+1/2})^2 + (\gamma - v_{\alpha+1/2})^2 \leq 2gh} = \frac{G_{\alpha+1/2}}{h} M_{\alpha+1/2}, \end{aligned} \quad (67)$$

where  $G_{\alpha+1/2}$  is defined by (25) and  $u_{\alpha+1/2}, v_{\alpha+1/2}$  are given by (26). The quantity  $N_{\alpha+1/2}$  satisfies the following moment relations

$$\int_{\mathbb{R}^2} \begin{pmatrix} 1 \\ \xi \\ \gamma \end{pmatrix} N_{\alpha+1/2} d\xi d\gamma = \begin{pmatrix} G_{\alpha+1/2} \\ u_{\alpha+1/2} G_{\alpha+1/2} \\ v_{\alpha+1/2} G_{\alpha+1/2} \end{pmatrix}, \quad \int_{\mathbb{R}^2} \begin{pmatrix} \xi^2 \\ \frac{\xi^2}{2} \\ \frac{\gamma^2}{2} \end{pmatrix} N_{\alpha+1/2} d\xi d\gamma = \begin{pmatrix} \left( \frac{u_{\alpha+1/2}^2}{2} + \frac{g}{4} h \right) G_{\alpha+1/2} \\ \left( \frac{v_{\alpha+1/2}^2}{2} + \frac{g}{4} h \right) G_{\alpha+1/2} \end{pmatrix}. \quad (68)$$

Notice that from (25), we can give a kinetic interpretation on the exchange terms under the form

$$G_{\alpha+1/2} = - \sum_{j=1}^N \left( \sum_{p=1}^{\alpha} l_p - \mathbf{1}_{j \leq \alpha} \right) \int_{\mathbb{R}^2} \left( \xi \frac{\partial M_j}{\partial x} + \gamma \frac{\partial M_j}{\partial y} \right) d\xi d\gamma, \quad (69)$$

for  $\alpha = 1, \dots, N$ .

Then we have the two following results.

**Proposition 4.1** *The functions  $\mathbf{u}^N$  defined by (22) and  $h$  are strong solutions of the system (23)-(24) if and only if the sets of equilibria  $\{M_{\alpha}\}_{\alpha=1}^N$ ,  $\{N_{\alpha+1/2}\}_{\alpha=0}^N$  are solutions of the kinetic equations defined by*

$$(\mathcal{B}_{\alpha}) \quad \frac{\partial M_{\alpha}}{\partial t} + \xi \frac{\partial M_{\alpha}}{\partial x} + \gamma \frac{\partial M_{\alpha}}{\partial y} - g \frac{\partial z_b}{\partial x} \frac{\partial M_{\alpha}}{\partial \xi} - g \frac{\partial z_b}{\partial y} \frac{\partial M_{\alpha}}{\partial \gamma} - N_{\alpha+1/2} + N_{\alpha-1/2} = Q_{\alpha}, \quad (70)$$

for  $\alpha = 1, \dots, N$ . The quantities  $Q_{\alpha} = Q_{\alpha}(t, x, y, \xi, \gamma)$  are “collision terms” equal to zero at the macroscopic level, i.e. they satisfy a.e. for values of  $(t, x, y)$

$$\int_{\mathbb{R}^2} Q_{\alpha} d\xi d\gamma = \int_{\mathbb{R}^2} \xi Q_{\alpha} d\xi d\gamma = \int_{\mathbb{R}^2} \gamma Q_{\alpha} d\xi d\gamma = 0. \quad (71)$$

**Proposition 4.2** *The solutions of (70) are entropy solutions if*

$$\begin{aligned} \frac{\partial H(M_{\alpha})}{\partial t} + \xi \frac{\partial H(M_{\alpha})}{\partial x} + \gamma \frac{\partial H(M_{\alpha})}{\partial y} - g \frac{\partial z_b}{\partial x} \frac{\partial H(M_{\alpha})}{\partial \xi} - g \frac{\partial z_b}{\partial y} \frac{\partial H(M_{\alpha})}{\partial \gamma} \\ \leq (H(N_{\alpha+1/2}) - H(N_{\alpha-1/2})), \end{aligned} \quad (72)$$

with the notation  $H(M) = H(M, \xi, \gamma, z_b)$  and  $H$  defined by (65). The integration in  $\xi, \gamma$  of relation (72) gives

$$\begin{aligned} \frac{\partial E_{\alpha}}{\partial t} + \frac{\partial}{\partial x} u_{\alpha} (E_{\alpha} + \frac{g}{2} h_{\alpha} h) + \frac{\partial}{\partial y} v_{\alpha} (E_{\alpha} + \frac{g}{2} h_{\alpha} h) \leq l_{\alpha} \left( \frac{u_{\alpha+1/2}^2 + v_{\alpha+1/2}^2}{2} + g z_b \right) G_{\alpha+1/2} \\ - l_{\alpha} \left( \frac{u_{\alpha-1/2}^2 + v_{\alpha-1/2}^2}{2} + g z_b \right) G_{\alpha-1/2}. \end{aligned}$$

**Proof of proposition 4.1** *The proof relies on averages w.r.t the variables  $\xi, \gamma$  of Eq. (70) against the vector  $(1, \xi, \gamma)^T$ . Using relations (64), (67), (68) and the properties of the collision terms (71), the quantities*

$$\int_{\mathbb{R}^2} (\mathcal{B}_{\alpha}) d\xi d\gamma, \quad \int_{\mathbb{R}^2} \xi (\mathcal{B}_{\alpha}) d\xi d\gamma, \quad \text{and} \quad \int_{\mathbb{R}^2} \gamma (\mathcal{B}_{\alpha}) d\xi d\gamma,$$

respectively give Eqs. (32) and (24). The sum for  $\alpha = 1$  to  $N$  of Eqs. (32) with (25) gives (23) that completes the proof. ■

**Proof of prop. 4.2** The proof is obtained multiplying (70) by  $H'_\alpha(\overline{M}_\alpha, \xi, \gamma, z_b)$ . Indeed, it is easy to see that

$$H'_\alpha(M_\alpha, \xi, \gamma, z_b) \frac{\partial M_\alpha}{\partial v} = \frac{\partial}{\partial v} H_\alpha(M_\alpha, \xi, \gamma, z_b),$$

for  $v = t, x, y, \xi, \gamma$ . Likewise for the quantity  $H'_\alpha(M_\alpha, \xi, \gamma, z_b) N_{\alpha+1/2}$ , we have

$$H'_\alpha(M_\alpha, \xi, \gamma, z_b) N_{\alpha+1/2} = H(N_{\alpha+1/2}, \xi, \gamma, z_b).$$

So finally, Eq. (70) multiplied by  $H'_\alpha(M_\alpha, \xi, \gamma, z_b)$  gives

$$\frac{\partial H_\alpha}{\partial t} + \xi \frac{\partial H_\alpha}{\partial x} + \gamma \frac{\partial H_\alpha}{\partial y} - g \frac{\partial z_b}{\partial x} \frac{\partial H_\alpha}{\partial \xi} - g \frac{\partial z_b}{\partial y} \frac{\partial H_\alpha}{\partial \gamma} \leq \left( \frac{\xi^2 + \gamma^2}{2} + g z_b \right) (N_{\alpha+1/2} - N_{\alpha-1/2}).$$

It remains to calculate the sum of the preceding relations from  $\alpha = 1, \dots, N$  and to integrate the obtained relation in  $\xi, \gamma$  over  $\mathbb{R}^2$  that completes the proof. ■

**Remark 4.3** If we introduce a  $(2N + 1) \times N$  matrix  $\mathcal{K}(\xi, \gamma)$  defined by

$$\mathcal{K}_{1,j} = 1, \quad \mathcal{K}_{i+1,j} = \xi \delta_{i,j}, \quad \mathcal{K}_{i+N+1,j+N} = \gamma \delta_{i,j},$$

for  $i, j = 1, \dots, N$  with  $\delta_{i,j}$  the Kronecker symbol. Then, using Prop. 4.1, we can write

$$U = \int_{\mathbb{R}^2} \mathcal{K}(\xi, \gamma) M(\xi, \gamma) d\xi d\gamma, \quad F_x(U) = \int_{\mathbb{R}^2} \xi \mathcal{K}(\xi, \gamma) M(\xi, \gamma) d\xi d\gamma, \quad (73)$$

$$F_y(U) = \int_{\mathbb{R}^2} \gamma \mathcal{K}(\xi, \gamma) M(\xi, \gamma) d\xi d\gamma, \quad S_e(U) = \int_{\mathbb{R}^2} \mathcal{K}(\xi, \gamma) N(\xi, \gamma) d\xi d\gamma, \quad (74)$$

with  $M(\xi, \gamma) = (M(U_1, \xi, \gamma), \dots, M(U_N, \xi, \gamma))^T$  and

$$N(\xi, \gamma) = \begin{pmatrix} N_{3/2}(\xi, \gamma) - N_{1/2}(\xi, \gamma) \\ \vdots \\ N_{N+1/2}(\xi, \gamma) - N_{N-1/2}(\xi, \gamma) \end{pmatrix}.$$

Hence, using the above notations, the layer-averaged Euler system (23)-(24) can be written under the form

$$\int_{\mathbb{R}^2} \mathcal{K}(\xi, \gamma) \left( \frac{\partial M(\xi, \gamma)}{\partial t} + \xi \frac{\partial M(\xi, \gamma)}{\partial x} + \gamma \frac{\partial M(\xi, \gamma)}{\partial y} - g \nabla_{x,y} z_b \cdot \nabla_{\xi,\gamma} M - N(\xi, \gamma) \right) d\xi d\gamma = 0.$$

## 5 Numerical scheme

The numerical scheme proposed in this section extends the results presented by some of the authors in [7, 11, 10, 6]. The objective is to propose – based on the kinetic interpretation obtained in prop. 4.1 – an accurate, robust and efficient numerical approximation of the system given in prop. 3.5 and that can be written under the form (53). First, we focus on the Euler part of the system (49)-(50) then in paragraph 5.6, a numerical scheme for the viscous terms is proposed.

Notice that, as a consequence of the layer-averaged discretization, the system (49)-(50) and the Boltzmann type equation (70) are only 2d  $(x, y)$  partial differential equations with source terms. Hence, the spacial approximation of the considered PDEs is performed on a 2d planar mesh.



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

We consider discrete times  $t^n$  with  $t^{n+1} = t^n + \Delta t^n$ . For the time discretisation of the layer-averaged Navier-Stokes system (53) we adopt the following scheme

$$U^{n+1} = U - \Delta t^n \left( \frac{\partial F_x(U)}{\partial x} + \frac{\partial F_y(U)}{\partial y} - \mathcal{S}_b(U) \right) + \Delta t^n S_e^{n+1} + \Delta t^n S_{v,f}^{n+p}, \quad (75)$$

where the superscript  $n$  has been omitted and the integer  $p = 0, 1/2, 1$  will be precised below.

Using the expressions (34)-(35) for the layer averaged model, the semi-discrete in time scheme (75) writes

$$h^{n+1} = h^{n+1/2} = h - \Delta t^n \sum_{\alpha=1}^N \nabla_{x,y} \cdot (h_\alpha \mathbf{u}_\alpha), \quad (76)$$

$$(h_\alpha \mathbf{u}_\alpha)^{n+1/2} = h_\alpha \mathbf{u}_\alpha - \Delta t^n \left( \nabla_{x,y} \cdot (h_\alpha \mathbf{u}_\alpha \otimes \mathbf{u}_\alpha) + \nabla_{x,y} \left( \frac{g}{2} h h_\alpha \right) + g h_\alpha \nabla_{x,y} z_b \right), \quad (77)$$

$$\begin{aligned} (h_\alpha \mathbf{u}_\alpha)^{n+1} &= (h_\alpha \mathbf{u}_\alpha)^{n+1/2} - \Delta t^n \left( \mathbf{u}_{\alpha+1/2}^{n+1} G_{\alpha+1/2} - \mathbf{u}_{\alpha-1/2}^{n+1} G_{\alpha-1/2} \right. \\ &\quad + \mu \nabla_{x,y} \cdot (h_\alpha^{n+p} \mathbf{T}_\alpha^{n+p}) - \mathbf{T}_{\alpha+1/2}^{n+p} \nabla_{x,y} z_{\alpha+1/2}^{n+p} + \mathbf{T}_{\alpha-1/2}^{n+p} \nabla_{x,y} z_{\alpha-1/2}^{n+p} \\ &\quad \left. + \mu_{\alpha+1/2} \frac{\mathbf{u}_{\alpha+1}^{n+p} - \mathbf{u}_\alpha^{n+p}}{h_{\alpha+1}^{n+p} + h_\alpha^{n+p}} - \mu_{\alpha-1/2} \frac{\mathbf{u}_\alpha^{n+p} - \mathbf{u}_{\alpha-1}^{n+p}}{h_\alpha^{n+p} + h_{\alpha-1}^{n+p}} - \kappa_\alpha \mathbf{u}_\alpha^{n+p} \right), \end{aligned} \quad (78)$$

$$G_{\alpha+1/2} = - \sum_{j=1}^N \left( \sum_{p=1}^\alpha l_p - \mathbf{1}_{j \leq \alpha} \right) \nabla_{x,y} \cdot (h_j \mathbf{u}_j), \quad (79)$$

for  $\alpha = 1, \dots, N$ . The vertical velocities  $\{w_\alpha\}_{\alpha=1}^N$  are defined by (31). The first two equations (76)-(77) consist in an explicit time scheme where the horizontal fluxes and the topography source term are taken into account whereas in Eq. (78) an implicit treatment of the exchange terms between layers is proposed. The implicit part of the scheme requires to solve a linear problem (see lemma. 5.1) but, on the contrary of previous work of some of the authors [11], it implies that the CFL condition (105) no more depends on the exchange terms.

When  $\mu = \kappa = 0$  in Eq. (78), Eqs. (76)-(78) correspond to the layer-averaged of the Euler system. The choice  $p = 1$  (resp.  $p = 1/2$ ) in Eq. (78) corresponds to an implicit (resp. semi-implicit) treatment of the viscous and friction terms whereas the choice  $p = 0$  implies an explicit treatment and requires a CFL condition. Notice that the advantages and limitations of an implicit or explicit discretization in time scheme for the viscous and friction parts of Eq. (78) are not detailed here.

### 5.2 Space discretisation

Let  $\Omega$  denote the computational domain with boundary  $\Gamma$ , which we assume is polygonal. Let  $T_h$  be a triangulation of  $\Omega$  for which the vertices are denoted by  $P_i$  with  $S_i$  the set of interior nodes and  $G_i$  the set of boundary nodes.

For the space discretization of the system (76)-(78), we use a finite volume technique for the Euler part – that is described below – and a finite element approach  $\mathbb{P}_1$  on  $T_h$  – for the viscous part that is described in paragraph 5.6.

### 5.3 Finite volume formalism for the Euler part

In this paragraph and in paragraph (5.4), we propose a space discretization for the model (76)-(78) without the viscous and friction terms i.e. the system

$$h^{n+1} = h^{n+1/2} = h - \Delta t^n \sum_{\alpha=1}^N \nabla_{x,y} \cdot (h_\alpha \mathbf{u}_\alpha), \quad (80)$$

$$(h_\alpha \mathbf{u}_\alpha)^{n+1/2} = h_\alpha \mathbf{u}_\alpha - \Delta t^n \left( \nabla_{x,y} \cdot (h_\alpha \mathbf{u}_\alpha \otimes \mathbf{u}_\alpha) + \nabla_{x,y} \left( \frac{g}{2} h h_\alpha \right) + g h_\alpha \nabla_{x,y} z_b \right), \quad (81)$$

$$(h_\alpha \mathbf{u}_\alpha)^{n+1} = (h_\alpha \mathbf{u}_\alpha)^{n+1/2} - \Delta t^n \left( \mathbf{u}_{\alpha+1/2}^{n+1} G_{\alpha+1/2} - \mathbf{u}_{\alpha-1/2}^{n+1} G_{\alpha-1/2} \right), \quad (82)$$

completed with (79).

We recall now the general formalism of finite volumes on unstructured meshes.

The dual cells  $C_i$  are obtained by joining the centers of mass of the triangles surrounding each vertex  $P_i$ . We use the following notations (see Fig. 2):

- $K_i$ , set of subscripts of nodes  $P_j$  surrounding  $P_i$ ,
- $|C_i|$ , area of  $C_i$ ,
- $\Gamma_{ij}$ , boundary edge between the cells  $C_i$  and  $C_j$ ,
- $L_{ij}$ , length of  $\Gamma_{ij}$ ,
- $\mathbf{n}_{ij}$ , unit normal to  $\Gamma_{ij}$ , outward to  $C_i$  ( $\mathbf{n}_{ji} = -\mathbf{n}_{ij}$ ).

If  $P_i$  is a node belonging to the boundary  $\Gamma$ , we join the centers of mass of the triangles adjacent to the boundary to the middle of the edge belonging to  $\Gamma$  (see Fig. 2) and we denote

- $\Gamma_i$ , the two edges of  $C_i$  belonging to  $\Gamma$ ,
- $L_i$ , length of  $\Gamma_i$  (for sake of simplicity we assume in the following that  $L_i = 0$  if  $P_i$  does not belong to  $\Gamma$ ),
- $\mathbf{n}_i$ , the unit outward normal defined by averaging the two adjacent normals.

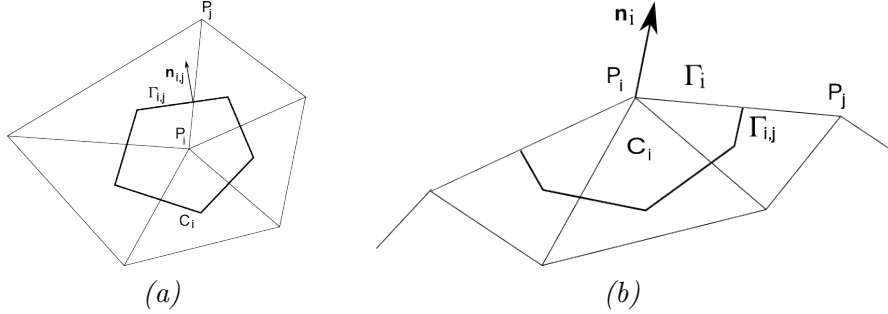


Figure 2: (a) Dual cell  $C_i$  and (b) Boundary cell  $C_i$ .

We define the piecewise constant functions  $U^n(x, y)$  on cells  $C_i$  corresponding to time  $t^n$  and  $z_b(x, y)$  as

$$U^n(x, y) = U_i^n, \quad z_b(x, y) = z_i, \quad \text{for } (x, y) \in C_i, \quad (83)$$

with  $U_i^n = (h_i^n, q_{x,1,i}^n, \dots, q_{x,N,i}^n, q_{y,1,i}^n, \dots, q_{y,N,i}^n)^T$  i.e.

$$U_i^n \approx \frac{1}{|C_i|} \int_{C_i} U(t^n, x, y) dx dy, \quad z_i \approx \frac{1}{|C_i|} \int_{C_i} z_b(x, y) dx dy.$$

We will also use the notation

$$U_{\alpha,i}^n \approx \frac{1}{|C_i|} \int_{C_i} U_\alpha(t^n, x, y) dx dy,$$

with  $U_\alpha$  defined by (62). A finite volume scheme for solving the system (80)-(81) is a formula of the form

$$U_i^{n+1/2} = U_i - \sum_{j \in K_i} \sigma_{i,j} \mathcal{F}_{i,j} - \sigma_i \mathcal{F}_{e,i}, \quad (84)$$

where using the notations of (75)

$$\sum_{j \in K_i} L_{i,j} \mathcal{F}_{i,j} \approx \int_{C_i} \left( \frac{\partial F_x(U)}{\partial x} + \frac{\partial F_y(U)}{\partial y} \right) dx dy, \quad (85)$$

with

$$\sigma_{i,j} = \frac{\Delta t^n L_{i,j}}{|C_i|}, \quad \sigma_i = \frac{\Delta t^n L_i}{|C_i|}.$$

Here we consider first-order explicit schemes where

$$\mathcal{F}_{i,j} = F(U_i, U_j, z_{b,i} - z_{b,j}, \mathbf{n}_{i,j}). \quad (86)$$

and

$$\mathcal{F}_{i,j} = F(U_i, U_j, z_{b,i} - z_{b,j}, \mathbf{n}_{i,j}) = \begin{pmatrix} F(U_{1,i}, U_{1,j}, z_{b,i} - z_{b,j}, \mathbf{n}_{i,j}) \\ \vdots \\ F(U_{N,i}, U_{N,j}, z_{b,i} - z_{b,j}, \mathbf{n}_{i,j}) \end{pmatrix} \quad (87)$$

and for the boundary nodes

$$\mathcal{F}_{e,i} = F(U_i, U_{e,i}, \mathbf{n}_i) = \begin{pmatrix} F(U_{1,i}, U_{1,e,i}, \mathbf{n}_i) \\ \vdots \\ F(U_{N,i}, U_{N,e,i}, \mathbf{n}_i) \end{pmatrix}. \quad (88)$$

Relation (84) tells how to compute the values  $U_i^{n+1/2}$  knowing  $U_i$  and discretized values  $z_{b,i}$  of the topography. Following (85), the term  $\mathcal{F}_{i,j}$  in (84) denotes an interpolation of the normal component of the flux  $F(U) \cdot \mathbf{n}_{i,j}$  along the edge  $C_{i,j}$ . The functions  $F(U_i, U_j, z_{b,i} - z_{b,j}, \mathbf{n}_{i,j}) \in \mathbb{R}^{2N+1}$  are the numerical fluxes, see [18].

In the next paragraph we define  $\mathcal{F}(U_i, U_j, z_{b,i} - z_{b,j}, \mathbf{n}_{i,j})$  using the kinetic interpretation of the system. The computation of the value  $U_{i,e}$ , which denotes a value outside  $C_i$  (see Fig. 2-(b)), defined such that the boundary conditions are satisfied, and the definition of the boundary flux  $F(U_i, U_{e,i}, \mathbf{n}_i)$  are described paragraph 5.7. Notice that we assume a flat topography on the boundaries i.e.  $z_{b,i} = z_{b,i,e}$ .

## 5.4 Discrete kinetic equation

The choice of a kinetic scheme is motivated by several arguments. First, the kinetic interpretation is a suitable starting point for building a stable numerical scheme. We will prove in paragraph 5.4 that the proposed kinetic scheme preserves positivity of the water depth and ensures a discrete local maximum principle for a tracer concentration (temperature, salinity...). Second, the construction of the kinetic scheme does not need the computation of the system eigenvalues. This point is very important here since these eigenvalues are not available in explicit analytical form, and they are hardly accessible even numerically. Furthermore, as previously mentioned, hyperbolicity of the multilayer model may not hold, and the kinetic scheme allows overcoming this difficulty.

### 5.4.1 Without topography

In a first step we consider a situation with flat bottom. Following prop. 4.1, the model (23)-(24) reduces, for each layer, to a classical Saint-Venant system with exchange terms and its kinetic interpretation (see Eq. (70)) is given by

$$\frac{\partial M_\alpha}{\partial t} + \xi \frac{\partial M_\alpha}{\partial x} + \gamma \frac{\partial M_\alpha}{\partial y} - N_{\alpha+1/2} + N_{\alpha-1/2} = Q_\alpha, \quad \alpha \in \{1, \dots, N\}, \quad (89)$$

with the notations defined in paragraph 4.2.

Let  $C_i$  be a cell, see Fig. 2. The integral over  $C_i$  of the convective part of the kinetic equation (89) gives

$$\int_{C_i} \left( \frac{\partial M_\alpha}{\partial t} + \xi \frac{\partial M_\alpha}{\partial x} + \gamma \frac{\partial M_\alpha}{\partial y} \right) dx dy \approx |C_i| \frac{\partial M_{\alpha,i}}{\partial t} + \sum_{j \in K_i} \int_{\Gamma_{i,j}} M_{\alpha,i,j} \begin{pmatrix} \xi \\ \gamma \end{pmatrix} \cdot \mathbf{n}_{i,j} dl, \quad (90)$$

with  $M_{\alpha,i} = M(U_{\alpha,i}, \xi, \gamma)$ ,  $\mathbf{n}_{i,j}$  being the outward normal to the cell  $C_i$ . The quantity  $M_{\alpha,i,j}$  is defined by the classical kinetic upwinding

$$M_{\alpha,i,j} = M_{\alpha,i} \zeta_{i,j} \mathbf{1}_{\zeta_{i,j} \geq 0} + M_{\alpha,j} \zeta_{i,j} \mathbf{1}_{\zeta_{i,j} \leq 0},$$

with

$$\zeta_{i,j} = \begin{pmatrix} \xi \\ \gamma \end{pmatrix} \cdot \mathbf{n}_{i,j}.$$

Therefore, the kinetic scheme applied for Eq.(89) is given by

$$f_{\alpha,i}^{n+1/2-} = \left( 1 - \frac{\Delta t^n}{|C_i|} \sum_{j \in K_i} L_{i,j} \zeta_{i,j} \mathbf{1}_{\zeta_{i,j} \geq 0} \right) M_{\alpha,i} - \frac{\Delta t^n}{|C_i|} \sum_{j \in K_i} L_{i,j} M_{\alpha,j} \zeta_{i,j} \mathbf{1}_{\zeta_{i,j} \leq 0}, \quad (91)$$

$$f_{\alpha,i}^{n+1-} = f_{\alpha,i}^{n+1/2-} + \Delta t^n \left( N_{\alpha+1/2,i}^{n+1-} - N_{\alpha-1/2,i}^{n+1-} \right), \quad (92)$$

with the exchange terms  $\{N_{\alpha+1/2,i}^{n+1-}\}_{\alpha=0}^N$  defined by

$$N_{\alpha+1/2,i}^{n+1-}(\xi, \gamma) = \frac{G_{\alpha+1/2,i}}{h_i} f_{\alpha+1/2,i}^{n+1-}. \quad (93)$$

Following (26) we can write

$$f_{\alpha+1/2,i}^{n+1-} = \begin{cases} f_{\alpha,i}^{n+1-} & \text{if } G_{\alpha+1/2} \leq 0 \\ f_{\alpha+1,i}^{n+1-} & \text{if } G_{\alpha+1/2} > 0 \end{cases}$$

leading to

$$N_{\alpha+1/2,i}^{n+1-}(\xi, \gamma) = \frac{|G_{\alpha+1/2,i}|_+}{h_i} f_{\alpha+1,i}^{n+1-} - \frac{|G_{\alpha+1/2,i}|_-}{h_i} f_{\alpha,i}^{n+1-}.$$

Notice that the previous definition is consistent with (67). From (69), we get

$$G_{\alpha+1/2,i} = -\frac{1}{|C_i|} \sum_{k=1}^N \left( \sum_{p=1}^\alpha l_p - \mathbf{1}_{k \leq \alpha} \right) \sum_{j \in K_i} L_{i,j} \int_{\mathbb{R}^2} (M_{k,i} \zeta_{i,j} \mathbf{1}_{\zeta_{i,j} \geq 0} + M_{k,j} \zeta_{i,j} \mathbf{1}_{\zeta_{i,j} \leq 0}) d\xi d\gamma.$$

By analogy with the computations in (64), we can recover the macroscopic quantities  $U_{\alpha,i}^{n+1}$  at time  $t^{n+1}$  by integration of the relation (92)

$$U_{\alpha,i}^{n+1} = \int_{\mathbb{R}^2} \begin{pmatrix} 1 \\ \xi \\ \gamma \end{pmatrix} f_{\alpha,i}^{n+1-} d\xi d\gamma. \quad (94)$$

The scheme (91) and the definition (94) allow to complete the definition of the macroscopic scheme (84),(87), (88) with the numerical flux given by the flux vector splitting formula [16]

$$\begin{aligned} \mathcal{F}_{i,j} &= F^+(U_i^n, \mathbf{n}_{i,j}) + F^-(U_j^n, \mathbf{n}_{i,j}) \\ &= \int_{\mathbb{R}^2} \mathcal{K}(\xi, \gamma) M_i \zeta_{i,j} \mathbf{1}_{\zeta_{i,j} \geq 0} d\xi d\gamma + \int_{\mathbb{R}^2} \mathcal{K}(\xi, \gamma) M_j \zeta_{i,j} \mathbf{1}_{\zeta_{i,j} \leq 0} d\xi d\gamma, \end{aligned} \quad (95)$$

where  $\mathcal{K}(\xi, \gamma)$  is defined in Remark 4.3 and  $M_i = (M_{1,i}, \dots, M_{N,i})^T$ .

Using (93), we rewrite the step (92) under the form

$$(\mathbf{I}_N + \Delta t \mathbf{G}_{N,i}) f^{n+1-} = f^{n+1/2-},$$

where  $\mathbf{I}_N$  is the identity matrix of size  $N$  and  $\mathbf{G}_{N,i}$  is defined by

$$\mathbf{G}_{N,i} = \begin{pmatrix} -\frac{|G_{3/2,i}|_-}{h_{1,i}^{n+1}} & -\frac{|G_{3/2,i}|_+}{h_{1,i}^{n+1}} & 0 & 0 & \dots & 0 \\ \frac{|G_{3/2,i}|_-}{h_{2,i}^{n+1}} & \ddots & \ddots & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & 0 & 0 \\ \vdots & 0 & \frac{|G_{\alpha-1/2,i}|_-}{h_{\alpha,i}^{n+1}} & -\frac{|G_{\alpha+1/2,i}|_- - |G_{\alpha-1/2,i}|_+}{h_{\alpha,i}^{n+1}} & -\frac{|G_{\alpha+1/2,i}|_+}{h_{\alpha,i}^{n+1}} & 0 \\ \vdots & \ddots & 0 & \ddots & \ddots & -\frac{|G_{N-1/2,i}|_+}{h_{N-1,i}^{n+1}} \\ 0 & \dots & 0 & 0 & \frac{|G_{N-1/2,i}|_-}{h_{N,i}^{n+1}} & \frac{|G_{N-1/2,i}|_+}{h_{N,i}^{n+1}} \end{pmatrix}.$$

Hence, the resolution of the discrete kinetic equation (92) requires to inverse the matrix

$$\begin{pmatrix} \mathbf{I}_N + \Delta t \mathbf{G}_{N,i} & 0 \\ 0 & \mathbf{I}_N + \Delta t \mathbf{G}_{N,i} \end{pmatrix}$$

and we have the following lemma.

**Lemma 5.1** *The matrix  $\mathbf{I}_N + \Delta t \mathbf{G}_{N,i}$*

- (i) *is invertible for any  $h_i^{n+1} > 0$ ,*
- (ii)  *$(\mathbf{I}_N + \Delta t \mathbf{G}_{N,i})^{-1}$  has only positive coefficients,*
- (iii) *for any vector  $T$  with non negative entries i.e.  $T_\alpha \geq 0$ , for  $1 \leq \alpha \leq N$ , one has*

$$\|(\mathbf{I}_N + \Delta t \mathbf{G}_{N,i})^{-t} T\|_\infty \leq \|T\|_\infty.$$

**Proof of lemma 5.1** (i) *For any  $h_i^{n+1} > 0$ , the matrix  $\mathbf{I}_N + \Delta t \mathbf{G}_{N,i}$  is a strictly dominant diagonal matrix and hence it is invertible.*

(ii) *Denoting  $\mathbf{G}_{N,i}^d$  (resp.  $\mathbf{G}_{N,i}^{nd}$ ) the diagonal (resp. non diagonal) part of  $\mathbf{G}_{N,i}$  we can write*

$$\mathbf{I}_N + \Delta t \mathbf{G}_{N,i} = (\mathbf{I}_N + \Delta t \mathbf{G}_{N,i}^d) \left( \mathbf{I}_N - (\mathbf{I}_N + \Delta t \mathbf{G}_{N,i}^d)^{-1} (-\Delta t \mathbf{G}_{N,i}^{nd}) \right),$$

*where all the entries of the matrix  $\mathbf{J}_{N,i} = (\mathbf{I}_N + \Delta t \mathbf{G}_{N,i}^d)^{-1} (-\Delta t \mathbf{G}_{N,i}^{nd})$ , are non negative and less than 1. And hence, we can write*

$$(\mathbf{I}_N + \Delta t \mathbf{G}_{N,i})^{-1} = \sum_{k=0}^{\infty} \mathbf{J}_{N,i}^k,$$

*proving all the entries of  $(\mathbf{I}_N + \Delta t \mathbf{G}_{N,i})^{-1}$  are non negative.*

(ii) *Let us consider the vector  $\mathbf{1}$  whose entries are all equal to 1. Since we have*

$$(\mathbf{I}_N + \Delta t \mathbf{G}_{N,i})^t \mathbf{1} = \mathbf{1},$$

*we also have  $\mathbf{1} = (\mathbf{I}_N + \Delta t \mathbf{G}_{N,i})^{-t} \mathbf{1}$ . Now let  $T$  be a vector whose entries  $\{T_\alpha\}_{1 \leq \alpha \leq N}$  are non negative, then*

$$(\mathbf{I}_N + \Delta t \mathbf{G}_{N,i})^{-t} \mathbf{T} \leq (\mathbf{I}_N + \Delta t \mathbf{G}_{N,i})^{-t} \mathbf{1} \|\mathbf{T}\|_\infty = \mathbf{1} \|\mathbf{T}\|_\infty,$$

*that completes the proof. ■*

### 5.4.2 With topography

The hydrostatic reconstruction scheme (HR scheme for short) for the Saint-Venant system has been introduced in [5] in the 1d case and described in 2d for unstructured meshes in [7]. The HR in the context of the kinetic description for the Saint-Venant system has been studied in [6].

In order to take into account the topography source and to preserve relevant equilibria, the HR leads to a modified version of (84) under the form

$$U_i^{n+1/2} = U_i^n - \sum_{j \in K_i} \sigma_{i,j} \mathcal{F}_{i,j}^* - \sigma_i \mathcal{F}_{i,e} + \sum_{j \in K_i} \sigma_{i,j} \mathcal{S}_{i,j}^*, \quad (96)$$

where

$$\mathcal{F}_{i,j}^* = F(U_{i,j}^*, U_{j,i}^*, \mathbf{n}_{i,j}), \quad \mathcal{S}_{i,j} = S(U_i, U_{i,j}^*, \mathbf{n}_{i,j}) = \begin{pmatrix} 0 \\ \frac{g}{2} l_1 (h_{i,j}^{*2} - h_i^2) \mathbf{n}_{i,j} \\ \vdots \\ \frac{g}{2} l_N (h_{i,j}^{*2} - h_i^2) \mathbf{n}_{i,j} \end{pmatrix}, \quad (97)$$

with

$$\begin{aligned} z_{b,i,j}^* &= \max(z_{b,i}, z_{b,j}), \quad h_{i,j}^* = \max(h_i + z_{b,i} - z_{b,i,j}^*, 0), \\ U_{i,j}^* &= (h_{i,j}^*, l_1 h_{i,j}^* u_{1,i}, \dots, l_N h_{i,j}^* u_{N,i}, l_1 h_{i,j}^* v_{1,i}, \dots, l_N h_{i,j}^* v_{N,i})^T. \end{aligned} \quad (98)$$

We would like here to propose a kinetic interpretation of the HR scheme, which means to interpret the above numerical fluxes as averages with respect to the kinetic variables of a scheme written on a kinetic function  $f$ . More precisely, we would like to approximate the solution to (70) by a kinetic scheme such that the associated macroscopic scheme is exactly (96)-(97) with homogeneous numerical flux  $\mathcal{F}$  given by (95). We denote  $M_{\alpha,i,j}^* = M(U_{\alpha,i,j}^*, \xi, \gamma)$  for any  $\alpha = 1, \dots, N$  and we consider the scheme

$$\begin{aligned} f_{\alpha,i}^{n+1/2-} &= M_{\alpha,i} - \frac{\Delta t^n}{|C_i|} \sum_{j \in K_i} L_{i,j} \zeta_{i,j} \mathbf{1}_{\zeta_{i,j} \geq 0} M_{\alpha,i,j}^* - \frac{\Delta t^n}{|C_i|} \sum_{j \in K_i} L_{i,j} M_{\alpha,j,i}^* \zeta_{i,j} \mathbf{1}_{\zeta_{i,j} \leq 0}, \\ &\quad - \frac{\Delta t^n}{|C_i|} \sum_{j \in K_i} L_{i,j} (M_{\alpha,i} - M_{\alpha,i,j}^*) \theta_{\alpha,i,j}, \end{aligned} \quad (99)$$

$$f_{\alpha,i}^{n+1-} = f_{\alpha,i}^{n+1/2-} + \Delta t^n (N_{\alpha+1/2,i}^{*,n+1-} - N_{\alpha-1/2,i}^{*,n+1-}), \quad (100)$$

where

$$\theta_{\alpha,i,j} = \begin{pmatrix} \xi - u_{\alpha,i} \\ \gamma - v_{\alpha,i} \end{pmatrix} \cdot \mathbf{n}_{i,j}.$$

For the exchange terms, by analogy with (93) we define

$$N_{\alpha+1/2,i}^{*,n+1-}(\xi, \gamma) = \frac{G_{\alpha+1/2,i}^*}{h_i} f_{\alpha+1/2,i}^{n+1-}, \quad (101)$$

and using (69) we get

$$G_{\alpha+1/2,i}^* = -\frac{1}{|C_i|} \sum_{k=1}^N \left( \sum_{p=1}^{\alpha} l_p - \mathbf{1}_{k \leq \alpha} \right) \sum_{j \in K_i} L_{i,j} \int_{\mathbb{R}^2} (M_{k,i,j}^* \zeta_{i,j} \mathbf{1}_{\zeta_{i,j} \geq 0} + M_{k,i,j}^* \zeta_{i,j} \mathbf{1}_{\zeta_{i,j} \leq 0}) d\xi d\gamma.$$

It is easy to see that in the previous formula, we have the moment relations

$$\int_{\mathbb{R}^2} (M_{\alpha,i} - M_{\alpha,i,j}^*) \theta_{\alpha,i,j} d\xi d\gamma = 0, \quad (102)$$

$$\int_{\mathbb{R}^2} \begin{pmatrix} \xi \\ \gamma \end{pmatrix} (M_{\alpha,i} - M_{\alpha,i,j}^*) \theta_{\alpha,i,j} d\xi d\gamma = \frac{g}{2} l_{\alpha} (h_{i,j}^{*2} - h_i^2) \mathbf{n}_{i,j}, \quad (103)$$

Using again (94), the integration of the set of equations (99)-(100), for  $\alpha = 1, \dots, N$ , multiplied by  $\mathcal{K}(\xi, \gamma)$  with respect to  $\xi, \gamma$  then gives the HR scheme (96)-(97) with (95), (98). Thus as announced, (99)-(100) is a kinetic interpretation of the HR scheme in 3d for an unstructured mesh.

There exists a velocity  $v_m \geq 0$  such that for all  $\alpha, i$ ,

$$|\xi| \geq v_m \text{ or } |\gamma| \geq v_m \Rightarrow M(U_{\alpha,i}, \xi, \gamma) = 0. \quad (104)$$

This means equivalently that  $|u_{\alpha,i}| + |v_{\alpha,i}| + \sqrt{2gh_i} \leq v_m$ . We consider a CFL condition strictly less than one,

$$\sigma_i v_m \leq \beta < \frac{1}{2} \quad \text{for all } i, \quad (105)$$

where  $\sigma_i = \Delta t^n \sum_{j \in K_i} L_{i,j} / |C_i|$ , and  $\beta$  is a given constant.

Then the following proposition holds.

**Proposition 5.1** *Under the CFL condition (105), the scheme (99)-(100) verifies the following properties.*

(i) *The macroscopic scheme derived from (99)-(100) using (94) is a consistent discretization of the layer-averaged Euler system (23)-(24).*

(ii) *The kinetic function remains nonnegative i.e.*

$$f_{\alpha,i}^{n+1-} \geq 0, \quad \forall (\xi, \gamma) \in \mathbb{R}^2, \forall i, \forall \alpha.$$

(iii) *The scheme (99)-(100) is kinetic well balanced i.e. at rest*

$$f_{\alpha,i}^{n+1-} = M_{\alpha,i}, \quad \forall (\xi, \gamma) \in \mathbb{R}^2, \forall i, \forall \alpha = 1, \dots, N. \quad (106)$$

**Proof of prop. 5.1** (i) *Since the Boltzmann type equations (70) are almost linear transport equations with source terms, the discrete kinetic scheme (99)-(100) is clearly a consistent discretization of (70). And therefore using the kinetic interpretation given in prop. 4.1, the macroscopic scheme obtained from (99)-(100) using (94) is a consistent discretization of the layer-averaged Euler system (23)-(24).*

(ii) *In (99)-(100) we have*

$$\frac{\Delta t^n}{|C_i|} \sum_{j \in K_i} L_{i,j} M_{\alpha,j,i}^* \zeta_{i,j} \mathbf{1}_{\zeta_{i,j} \leq 0} \leq 0,$$

and the HR (98) ensures

$$M_{\alpha,i,j}^* \leq M_{\alpha,i}, \quad \forall (\xi, \gamma) \in \mathbb{R}^2, \forall \alpha$$

leading to

$$f_{\alpha,i}^{n+1/2-} \geq \left( 1 - \frac{\Delta t^n}{|C_i|} \sum_{j \in K_i} L_{i,j} (\zeta_{i,j} \mathbf{1}_{\zeta_{i,j} \geq 0} + \theta_{\alpha,i,j} \mathbf{1}_{\theta_{\alpha,i,j} \geq 0}) \right) M_{\alpha,i},$$

But  $\zeta_{i,j} \mathbf{1}_{\zeta_{i,j} \geq 0} \leq \max\{|\xi|, |\gamma|\}$ ,  $\theta_{\alpha,i,j} \mathbf{1}_{\theta_{\alpha,i,j} \geq 0} \leq \max\{|\xi - u_{\alpha,i}|, |\gamma - v_{\alpha,i}|\}$  and therefore

$$\frac{\Delta t^n}{|C_i|} \sum_{j \in K_i} L_{i,j} (\zeta_{i,j} \mathbf{1}_{\zeta_{i,j} \geq 0} + \theta_{\alpha,i,j} \mathbf{1}_{\theta_{\alpha,i,j} \geq 0}) \leq \sigma_i (\max\{|\xi|, |\gamma|\} + \max\{|\xi - u_{\alpha,i}|, |\gamma - v_{\alpha,i}|\}) \leq 1,$$

where (104), (105) have been used, proving  $f_{\alpha,i}^{n+1/2-} \geq 0$  for any  $\xi \in \mathbb{R}$  and any  $\alpha \in \{1, \dots, N\}$ . Now using the results of lemma 5.1, it ensures that  $f_{\alpha,i}^{n+1-}$  defined by (100) satisfies  $f_{\alpha,i}^{n+1-} \geq 0$  for any  $(\xi, \gamma) \in \mathbb{R}^2$  and any  $\alpha \in \{1, \dots, N\}$ , proving (ii).

(iii) *Considering the situation at rest i.e.  $u_{\alpha,i} = v_{\alpha,i} = 0$ ,  $\forall \alpha, i$  and  $h_i + z_{b,i} = h_j + z_{b,j}$ ,  $\forall i, j$  we have*

$$M_{\alpha,i} = M_{\alpha,i,j}^*, \quad \forall \alpha, i, j.$$

From (99)-(100), this gives (106). ■

## 5.5 Macroscopic scheme

The numerical scheme for the system (80)-(82) is given by (94),(99), (100) and requires to calculate fluxes having the form

$$F_{sv}(U) = \begin{pmatrix} F_h \\ F_{hu} \\ F_{hv} \end{pmatrix} = \int_{n_x \xi + n_y \gamma \geq 0} \begin{pmatrix} 1 \\ \xi \\ \gamma \end{pmatrix} (n_x \xi + n_y \gamma) M(U, \xi, \gamma) d\xi d\gamma.$$

with  $M$  given by (63),  $n_x$  and  $n_y$  being the components of a normal unit vector  $\mathbf{n}$ . Defining the change of variables

$$\xi = u + cz_1, \quad \gamma = v + cz_2,$$

we can write

$$F_{sv}(U) = h \int_{n_x(cz_1+u)+n_y(cz_2+v) \geq 0} (n_x(cz_1+u) + n_y(cz_2+v)) \begin{pmatrix} 1 \\ u + cz_1 \\ v + cz_2 \end{pmatrix} \chi_0(z_1, z_2) dz_1 dz_2,$$

where  $\chi_0$  is defined by (60). A second change of variables

$$y_1 = n_x z_1 + n_y z_2, \quad y_2 = n_x z_2 - n_y z_1, \quad \tilde{u} = n_x u + n_y v.$$

gives

$$F_{sv}(U) = h \int_{\{y_1 \geq -\frac{\tilde{u}}{c}\} \times \mathbb{R}} (\tilde{u} + cy_1) \begin{pmatrix} 1 \\ u + cn_x y_1 \\ v + cn_y y_1 \end{pmatrix} \chi_0(y_1, y_2) dy_1 dy_2, \quad (107)$$

since  $\chi_0$  is odd. The details of the computations of formula (107) is given in appendix B.

Using the properties obtained at the kinetic level for the resolution of the system (23)-(24), the following proposition holds.

**Proposition 5.2** *Under the CFL condition (105), the scheme (94),(99), (100) satisfies the following properties.*

- (i) *The macroscopic scheme derived from (99)-(100) using (94) is a consistent discretization of the layer-averaged Euler system (23)-(24).*
- (ii) *The water depth remains nonnegative i.e.*

$$h_i^{n+1} \geq 0, \quad \forall i, \quad \text{when } h_i^n \geq 0 \quad \forall i.$$

- (iii) *The scheme (94),(99), (100) is well-balanced.*

**Proof of prop. 5.2** *The proof is similar to the one given in prop. 5.1. ■*

## 5.6 The discrete layer-averaged Navier-Stokes system

In this paragraph, we detail the space discretisation of the viscous terms. Several expressions have been obtained for the viscous terms, see paragraph 3.2. In this paragraph, we do not give a numerical scheme for the model (78) but for the simplified expression depicted in Proposition 3.5.

Rewriting (78) under the form

$$(h_\alpha \mathbf{u}_\alpha)^{n+1} = \widetilde{h_\alpha \mathbf{u}_\alpha} + \Delta t^n S_{v,f}(U), \quad (108)$$

with  $S_{v,f}(U) = (S_{v,f,1}, \dots, S_{v,f,N})^T$  and

$$\widetilde{h_\alpha \mathbf{u}_\alpha} = h_\alpha \mathbf{u}_\alpha - \Delta t^n \left( \nabla_{x,y} \cdot (h_\alpha \mathbf{u}_\alpha \otimes \mathbf{u}_\alpha) - \nabla_{x,y} \left( \frac{g}{2} h h_\alpha \right) - g h_\alpha \nabla_{x,y} z_b \right)$$



$$\begin{aligned}
& + \mathbf{u}_{\alpha+1/2}^{n+1} G_{\alpha+1/2} - \mathbf{u}_{\alpha-1/2}^{n+1} G_{\alpha-1/2} \Big), \\
S_{v,f,\alpha} &= \nabla_{x,y} \cdot (h_\alpha \mathbf{T}_\alpha^0) + \Gamma_{\alpha+1/2} (\mathbf{u}_{\alpha+1} - \mathbf{u}_\alpha) - \Gamma_{\alpha-1/2} (\mathbf{u}_\alpha - \mathbf{u}_{\alpha-1}) - \kappa_\alpha \mathbf{u}_\alpha,
\end{aligned}$$

with the definitions (42),(43),(47), (48),(51) for  $\mathbf{T}_\alpha^0$ ,  $\Gamma_{\alpha\pm 1/2}$ . It remains to give a fully discrete scheme for the viscous and friction terms  $\{S_{v,f,\alpha}\}$ .

The discretization of (108) is done using a finite element / finite difference approximation obtained as follows. We depart from the triangulation defined in paragraph (5.2) and we use the cells values of the variables – inherited from the finite volume framework – to define a  $\mathbb{P}_1$  approximation of the variables.

Notice that, compared to the advection and pressure terms, the discretization of the viscous terms raises less difficulties and we propose a stable scheme that will be extended to more general rheology terms [24] and more completely analysed in a forthcoming paper.

Using a classical  $\mathbb{P}_1$  finite element type approximation with mass lumping of Eq. (108), we get

$$\begin{aligned}
\mathbf{U}_\alpha^{n+1} &= \tilde{\mathbf{U}}_\alpha - \Delta t^n \left( \mathcal{K}_{\alpha+1} \mathbf{U}_{\alpha+1} + \mathcal{K}_\alpha \mathbf{U}_\alpha + \mathcal{K}_{\alpha-1} \mathbf{U}_{\alpha-1} \right) \\
&+ \Delta t^n \mathcal{G}_{\alpha+1/2} (\mathbf{U}_{\alpha+1} - \mathbf{U}_\alpha) - \Delta t^n \mathcal{G}_{\alpha-1/2} (\mathbf{U}_\alpha - \mathbf{U}_{\alpha-1}) - \Delta t^n \kappa_\alpha \mathbf{U}_\alpha,
\end{aligned} \tag{109}$$

with the matrices

$$\begin{aligned}
\mathcal{K}_\alpha &= \frac{1}{2} \int_\Omega \left( \frac{h_{\alpha,j}}{h_{\alpha+1,j} + h_{\alpha,j}} + \frac{h_{\alpha,j}}{h_{\alpha,j} + h_{\alpha-1,j}} \right) \nabla_{x,y} \varphi_i \cdot \nabla_{x,y} \varphi_j \, dx dy, \\
\mathcal{K}_{\alpha\pm 1} &= \frac{\mu_{\alpha\pm 1/2}}{2} \int_\Omega \frac{h_{\alpha\pm 1,j}}{h_{\alpha+1,j} + h_{\alpha,j}} \nabla_{x,y} \varphi_i \cdot \nabla_{x,y} \varphi_j \, dx dy, \\
\mathcal{G}_{\alpha+1/2} &= \mu_{\alpha+1/2} \int_\Omega \frac{1 + |\nabla_{x,y} z_{\alpha+1/2}|^2}{h_{\alpha+1} + h_\alpha} \varphi_i \cdot \varphi_j \, dx dy,
\end{aligned}$$

where the expression of the viscous terms given in prop. 3.5 is used. We have presented an explicit in time version of (109) that is stable under a classical CFL condition. An implicit or semi-implicit version of (109) can also be used.

The main purpose of this paper is to propose a stable and robust numerical approximation of the incompressible Euler system with free surface. Voluntarily, we give few details concerning the numerical approximation of the dissipative terms:

- the viscous and friction terms are dissipative and hence a reasonable approximation leads to a stable numerical scheme.
- In this paper, we consider a simplified newtonian rheology for the fluid (some vorticity terms have been neglected, see paragraph 2.2), the numerical approximation of the general (layer-averaged) rheology [24] will be studied in a forthcoming paper.

## 5.7 Boundary conditions

The contents of this paragraph slightly differ from previous works of one of the authors [26] and valid for the classical Saint-Venant system. First, we focus on the boundary conditions for the layer-averaged Euler system i.e. the system (76)-(77) for  $\mu = 0$ ,  $\kappa = 0$  and then for the viscous part.

### 5.7.1 Layer-averaged Euler system

In this paragraph we detail the computation of the boundary flux  $\mathcal{F}(\mathbf{U}_i, \mathbf{U}_{e,i}, \mathbf{n}_i)$  appearing in (84),(87),(88). The variable  $\mathbf{U}_{i,e}^n$  can be interpreted as an approximation of the solution in a ghost cell adjacent to the boundary. As before we introduce the vector

$$\mathbf{U}_{i,e} = (h_{i,e}^n, (hu)_{1,i,e}, \dots, (hu)_{N,i,e}, (hv)_{1,i,e}, \dots, (hv)_{N,i,e})^T,$$

and we will use the flux vector splitting form associated to the kinetic formulation (95)

$$\mathcal{F}(U_i, U_{i,e}, \mathbf{n}_i) = F^+(U_i, \mathbf{n}_i) + F^-(U_{i,e}, \mathbf{n}_i) \quad (110)$$

with  $U_{i,e}^n$  defined according to the boundary type.

**Solid wall** If we consider a node  $i_0$  belonging to a solid wall, we prescribe a slip condition written

$$\mathbf{u}_\alpha \cdot \mathbf{n}_{i_0} = 0, \quad (111)$$

for  $\alpha = 1, \dots, N$ . We assume the continuity of the water depth  $h_{i_0,e} = h_{i_0}$  and of the tangential component of velocity.

From (107) with (111) we obtain

$$F_h^+(U_{\alpha,i_0}) + F_h^-(U_{\alpha,i_0,e}) = 0,$$

and

$$\left( \frac{F_{hu}^+(U_{\alpha,i_0}) + F_{hu}^-(U_{\alpha,i_0,e})}{F_{hv}^+(U_{\alpha,i_0}) + F_{hv}^-(U_{\alpha,i_0,e})} \right) \cdot \mathbf{n}_{i_0} = \frac{gh_{\alpha,i_0}h_{i_0}}{2}, \quad \left( \frac{F_{hu}^+(U_{\alpha,i_0}) + F_{hu}^-(U_{\alpha,i_0,e})}{F_{hv}^+(U_{\alpha,i_0}) + F_{hv}^-(U_{\alpha,i_0,e})} \right) \cdot \mathbf{t}_{i_0} = 0,$$

for  $\alpha = 1, \dots, N$  for a vector  $\mathbf{t}_{i_0}$  orthogonal to  $\mathbf{n}_{i_0}$ . The condition (111) is therefore prescribed weakly but a posteriori, in order to be sure that (111) is satisfied, we can apply

$$\mathbf{u}_{\alpha,i_0,e} = \mathbf{u}_{\alpha,i_0} - (\mathbf{u}_{\alpha,i_0} \cdot \mathbf{n}_{i_0}) \mathbf{n}_{i_0}.$$

**Fluid boundary** Even if the considered model is more complex than the Shallow water system, we can consider that the type of the flow depends, for each layer, on the value of the Froude number  $Fr_\alpha = \frac{|\mathbf{u}_\alpha|}{\sqrt{gh}}$ , a flow is said torrential, for  $|\mathbf{u}_\alpha| > \sqrt{gh}$  and fluvial, for  $|\mathbf{u}_\alpha| < \sqrt{gh}$ .

Generally, for the fluid boundaries, the conditions prescribed by the user depend on the type of the flow defined by this criterium.

We have also to notice that with  $\mathbf{n}$  the outward unit normal to the boundary edge, an inflow boundary corresponds to  $\mathbf{u}_\alpha \cdot \mathbf{n} < 0$  and an outflow one to  $\mathbf{u}_\alpha \cdot \mathbf{n} > 0$ .

We will treat the following cases: for a fluvial flow boundary, we distinguish the cases where the flux or the water depth are given, while for a torrential flow we distinguish the inflow or outflow boundaries.

**Fluvial boundary. Flux given** We consider first a fluvial boundary, so we assume that

$$|\mathbf{u}_\alpha| < \sqrt{gh}. \quad (112)$$

If for each layer, the flux  $\mathbf{q}_{g,\alpha}$  is given, then we wish to impose

$$(F_h(\mathbf{U}_{\alpha,i}) + F_h(\mathbf{U}_{e,\alpha,i})) \cdot \mathbf{n}_i = \mathbf{q}_{g,\alpha} \cdot \mathbf{n}_i, \quad F_h(\mathbf{U}_{e,\alpha,i}) \cdot \mathbf{t}_i = q_{g,\alpha} \cdot \mathbf{t}_i, \quad (113)$$

with  $\mathbf{n}_i \cdot \mathbf{t}_i = 0$ . The value  $\mathbf{q}_{g,\alpha}$  depends on the value of the prescribed flux along the vertical axis.

If one imposes (113) directly leads to instabilities (especially because the numerical values are not necessarily in the regime of validity of this condition). We propose to discretize it in a weak form. We denote

$$a_1 = \mathbf{q}_{g,\alpha} \cdot \mathbf{n}_i - F_h(\mathbf{U}_{\alpha,i}) \cdot \mathbf{n}_i. \quad (114)$$

If  $a_1 \geq 0$ , we prescribe

$$F_h(\mathbf{U}_{e,\alpha,i}) = 0, \quad F_{hu}(\mathbf{U}_{e,\alpha,i}) = 0, \quad \text{and} \quad F_{hv}(\mathbf{U}_{e,\alpha,i}) = 0.$$

If  $a_1 < 0$ , we have to write a third equation to be able to compute the three components of  $\mathbf{U}_{e,\alpha}$  and by analogy with what is done for the Saint-Venant system – where the Riemann invariant related to the outgoing characteristic is preserved – we assume the quantity  $\mathbf{u}_\alpha \cdot \mathbf{n}$  is constant though the interface, i.e.

$$\mathbf{u}_{e,\alpha,i} \cdot \mathbf{n}_i - 2\sqrt{gh_{e,i}} = \mathbf{u}_{\alpha,i} \cdot \mathbf{n}_i - 2\sqrt{gh_i}. \quad (115)$$

As (112) is satisfied, the eigenvalue  $u_{e,\alpha,i} \cdot \mathbf{n}_i - 2\sqrt{gh_{e,i}}$  is positive.

We use the equations (113) and (115) to compute  $h_{e,i}$  and  $\mathbf{u}_{e,\alpha,i} \cdot \mathbf{n}_i$ . We denote  $a_2 = \mathbf{u}_{\alpha,i} \cdot \mathbf{n}_i - 2\sqrt{gh_i}$  and

$$m = \frac{\mathbf{u}_{e,\alpha,i} \cdot \mathbf{n}_i}{\sqrt{gh_{e,i}}}. \quad (116)$$

Then the equation (115) gives

$$\sqrt{gh_{e,i}} (m - 2) = a_2 \quad (117)$$

and using the definition of  $F_h$  (see (139)) with (113),(114) we have

$$\frac{h_{e,i}}{\pi} \int_{z \leq \frac{-\mathbf{u}_{e,\alpha,i} \cdot \mathbf{n}_i}{\sqrt{\frac{gh_{e,i}}{2}}}} \left( \mathbf{u}_{e,\alpha,i} \cdot \mathbf{n}_i + \sqrt{\frac{gh_{e,i}}{2}} z \right) \sqrt{1 - \frac{z^2}{4}} dz = a_1, \quad (118)$$

or using (115)

$$\psi(h_{e,i}) = a_1, \quad (119)$$

with

$$\psi(h_{e,i}) = \frac{h_{e,i}}{\pi} \int_{z \leq \frac{-(2\sqrt{gh_{e,i}} + a_2)}{\sqrt{\frac{gh_{e,i}}{2}}}} \left( 2\sqrt{gh_{e,i}} + a_2 + \sqrt{\frac{gh_{e,i}}{2}} z \right) \sqrt{1 - \frac{z^2}{4}} dz.$$

It is easy to see that  $h \mapsto \psi(h)$  is a growing function of  $h$  with  $\psi(0) = 0$  and  $\psi(+\infty) = +\infty$  and therefore, Eq. (118) admits a unique solution for any  $a_1 > 0$ . Using (117), Eq. (119) is equivalent to solve for  $m$

$$\Psi(m) = a_2, \quad (120)$$

with

$$\Psi(m) = K \frac{(m - 2)}{\phi(m)^{\frac{1}{3}}},$$

and  $K = (\sqrt{2}g a_1)^{\frac{1}{3}}$

$$\phi(m) = \frac{1}{\pi} \int_{z \leq -\sqrt{2}m} (\sqrt{2}m + z) \sqrt{1 - \frac{z^2}{4}} dz.$$

In practice, we use a Newton-Raphson algorithm to solve an equivalent form of Eq. (120), namely

$$m - 2 - \frac{a_2}{K} \phi(m)^{1/3} = 0.$$

Once the above equation has been solved, from (116),(117) we deduce

$$h_{e,i} = \frac{1}{g} \left( \frac{a_2}{m - 2} \right)^2, \quad \mathbf{u}_{e,\alpha,i} \cdot \mathbf{n}_i = \frac{a_2 m}{m - 2}.$$

**Remark 5.2** Notice that in the procedure proposed to calculate  $\mathbf{u}_{e,\alpha,i}$ ,  $h_{e,i}$ , even if  $h_{e,i}$  represents a total water depth, a different value of  $h_{e,i}$  is calculated for each layer  $\alpha$ .  $h_{e,i}$  is only used to ensure (113).

**Fluvial boundary. Water depth given** We verify that the flow is actually fluvial, i.e.

$$(\mathbf{u}_{\alpha,i} \cdot \mathbf{n}_i - \sqrt{gh_i})(\mathbf{u}_{\alpha,i} \cdot \mathbf{n}_i + \sqrt{gh_i}) \leq 0. \quad (121)$$

Since the water depth is given, we write

$$h_{e,i} = h_{g,i}. \quad (122)$$

We assume the continuity of the tangential component

$$(h\mathbf{u})_{e,\alpha,i} \cdot \mathbf{t}_i = (h\mathbf{u})_{\alpha,i} \cdot \mathbf{t}_i, \quad (123)$$

with  $\mathbf{t}_i \cdot \mathbf{n}_i = 0$ . To define completely  $\mathbf{u}_{e,\alpha,i}$ , we assume, as in the previous case, that the Riemann invariant is constant along the outgoing characteristic (115), so we obtain

$$\mathbf{u}_{e,\alpha,i} \cdot \mathbf{n}_i = \mathbf{u}_{\alpha,i} \cdot \mathbf{n}_i + 2\sqrt{g}(\sqrt{h_i} - \sqrt{h_{g,i}}). \quad (124)$$

Sometimes it appears that the numerical values do not satisfy the condition (121), then the flow is in fact torrential and

- if  $\mathbf{u}_{\alpha,i} \cdot \mathbf{n}_i > 0$ , the condition (122) cannot be satisfied (see Sec. 6.2.4),
- if  $\mathbf{u}_{\alpha,i} \cdot \mathbf{n}_i < 0$ , one condition is missing and we prescribe  $\mathbf{u}_{e,\alpha,i} \cdot \mathbf{n}_i = \mathbf{u}_{\alpha,i} \cdot \mathbf{n}_i$ .

**Torrential inflow boundary** For a torrential inflow boundary we assume that the water depth and the flux are given, then we prescribe

$$h_{e,i} = h_{g,i}, \quad (h\mathbf{u})_{e,\alpha,i} \cdot \mathbf{t}_i = (h\mathbf{u})_{g,\alpha,i} \cdot \mathbf{t}_i,$$

and

$$(F_h(\mathbf{U}_{\alpha,i}) + F_h(\mathbf{U}_{e,\alpha,i})) \cdot \mathbf{n}_i = \mathbf{q}_{g,\alpha} \cdot \mathbf{n}_i = (h\mathbf{u})_{g,\alpha,i} \cdot \mathbf{n}_i.$$

In this case we have to compute  $(h\mathbf{u})_{e,\alpha,i} \cdot \mathbf{n}_i$  or  $\mathbf{u}_{e,\alpha,i} \cdot \mathbf{n}_i$ . We consider an inflow boundary, so  $(h\mathbf{u})_{g,\alpha,i} \cdot \mathbf{n}_i < 0$  therefore using the notation (114) we have  $a_1 < 0$ . By analogy with the previous section we denote

$$m = \frac{\mathbf{u}_{e,\alpha,i} \cdot \mathbf{n}_i}{\sqrt{gh_{g,i}}},$$

then the equation for  $m$  is (see (116)-(118))

$$\phi(m) = \sqrt{\frac{2}{g}} \frac{a_1}{h_{g,i}^{3/2}}.$$

As in the paragraph entitled *Flux given*, the above equation has a unique solution  $m < 2$  for  $a_1 < 0$ .

**Torrential outflow boundary** In the case of a torrential outflow boundary, we do not prescribe any condition. We assume that the two Riemann invariants are constant along the outgoing characteristics leading to

$$\begin{aligned} \mathbf{u}_{e,\alpha,i} \cdot \mathbf{n}_i - 2\sqrt{gh_{e,i}} &= \mathbf{u}_{\alpha,i} \cdot \mathbf{n}_i - 2\sqrt{gh_i}, \\ \mathbf{u}_{e,\alpha,i} \cdot \mathbf{n}_i + 2\sqrt{gh_{e,i}} &= \mathbf{u}_{\alpha,i} \cdot \mathbf{n}_i + 2\sqrt{gh_i}, \end{aligned}$$

and we deduce  $h_{e,i} = h_i$ ,  $\mathbf{u}_{e,\alpha,i} \cdot \mathbf{n}_i = \mathbf{u}_{\alpha,i} \cdot \mathbf{n}_i$ . We assume that we also have  $(h\mathbf{u})_{e,\alpha,i} \cdot \mathbf{t}_i = (h\mathbf{u})_{\alpha,i} \cdot \mathbf{t}_i$ .

### 5.7.2 Layer-averaged Navier-Stokes system

Because of the fractional step we use, the boundary conditions for the layer-averaged Euler system are, to some extent, independent from the one used for the rheology terms.

For the resolution of Eq. (108), boundary conditions associated with the operator

$$\nabla_{x,y} \cdot (h_\alpha \mathbf{T}_\alpha^0),$$

have to be specified and usually we prescribe homogeneous Neumann boundary conditions (corresponding to an imposed stress). Of course, in particular cases, Dirichlet or Robin type boundary conditions can also be considered.

## 5.8 Toward second order schemes

In order to improve the accuracy of the results the first-order scheme defined in paragraphs 5.3-5.5 can be extended to a formally second-order one using a MUSCL like extension (see [46]).

### 5.8.1 Second order reconstruction for the layer-averaged Euler system

In the definition of the flux (97), we replace the piecewise constant values  $\mathbf{U}_{i,j}, \mathbf{U}_{j,i}$  by more accurate reconstructions deduced from piecewise linear approximations, namely the values  $\tilde{\mathbf{U}}_{i,j}, \tilde{\mathbf{U}}_{j,i}$  reconstructed on both sides of the interface. The reconstruction procedure is similar to the one used and described in [7, paragraph. 5.1].

The second order reconstruction is only applied for the horizontal fluxes. For the exchange terms along the vertical axis involving the quantities  $G_{\alpha \pm 1/2}$ , we keep the first order approximation. Despite this, we recover over the simulations (see paragraphs 6.1, 6.2) a second order type convergence curve.

### 5.8.2 Modified Heun scheme

The explicit time scheme (76)-(77) used in the previous paragraphs corresponds to a first order explicit Euler scheme. The second-order accuracy in time is usually recovered by the Heun method [17] that is a slight modification of the second order Runge-Kutta method. More precisely, for a dynamical system written under the form

$$\frac{\partial y}{\partial t} = f(y), \quad (125)$$

the Heun scheme consists in defining  $y^{n+1}$  by

$$y^{n+1} = y(t^n + \Delta t^n) = \frac{y^n + \tilde{y}^{n+2}}{2}, \quad (126)$$

with

$$\tilde{y}^{n+1} = y^n + \Delta t^n f(y^n, t^n), \quad \tilde{y}^{n+2} = \tilde{y}^{n+1} + \Delta t^n f(\tilde{y}^{n+1}, t^{n+1}). \quad (127)$$

But the scheme defined by (126) does not preserve the invariant domains. Indeed, the time step being given by a CFL condition,  $\Delta t^n$  in the relation (127) should be replaced by  $\tilde{\Delta t}^{n+1}$  i.e. the time step satisfying the CFL condition and calculated using  $\tilde{y}^{n+1}$ . Thus in situations where the time step strongly varies from one iteration to another, the Heun scheme does not preserve the positivity of the scheme.

To overcome this difficulty, we propose an improvement of the Heun scheme

**Proposition 5.3** *The scheme defined by  $y^{n+1} = (1 - \gamma)y^n + \gamma\tilde{y}^{n+2}$  with*

$$\tilde{y}^{n+1} = y^n + \Delta t_1^n f(y^n), \quad \tilde{y}^{n+2} = \tilde{y}^{n+1} + \Delta t_2^n f(\tilde{y}^{n+1}),$$

and

$$\Delta t^n = \frac{2\Delta t_1^n \Delta t_2^n}{\Delta t_1^n + \Delta t_2^n}, \quad \gamma = \frac{(\Delta t^n)^2}{2\Delta t_1^n \Delta t_2^n},$$

is second order and compatible with a CFL constraint. Since  $\gamma \geq 0$ ,  $y^{n+1}$  is a convex combination of  $y^n$  and  $\tilde{y}^{n+2}$  so the scheme preserves the positivity. For the previous relations  $\Delta t_1^n$  and  $\Delta t_2^n$  respectively satisfy the CFL conditions associated with  $y^n$  and  $\tilde{y}^{n+1}$ .

When  $\Delta t_1^n = \Delta t_2^n = \Delta t^n$ , the scheme reduces to the classical Heun scheme with  $\alpha = \gamma = 1/2$ .

**Proof of proposition 5.3** Using (125), a Taylor expansion of  $y(t^n + \Delta t^n)$  gives

$$y(t^n + \Delta t^n) = y^n + \Delta t^n f(y^n) + \frac{(\Delta t^n)^2}{2} f(y^n) f'(y^n) + \mathcal{O}((\Delta t^n)^3).$$

Using the definitions given in the proposition, we have

$$\tilde{y}^{n+2} = y^n + (\Delta t_1^n + \Delta t_2^n) f(y^n) + \Delta t_1^n \Delta t_2^n f(y^n) f'(y^n) + \mathcal{O}(\Delta t_2^n (\Delta t_1^n)^2),$$

and a simple calculus gives  $\alpha y^n + \beta \tilde{y}^{n+1} + \gamma \tilde{y}^{n+2} - y(t^n + \Delta t^n) = \mathcal{O}((\Delta t^n)^3)$ , that completes the proof. ■

## 6 Numerical validations

In this section, we use the numerical scheme to simulate several test cases: analytical solutions or in situ measurements, stationary or non-stationary solutions, for the Euler and Navier-Stokes systems. The obtained results emphasize the accuracy of the numerical procedure in a wide range of typical applications.

The numerical simulations presented in this section have been obtained with the code Freshkiss3d [1] where the numerical scheme presented in this paper is implemented.

### 6.1 Thacker' analytical solution

The proposed model is an approximation of the 3d Navier-Stokes and hence it is more general than the Saint-Venant system [13, 33]. But when considering flows with

$$\frac{\partial u}{\partial z} = \frac{\partial v}{\partial z} = 0,$$

the solutions of the model depicted in this paper should corresponds to the solutions of the Saint-Venant system. The Thacker' solution [44], corresponds to a periodic oscillation in a parabolic bowl. This analytical solution for the 2d Saint-Venant system is also an analytical solution for the 3d incompressible and hydrostatic Euler system.

From [44], considering the parameters  $(h_0, \eta_0, a) \in \mathbb{R}_+^* \times \mathbb{R}^2$

$$z_b(x, y) = -h_0 \left( 1 - \frac{x^2 + y^2}{a^2} \right), \quad (128)$$

then

$$\begin{cases} h = \frac{\eta_0 h_0}{a^2} \left( 2 \left( x - \frac{L}{2} \right) \cos \omega t + 2 \left( y - \frac{L}{2} \right) \sin \omega t - \eta_0 \right) - z_b(x, y) \\ u = -\eta_0 \omega \sin \omega t \\ v = \eta_0 \omega \cos \omega t \\ w = \mathbf{u} \cdot \nabla_{x,y} z_b(x, y) \end{cases} \quad (129)$$

with  $\omega = \sqrt{2g h_0}/a$  is solution of the free surface incompressible Euler system (17)-(18) completed with (4),(7).

The geometrical domain is defined by  $(x, y) \in [-L/2, L/2]^2$ , with the parameter values  $L = 4$ ,  $h_0 = 1$  m,  $a = 1.5$  m,  $\eta_0 = 0.5$  m, and the topography is given by (128). The initial conditions correspond to (129) at time  $t = 0$  s.

In order to evaluate the convergence rate of the simulated solution  $h_{sim}$  towards the analytical one  $h_{anal}$ , we plot the error rate versus the space discretization for five unstructured meshes with 1273, 1104, 30441, 59473 and 98137 nodes. We have plotted (see Fig. 3) the  $\log(L^2 - error)$  over the water depth at time  $T = 2\pi/\omega$  seconds versus  $\log(h_{a0}/h_a)$  (we denote by  $h_a$  the average edge length and  $h_{a0}$  the average edge length of the coarser mesh) for the first and second-order schemes and they are compared to the theoretical order.

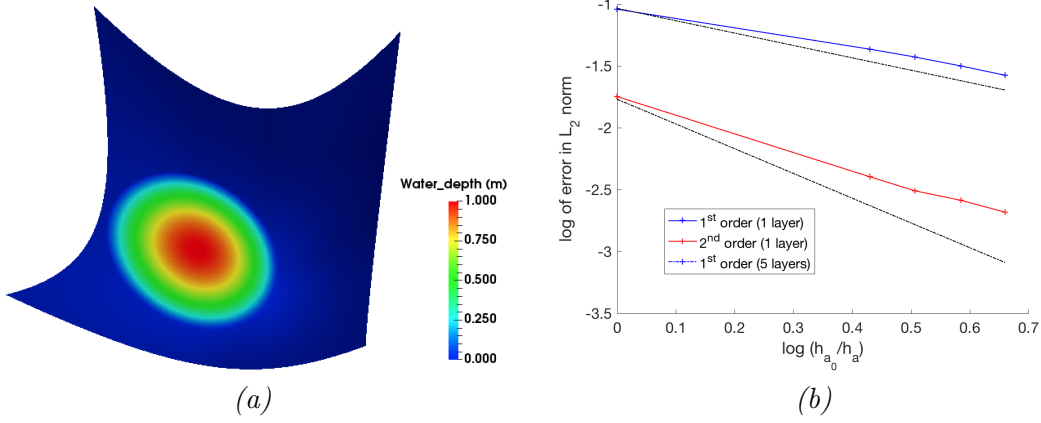


Figure 3: Parabolic bowl: (a) the geometrical domain and the water depth, (b) error between the analytical water depth and the simulated one with the five unstructured meshes. The curves for the first order scheme (space and time) and its second order extension (space and time) are compared to the first and second order theoretical curves (dashed lines).

Notice that over Fig. 3-(b), we have plotted the simulated solution obtained with 1 layer – that corresponds to the classical Saint-Venant system – and the simulated solution with 5 layers. The obtained convergence curves coincide proving the stability of the numerical scheme for the layers averaged system. The analytical solution is non stationary and hence, the errors due to the time scheme are combined with the one induced by the space discretization and it is the reason why we do not recover the theoretical order of convergence.

## 6.2 Stationary analytical solutions

We now compare our numerical model with stationary analytical solutions for the free surface Euler system proposed by some of the authors in [22].

We consider as geometrical domain a channel  $(x, y) \in [0, x_{max}] \times [0, 2]$ . The analytical solution given in [22, Prop. 3.1] and defined by

$$z_b = \bar{z}_b - h_0 - \frac{\alpha^2 \beta^2}{2g \sin^2(\beta h_0)}, \quad (130)$$

$$u_{\alpha, \beta} = \frac{\alpha \beta}{\sin(\beta h_0)} \cos(\beta(z - z_b)), \quad (131)$$

$$v_{\alpha, \beta} = 0,$$

$$w_{\alpha, \beta} = \alpha \beta \left( \frac{\partial z_b}{\partial x} \frac{\cos(\beta(z - z_b))}{\sin(\beta h_0)} + \frac{\partial h_0}{\partial x} \frac{\sin(\beta(z - z_b)) \cos(\beta h_0)}{\sin^2(\beta h)} \right),$$

with  $\alpha = 1 \text{ m}^2 \cdot \text{s}^{-1}$ ,  $\beta = 1 \text{ m}^{-1}$ ,  $\bar{z}_b = cst$ ,  $x_{max} = 20 \text{ m}$  and

$$h_0(x, y) = \frac{1}{2} + \frac{3}{2} \frac{1}{1 + (x - \frac{1}{2}x_{max})^2} - \frac{1}{2} \frac{1}{2 + (x - \frac{2}{3}x_{max})^2}, \quad (132)$$

is a stationary regular analytical solution of the incompressible and hydrostatic Euler system with free surface (17)-(18),(4),(7) with  $p^a = 0$ .

In order to obtain the simulated solution, we consider the topography defined by (130),(132) and we impose the following boundary conditions

- solid wall for the two boundaries  $y = 0$  m and  $y = 2$  m,
- given water depth  $h_0(x_{max}, y)$  at  $x = x_{max} = 20$  m,
- given flux defined by (131) at  $x = 0$  m.

We have performed the simulations for several unstructured meshes and an adapted number of layers so that each 3d element of the mesh can be approximatively considered as a regular polyhedron (see remark 6.1), the considered meshes have 290 nodes and 2 layers, 597 nodes and 4 layers, 1010 nodes and 8 layers, 2112 nodes and 17 layers.

Over Fig. 4(a), we have depicted the features of the analytical solution we use for the convergence test, it clearly appears over Fig. 4(a) that the velocity profile of chosen analytical solution varies along the  $z$  axis. Figure 4-(b) gives the convergence curve towards the analytical solution i.e. the  $\log(L^2 - error)$  of the water depth – at time  $T = 300$  seconds when the stationary regime is reached – versus  $\log(h_{a0}/h_a)$  for the first and second-order schemes and they are compared to the theoretical order.

**Remark 6.1** *In the case of a flat topography and a flat free surface, the 3d elements composing the mesh are triangular prisms. But in more complex situations, the lower and upper triangles are not necessarily parallel and the 3d elements of the mesh are a portion of a vertical cylinder with a triangular basis.*

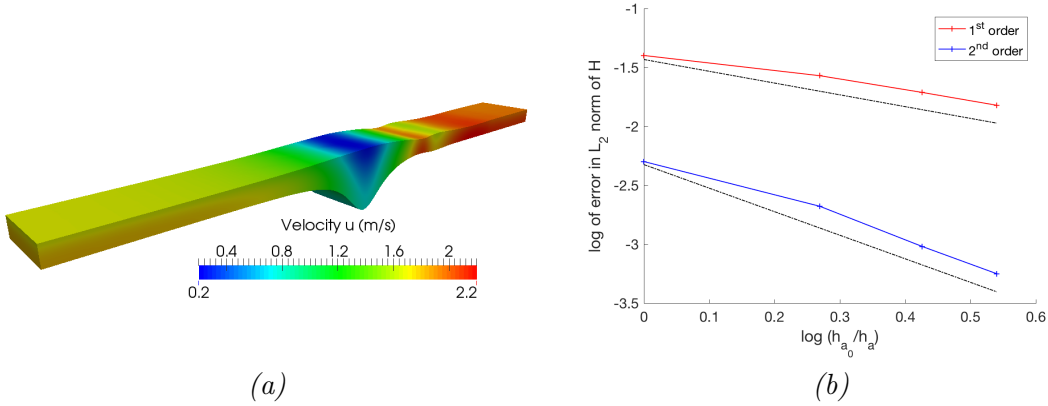


Figure 4: (a) Surface level of the analytical solution (130)-(132) and horizontal velocity  $u_{\alpha,\beta}$ , (b) error between the analytical solution and the simulated one with the six meshes, first order (space and time) and second order extension (space and time) schemes. The first and second order theoretical curves correspond to the dashed lines.

**Remark 6.2** *Following the results given in [22, paragraph 3.4], it is possible to obtain stationary analytical solutions with discontinuities for the Euler system. In this case, the unknowns are not given by algebraic expressions but are obtained through the resolution of an ODE involving only the water depth  $h$ .*

*The numerical scheme has been used in the context of such a discontinuous analytical solution. As planned, for the first and second order schemes, we recover a first order convergence of the simulated solution towards the analytical one because of the discontinuity of the reference solution.*



### 6.3 Non-stationary analytical solutions for Navier-Stokes

Considering the Navier-Stokes system (1)-(3) completed with the boundary conditions (7)-(9), the following proposition holds.

**Proposition 6.1** *For some  $\alpha \in \mathbb{R}_+$ ,  $(\beta, \gamma) \in \mathbb{R}^2$ ,  $t_0 \in \mathbb{R}$ ,  $t_1 \in \mathbb{R}_+^*$  let us consider the functions  $u, v, w, h, p$  defined for  $(x, y) \in (-L/2, L/2) \times (-l/2, l/2)$ ,  $t \geq t_0$  by*

$$h(t, x, y) = \frac{\alpha}{t - t_0 + t_1}, \quad (133)$$

$$u(t, x, y, z) = \beta \left( z - \frac{h}{2} \right) + \frac{\gamma h}{\alpha} (x + y), \quad (134)$$

$$v(t, x, y, z) = \frac{(1 - \gamma)\beta}{\gamma} \left( z - \frac{h}{2} \right) + \frac{(1 - \gamma)h}{\alpha} (x + y), \quad (135)$$

$$w(t, x, y, z) = -\frac{z}{t - t_0 + t_1}, \quad (136)$$

$$p(t, x, y, z) = p^a(t, x, y) + g(h - z), \quad (137)$$

with at the free surface  $p^a(t, x, y) = p^{a,0}(t)$ ,  $W = \beta$  in (8) and at the bottom  $\kappa(h, \mathbf{U}) = \mu\beta \text{sign}(\mathbf{U}_b)$  in (6). For the choices (134), (135), the friction law defined by  $\mu\beta \text{sign}(\mathbf{U}_b)$  corresponds to a dissipation if  $x + y \leq \frac{\alpha\beta}{2\gamma}$  that is satisfied if  $\gamma(L + l) \leq \alpha\beta$ .

Then Eqs. (133)-(137) are solutions of the 3d incompressible, hydrostatic Navier-Stokes system with free surface (11)-(14) completed with the kinematic boundary conditions (7), (4).

The appropriate boundary conditions in  $x = \pm L/2$  are determined by the expressions of  $u, v, w, h, p$  given above.

Choosing the viscosity  $\mu = 0$ , the variables  $u, v, w, h, p$  become analytical solutions of the 3d Euler system (17)-(18) completed with the boundary conditions (4), (7).

**Proof of prop. 6.1** *The proof of prop. 6.1 relies on very simple computations and is not detailed here.*

■

We have performed the simulations for several unstructured meshes of the geometrical domain  $(x, y) \in [0, 5] \times [0, 1]$  and an adapted number of layers so that each 3d element of the mesh can be approximatively considered as a regular polyhedron, the considered meshes have 483 nodes and 3 layers, 700 nodes and 6 layers, 1306 nodes and 10 layers, 2781 nodes and 20 layers.

For  $L = 2$  m,  $\alpha = 1$  m.s,  $t_0 = 0$  s,  $t_1 = 0.5$  s,  $\beta = 2.5$  s<sup>-1</sup>,  $\gamma = 1$ , over Fig. 5-(a), we have depicted the features of the analytical solution – at time  $T = 1$  second – we use for the convergence test. Figure 5-(b) gives the convergence curve towards the analytical solution i.e. the  $\log(L^2 - \text{error})$  of the water depth – at time  $T = 1$  second – versus  $\log(h_{a0}/h_a)$  for the first and second-order schemes and they are compared to the theoretical order (we denote by  $h_a$  the average edge length and  $h_{a0}$  the average edge length of the coarser mesh). Notice that in this test case, the errors due to the space and time discretization are combined, this explains why the theoretical orders of convergence are not exactly obtained. Moreover, the boundary conditions (inflow prescribed) play an important role and since their numerical treatment is only at the first order in space, this also explains the difference between the theoretical and observed orders of convergence.

### 6.4 Simulation of a tsunami

In this section, we test our discrete model in the case of a real tsunami propagation. We perform simulation of a large scale phenomenon where measurements (free surface variations recorded by buoys) are available. Simulation of tsunami waves generated by earthquakes is very important in Earth science for hazard assessment and for recovering earthquakes characteristics. Indeed, tsunami waves can be analyzed to recover the earthquake source that generated the tsunami and are now classically used in

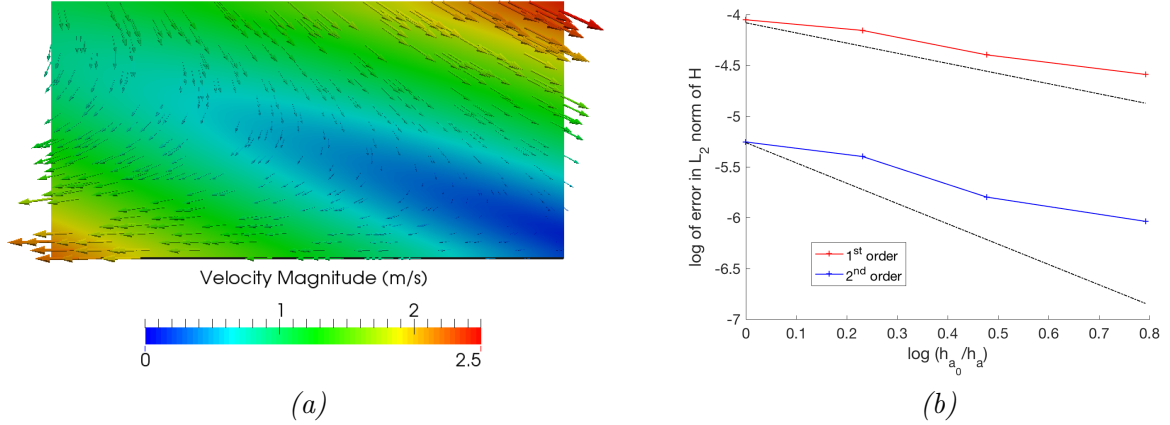


Figure 5: Analytical solution given in prop. 6.1 : (a) slice of the fluid domain for  $y = 0$  and velocity field at time  $T = 1$  s, (b) convergence curve towards the reference solution, first order (space and time) and second order extension (space and time) schemes. The first and second order theoretical curves correspond to the dashed lines.

joint inversion methods. It has been shown that tsunami waves provide strong constraints on the spatial distribution of the source, especially in shallow regions [35]. In some cases, far-field tsunami gauges may help constrain the earthquake source process even though they are affected by the compressibility of the water column and of the Earth [48, 35].

The 2014 Iquique earthquake struck off the coast of Chile on 1 April, with a moment magnitude of 8.1, at 20:46 local time (23:46 UTC). The epicenter of the earthquake was approximately 95 kilometres (59 mi) northwest of Iquique, see Fig. 6.

We have carried out simulations of the tsunami induced by the earthquake using

- a topography obtained from the National Oceanic and Atmospheric Administration (NOAA, [2]) using the ETOPO1 data (1-arc minute global relief model),
- an unstructured mesh whose dimensions – a square of  $2224.2 \text{ km}^2$  – corresponds to the domain covered by Fig. 6,
- a source (displacement of the seafloor due to the earthquake, see Fig. 6). This source is obtained by computing the 3D final displacements of the seafloor induced by the earthquake coseismic slip. This coseismic slip has been itself retrieved by inversion of numerous geodetic and seismic data and corresponds to the model determined by [45].

A series of simulation has been performed using several meshes and we present “converged” results in the sense that a finer mesh would give the same results. The results given over Fig. 7 have been obtained with a mesh containing 545821 nodes and 5 layers. We confront the numerical resolution – the first order scheme (space and time) and the second order scheme (space and time) – with the DART measurements (obtained from the NOAA website <http://www.ndbc.noaa.gov/dart.shtml>).

For the initial conditions at time  $t_0$ , we start from an ocean at rest and the seism occurs at time  $t_0$  ( $t_0$  is here 2014/04/01, 23h47mn25s, ). In the simulation, we do not consider the Coriolis force, the tides and the ocean currents.

The simulation has been carried out with a Mac book air 1.7 GHz Intel core i7 and takes 35 minutes.

Fig. 7 shows that the second order scheme significantly improves the results both for the amplitude and phase of the water waves. The second order scheme makes it possible to very accurately reproduce the shape of the first wave at the closest DART buoy 32401 located at 287 km from the epicenter. The two following peaks in the waveform are quite well reproduced up to about  $1.755 \times 10^5$  s. This is also

the case at the DART buoy 32402 located 853 km from the epicenter. The arrival time of the first wave is very well reproduced at the three DART buoys, slightly better with the second order scheme. At the most distant buoy 32412 (1650 km from the source), the first order scheme is not able to reproduce the recorded wave. The second order scheme quite well reproduces the first wave but not the rest of the waveform, possibly due to Earth curvature effects that are not taken into account here. Note that only the low frequency signal is reproduced by the model. Higher frequency variations of the waveform may be related to the spatio-temporal heterogeneity of the real source, small wavelength fluctuation of the topography not accounted for here, and possibly to non-hydrostatic effects [25, 3].

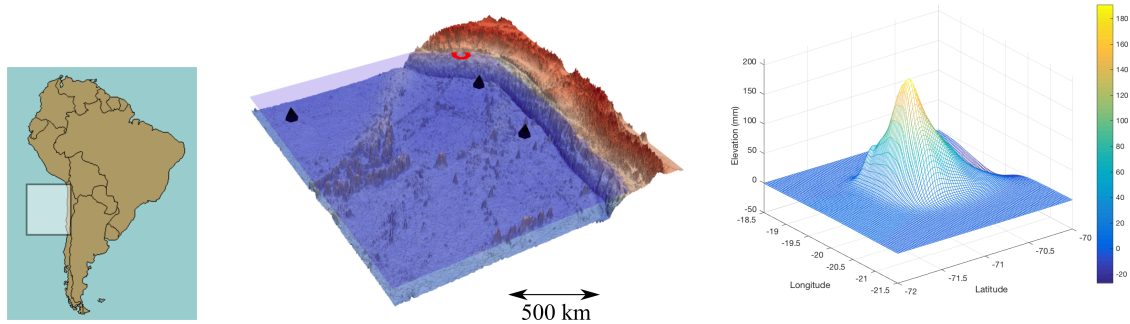


Figure 6: (Left) The Pacific ocean with the zone of interest; (Center) The Chilean coast with the epicenter shown (red cone) and the location of the three DART buoys (black boxes); (Right) Vertical displacement (in centimeter) of the topography due to the earthquake. Horizontal displacements (not shown here) are also taken into account in the simulation.

## 7 Conclusion

In this paper, we have presented a layer-averaged version of the incompressible, hydrostatic Euler and Navier-Stokes systems with free surface. Based on a kinetic interpretation of the system, we have derived a stable, robust and efficient numerical scheme in a finite volume framework on unstructured meshes. The numerical scheme is endowed with strong stability properties (domain invariant, well-balancing, wet/dry interfaces treatment, ...).

The numerical scheme is successfully validated with analytical solutions and complex test cases, like a tsunami propagation, proving its accuracy and efficiency.

**Acknowledgements** The authors thank Quentin Bl  tery for his involvement in the analysis of the 2014 Iquique earthquake (paragraph 6.4). We also thank Rapha  l Grandin for its contribution in the computations of the bottom displacements necessary for the tsunami simulations. This work has been partially funded by the ERC Contract No. ERC-CG-2013-PE10-617472 SLIDEQUAKES.

## References

- [1] *Freshkiss3d home page*, <http://freshkiss3d.gforge.inria.fr>, 2017.
- [2] *NOAA home page*, <https://www.ngdc.noaa.gov/mgg/global/global.html>, 2017.
- [3] N. A  siouene, M.-O. Bristeau, E. Godlewski, and J. Sainte-Marie, *A robust and stable numerical scheme for a depth-averaged Euler system* - <https://hal.inria.fr/hal-01162109>, June 2015.
- [4] E. Audusse, *A multilayer Saint-Venant model : Derivation and numerical validation*, Discrete Contin. Dyn. Syst. Ser. B **5** (2005), no. 2, 189–214.

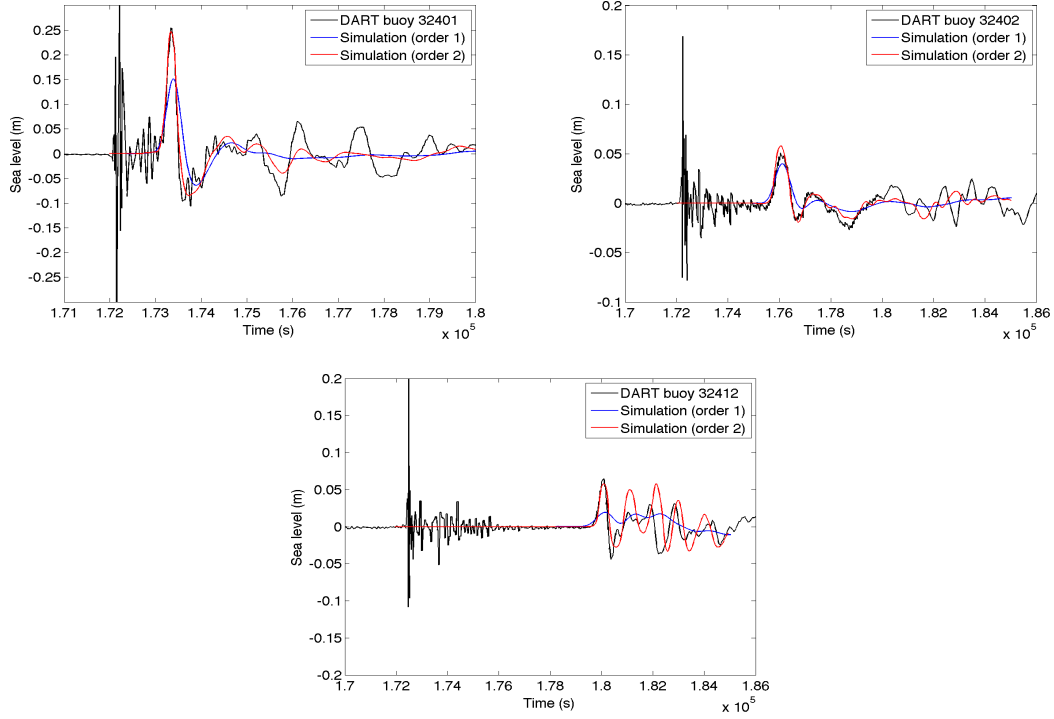


Figure 7: Comparison between the sea level variations recorded by the DART buoys (after detiding using a low pass Butterworth filter of order 4 with a cutoff frequency of 4 hours) and the simulation ( $1^{st}$  and  $2^{nd}$  order schemes). The simulation results have been filtered using the same Butterworth filter.

- [5] 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.
- [6] 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
- [7] E. Audusse and M.-O. Bristeau, *A well-balanced positivity preserving second-order scheme for Shallow Water flows on unstructured meshes*, J. Comput. Phys. **206** (2005), no. 1, 311–333.
- [8] ———, *Finite-volume solvers for a multilayer Saint-Venant system*, Int. J. Appl. Math. Comput. Sci. **17** (2007), no. 3, 311–319.
- [9] E. Audusse, M.-O. Bristeau, and A. Decoene, *Numerical simulations of 3d free surface flows by a multilayer Saint-Venant model*, Internat. J. Numer. Methods Fluids **56** (2008), no. 3, 331–350.
- [10] E. Audusse, M.-O. Bristeau, M. Pelanti, and J. Sainte-Marie, *Approximation of the hydrostatic Navier-Stokes system for density stratified flows by a multilayer model. Kinetic interpretation and numerical validation.*, J. Comp. Phys. **230** (2011), 3453–3478.
- [11] E. Audusse, M.-O. Bristeau, B. Perthame, and J. Sainte-Marie, *A multilayer Saint-Venant system with mass exchanges for Shallow Water flows. Derivation and numerical validation*, ESAIM: M2AN **45** (2011), 169–200.
- [12] E. Audusse, M.-O. Bristeau, and J. Sainte-Marie, *Kinetic entropy for layer-averaged hydrostatic Navier-Stokes equations*, submitted, 2017.

- [13] 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.
- [14] F. Berthelin and F. Bouchut, *Relaxation to isentropic gas dynamics for a BGK system with single kinetic entropy*, Methods Appl. Anal. **9** (2002), no. 2, 313–327. MR 1957492
- [15] F. Bouchut, *Construction of BGK models with a family of kinetic entropies for a given system of conservation laws*, J. Stat. Phys. **95** (1999), 113–170.
- [16] ———, *Entropy satisfying flux vector splittings and kinetic BGK models*, Numer. Math. **94** (2003), 623–672.
- [17] ———, *An introduction to finite volume methods for hyperbolic conservation laws*, ESAIM Proc. **15** (2004), 107–127.
- [18] ———, *Nonlinear stability of finite volume methods for hyperbolic conservation laws and well-balanced schemes for sources*, Birkhäuser, 2004.
- [19] F. Bouchut and T. Morales de Luna, *An entropy satisfying scheme for two-layer shallow water equations with uncoupled treatment*, M2AN Math. Model. Numer. Anal. **42** (2008), 683–698.
- [20] F. Bouchut and M. Westdickenberg, *Gravity driven shallow water models for arbitrary topography*, Comm. in Math. Sci. **2** (2004), 359–389.
- [21] F. Bouchut and V. Zeitlin, *A robust well-balanced scheme for multi-layer shallow water equations*, Discrete Contin. Dyn. Syst. Ser. B **13** (2010), 739–758.
- [22] A.-C. Boulanger and J. Sainte-Marie, *Analytical solutions for the free surface hydrostatic Euler equations*, Commun. Math. Sci. **11** (2013), no. 4, 993–1010.
- [23] Y. Brenier, *Homogeneous hydrostatic flows with convex velocity profiles*, Nonlinearity **12** (1999), no. 3, 495–512.
- [24] M.-O. Bristeau, B. Di-Martino, C. Guichard, and J. Sainte-Marie, *Layer-averaged Euler and Navier-Stokes equations*, Commun. Math. Sci. **15** (2017), no. 5, 1221–1246.
- [25] 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.
- [26] M.O. Bristeau and B. Coussin, *Boundary Conditions for the Shallow Water Equations solved by Kinetic Schemes*, Research Report RR-4282, INRIA, 2001.
- [27] M.-J. Castro, J. Macías, and C. Parés, *A Q-scheme for a class of systems of coupled conservation laws with source term. Application to a two-layer 1-D shallow water system*, M2AN Math. Model. Numer. Anal. **35** (2001), no. 1, 107–127.
- [28] M.J. Castro, J.A. García-Rodríguez, J.M. González-Vida, J. Macías, C. Parés, and M.E. Vázquez-Cendón, *Numerical simulation of two-layer shallow water flows through channels with irregular geometry*, J. Comput. Phys. **195** (2004), no. 1, 202–235.
- [29] V. Casulli, *A semi-implicit numerical method for the free-surface Navier-Stokes equations*, Internat. J. Numer. Methods Fluids **74** (2014), no. 8, 605–622. MR 3168787
- [30] E.D. Fernández-Nieto, G. Garres-Dias, A. Mangeney, and G. Narbona-Reina, *A multilayer shallow model for dry granular flows with the  $\mu(I)$ -rheology: Application to granular collapse on erodible beds*, Journal of Fluid Mechanics (2016).

- [31] E.D. Fernández-Nieto, E.H. Koné, and T. Chacón Rebollo, *A multilayer method for the hydrostatic Navier-Stokes equations: a particular weak solution*, Journal of Scientific Computing **60** (2014), no. 2, 408–437.
- [32] S. Ferrari and F. Saleri, *A new two-dimensional Shallow Water model including pressure effects and slow varying bottom topography*, M2AN Math. Model. Numer. Anal. **38** (2004), no. 2, 211–234.
- [33] 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.
- [34] E. Grenier, *On the derivation of homogeneous hydrostatic equations*, ESAIM: M2AN **33** (1999), no. 5, 965–970.
- [35] A. Gusman, S. Murotani, K. Satake, M. Heidarzadeh, E. Gunawan, S. Watada, and B. Schurr, *Fault slip distribution of the 2014 iquique, chile, earthquake estimated from ocean-wide tsunami waveforms and gps data*, Geophysical Research Letters **42** (2015), no. 4, 1053–1060, 2014GL062604.
- [36] J.-M. Hervouet, *Hydrodynamics of free surface flows: Modelling with the finite element method*, Wiley, 2007.
- [37] C.D. Levermore and M. Sammartino, *A shallow water model with eddy viscosity for basins with varying bottom topography*, Nonlinearity **14** (2001), no. 6, 1493–1515.
- [38] F. Marche, *Derivation of a new two-dimensional viscous shallow water model with varying topography, bottom friction and capillary effects*, European Journal of Mechanic /B **26** (2007), 49–63.
- [39] N. Masmoudi and T. Wong, *On the  $H^s$  theory of hydrostatic Euler equations*, Archive for Rational Mechanics and Analysis **204** (2012), no. 1, 231–271.
- [40] L. V. Ovsyannikov, *Two-layer shallow water models*, Prikl. Mekh. Tekh. Fiz. **2** (1979), 3–14.
- [41] B. Perthame, *Kinetic formulation of conservation laws*, Oxford University Press, 2002.
- [42] B. Perthame and C. Simeoni, *A kinetic scheme for the Saint-Venant system with a source term*, Calcolo **38** (2001), no. 4, 201–231.
- [43] J. Sainte-Marie, *Vertically averaged models for the free surface Euler system. Derivation and kinetic interpretation*, Math. Models Methods Appl. Sci. (M3AS) **21** (2011), no. 3, 459–490.
- [44] W. C. Thacker, *Some exact solutions to the non-linear shallow-water wave equations*, J. Fluid Mech. **107** (1981), 499–508.
- [45] M. Vallée, R. Grandin, S. Ruiz, B. Delouis, C. Vigny, E. Rivera, E. Aissaoui, S. Alleyer, Q. Bletery, C. Satriano, N. Poiata, P. Bernard, J.-P. Vilotte, and B. Schurr, *Complex rupture of an apparently simple asperity during the 2014/04/01 pisagua earthquake (northern chile,  $m_w=8.1$ )*, EGU General Assembly Conference Abstracts, vol. 18, EGU2016-8660, 2016.
- [46] B. Van Leer, *Towards the ultimate conservative difference scheme. V. A second-order sequel to Godunov’s method*, J. Comput. Phys. **135** (1997), no. 2, 227–248, With an introduction by Ch. Hirsch, Commemoration of the 30th anniversary {of J. Comput. Phys.}. MR 1486274
- [47] C. B. Vreugdenhil, *Two-layer shallow-water flow in two dimensions, a numerical study*, J. Comput. Phys. **33** (1979), 169–184.
- [48] S. Watada, S. Kusumoto, and K. Satake, *Traveltime delay and initial phase reversal of distant tsunamis coupled with the self-gravitating elastic earth*, Journal of Geophysical Research: Solid Earth **119** (2014), no. 5, 4287–4310, 2013JB010841.

## Appendix A

**Proof of prop. 3.4** Using the results of prop. 3.1, it remains to obtain the expression for the layer-averaged viscous terms.

The layer-averaging of the viscous terms appearing in (12) gives

$$\begin{aligned}
\int_{z_{\alpha+1/2}}^{z_{\alpha+1/2}} \mu \Delta u \, dz &= \int_{z_{\alpha+1/2}}^{z_{\alpha+1/2}} \left( \frac{\partial \hat{T}_{xx}}{\partial x} + \frac{\partial \hat{T}_{xy}}{\partial y} + \frac{\partial \hat{T}_{xz}}{\partial z} \right) dz \\
&= \frac{\partial}{\partial x} \left( h_{\alpha} \hat{T}_{xx,\alpha} \right) + \frac{\partial}{\partial y} \left( h_{\alpha} \hat{T}_{xy,\alpha} \right) \\
&\quad + \hat{T}_{xz,\alpha+1/2} - \hat{T}_{xz,\alpha-1/2} \\
&\quad - \frac{\partial z_{\alpha+1/2}}{\partial x} \hat{T}_{xx,\alpha+1/2} - \frac{\partial z_{\alpha+1/2}}{\partial y} \hat{T}_{xy,\alpha+1/2} \\
&\quad + \frac{\partial z_{\alpha-1/2}}{\partial x} \hat{T}_{xx,\alpha-1/2} + \frac{\partial z_{\alpha-1/2}}{\partial y} \hat{T}_{xy,\alpha-1/2},
\end{aligned} \tag{138}$$

with

$$h_{\alpha} \hat{T}_{xx,\alpha} = \int_{z_{\alpha-1/2}}^{z_{\alpha-1/2}} \hat{T}_{xx} dz.$$

The boundary conditions at the bottom (13) and at the free surface (14) imply that in (138)

$$\hat{T}_{xz,N+1/2} - \frac{\partial \eta}{\partial x} \hat{T}_{xx,N+1/2} - \frac{\partial \eta}{\partial y} \hat{T}_{xy,N+1/2} = 0,$$

and

$$\hat{T}_{xz,1/2} - \frac{\partial z_b}{\partial x} \hat{T}_{xx,1/2} - \frac{\partial z_b}{\partial y} \hat{T}_{xy,1/2} = \kappa u_1.$$

The above expression for  $\hat{T}_{xx,\alpha}$  and relation (138) have been obtained using the following formal computation

$$\begin{aligned}
\int_{z_{\alpha+1/2}}^{z_{\alpha+1/2}} \frac{\partial \hat{T}_{xx}}{\partial x} dz &= \frac{\partial}{\partial x} \int_{z_{\alpha+1/2}}^{z_{\alpha+1/2}} \hat{T}_{xx} dz - \frac{\partial z_{\alpha+1/2}}{\partial x} \hat{T}_{xx,\alpha+1/2} + \frac{\partial z_{\alpha-1/2}}{\partial x} \hat{T}_{xx,\alpha-1/2} \\
&= \frac{\partial}{\partial x} \left( h_{\alpha} \hat{T}_{xx,\alpha} \right) - \frac{\partial z_{\alpha+1/2}}{\partial x} \hat{T}_{xx,\alpha+1/2} + \frac{\partial z_{\alpha-1/2}}{\partial x} \hat{T}_{xx,\alpha-1/2}.
\end{aligned}$$

The definitions (37),(38) are motivated by the following computation

$$\begin{aligned}
\frac{h_{\alpha+1} + h_{\alpha}}{2} \hat{T}_{xx,\alpha+1/2} &= \mu_{\alpha+1/2} \int_{z_{\alpha}}^{z_{\alpha+1}} \frac{\partial u}{\partial x} dz \\
&= \mu_{\alpha+1/2} \frac{\partial}{\partial x} \int_{z_{\alpha}}^{z_{\alpha+1}} u dz - \mu_{\alpha+1/2} \frac{\partial z_{\alpha+1}}{\partial x} u_{\alpha+1} + \mu_{\alpha+1/2} \frac{\partial z_{\alpha}}{\partial x} u_{\alpha} \\
&= \mu_{\alpha+1/2} \frac{\partial}{\partial x} \left( \frac{h_{\alpha}}{2} u_{\alpha} + \frac{h_{\alpha+1}}{2} u_{\alpha+1} \right) - \mu_{\alpha+1/2} \frac{\partial z_{\alpha+1}}{\partial x} u_{\alpha+1} + \mu_{\alpha+1/2} \frac{\partial z_{\alpha}}{\partial x} u_{\alpha} \\
&= \mu_{\alpha+1/2} \left( \frac{h_{\alpha}}{2} \frac{\partial u_{\alpha}}{\partial x} + \frac{h_{\alpha+1}}{2} \frac{\partial u_{\alpha+1}}{\partial x} \right) - \mu_{\alpha+1/2} \frac{\partial z_{\alpha+1/2}}{\partial x} (u_{\alpha+1} - u_{\alpha}).
\end{aligned}$$

In order to prove the energy balance (41) we use the results of prop. 3.2 obtained for the layer-averaged Euler system and it remains to consider the viscous and frictions terms multiplied by  $\mathbf{u}_{\alpha}$ .

Let us define  $\mathbf{R}_{\alpha}$

$$\begin{aligned}\mathbf{R}_\alpha &= \begin{pmatrix} R_{x,\alpha} \\ R_{y,\alpha} \end{pmatrix} = \nabla_{x,y} \cdot (h_\alpha \mathbf{T}_\alpha) - \mathbf{T}_{\alpha+1/2} \nabla_{x,y} z_{\alpha+1/2} + \mathbf{T}_{\alpha-1/2} \nabla_{x,y} z_{\alpha-1/2} \\ &\quad + \mu_{\alpha+1/2} \frac{\mathbf{u}_{\alpha+1} - \mathbf{u}_\alpha}{h_{\alpha+1} + h_\alpha} - \mu_{\alpha-1/2} \frac{\mathbf{u}_\alpha - \mathbf{u}_{\alpha-1}}{h_\alpha + h_{\alpha-1}} - \kappa_\alpha \mathbf{u}_\alpha.\end{aligned}$$

We write

$$\begin{aligned}R_{x,\alpha} u_\alpha &= \frac{\partial(u_\alpha h_\alpha T_{xx,\alpha})}{\partial x} + \frac{\partial(u_\alpha h_\alpha T_{xy,\alpha})}{\partial y} \\ &\quad + \mu_{\alpha+1/2} \frac{u_{\alpha+1} + u_\alpha}{2} \frac{u_{\alpha+1} - u_\alpha}{h_{\alpha+1} + h_\alpha} - \mu_{\alpha-1/2} \frac{u_\alpha + u_{\alpha-1}}{2} \frac{u_\alpha - u_{\alpha-1}}{h_\alpha + h_{\alpha-1}} \\ &\quad - \frac{\partial z_{\alpha+1/2}}{\partial x} T_{xx,\alpha+1/2} u_\alpha - \frac{\partial z_{\alpha+1/2}}{\partial y} T_{xy,\alpha+1/2} u_\alpha \\ &\quad + \frac{\partial z_{\alpha-1/2}}{\partial x} T_{xx,\alpha-1/2} u_\alpha + \frac{\partial z_{\alpha-1/2}}{\partial y} T_{xy,\alpha-1/2} u_\alpha \\ &\quad - h_\alpha T_{xx,\alpha} \frac{\partial u_\alpha}{\partial x} - h_\alpha T_{xy,\alpha} \frac{\partial u_\alpha}{\partial y} \\ &\quad - \frac{\mu_{\alpha+1/2}}{2} \frac{(u_{\alpha+1} - u_\alpha)^2}{h_{\alpha+1} + h_\alpha} - \frac{\mu_{\alpha-1/2}}{2} \frac{(u_\alpha - u_{\alpha-1})^2}{h_\alpha + h_{\alpha-1}} - \kappa_\alpha u_\alpha^2,\end{aligned}$$

and using the closure relation (39), it comes for  $\alpha = 2, \dots, N-1$

$$\begin{aligned}R_{x,\alpha} u_\alpha &= \frac{\partial(u_\alpha h_\alpha T_{xx,\alpha})}{\partial x} + \frac{\partial(u_\alpha h_\alpha T_{xy,\alpha})}{\partial y} \\ &\quad + \mu_{\alpha+1/2} \frac{u_{\alpha+1} + u_\alpha}{2} \frac{u_{\alpha+1} - u_\alpha}{h_{\alpha+1} + h_\alpha} - \mu_{\alpha+1/2} \frac{u_\alpha + u_{\alpha-1}}{2} \frac{u_\alpha - u_{\alpha-1}}{h_\alpha + h_{\alpha-1}} \\ &\quad - T_{xx,\alpha+1/2} \left( \frac{h_\alpha}{2} \frac{\partial u_\alpha}{\partial x} + \frac{\partial z_{\alpha+1/2}}{\partial x} u_\alpha \right) - T_{xy,\alpha+1/2} \left( \frac{h_\alpha}{2} \frac{\partial u_\alpha}{\partial y} + \frac{\partial z_{\alpha+1/2}}{\partial y} u_\alpha \right) \\ &\quad + T_{xx,\alpha-1/2} \left( -\frac{h_\alpha}{2} \frac{\partial u_\alpha}{\partial x} + \frac{\partial z_{\alpha-1/2}}{\partial x} \frac{u_\alpha}{2} \right) + T_{xy,\alpha-1/2} \left( -\frac{h_\alpha}{2} \frac{\partial u_\alpha}{\partial y} + \frac{\partial z_{\alpha-1/2}}{\partial y} \frac{u_\alpha}{2} \right) \\ &\quad - \frac{\mu_{\alpha+1/2}}{2} \frac{(u_{\alpha+1} - u_\alpha)^2}{h_{\alpha+1} + h_\alpha} - \frac{\mu_{\alpha-1/2}}{2} \frac{(u_\alpha - u_{\alpha-1})^2}{h_\alpha + h_{\alpha-1}} - \kappa_\alpha u_\alpha^2.\end{aligned}$$

An analogous relation can be easily obtained for  $R_{y,\alpha} v_\alpha$  and computing the sum over the layers of the obtained quantity, we get

$$\begin{aligned}\sum_{\alpha=1}^N \mathbf{R}_\alpha \cdot \mathbf{u}_\alpha &= \sum_{\alpha=1}^N \left( \frac{\partial(u_\alpha h_\alpha T_{xx,\alpha})}{\partial x} + \frac{\partial(u_\alpha h_\alpha T_{xy,\alpha})}{\partial y} + \frac{\partial(v_\alpha h_\alpha T_{yx,\alpha})}{\partial x} + \frac{\partial(v_\alpha h_\alpha T_{yy,\alpha})}{\partial y} \right) \\ &\quad - \sum_{\alpha=1}^{N-1} \frac{h_{\alpha+1} + h_\alpha}{2\mu} \left( T_{xx,\alpha+1/2}^2 + T_{xy,\alpha+1/2}^2 + T_{yx,\alpha+1/2}^2 + T_{yy,\alpha+1/2}^2 \right) \\ &\quad - \sum_{\alpha=1}^{N-1} \mu \frac{(\mathbf{u}_{\alpha+1} - \mathbf{u}_\alpha)^2}{h_{\alpha+1} + h_\alpha} - \kappa u_1^2,\end{aligned}$$

proving the result. ■

## Appendix B

Simple computations give

$$F_h = h \int_{\{y_1 \geq -\frac{\tilde{u}}{c}\} \times \mathbb{R}} (\tilde{u} + cy_1) \chi_0(y_1, y_2) dy_1 dy_2$$



$$\begin{aligned}
&= h \int_{y_1=-\frac{\tilde{u}}{c}}^{+\infty} (\tilde{u} + cy_1) \left( \int_{-\infty}^{+\infty} \chi_0(y_1, y_2) dy_2 \right) dy_1 = \frac{h}{\pi} \int_{y_1=-\frac{\tilde{u}}{c}}^{+\infty} (\tilde{u} + cy_1) \sqrt{1 - \frac{y_1^2}{4}} dy_1, \quad (139) \\
F_{hu} &= h \int_{\{y_1 \geq -\frac{\tilde{u}}{c}\} \times \mathbb{R}} (\tilde{u} + cy_1)(u + cn_x y_1) \chi_0(y_1, y_2) dy_1 dy_2 \\
&= \frac{h}{\pi} \int_{y_1=-\frac{\tilde{u}}{c}}^{+\infty} (\tilde{u} + cy_1)(u + cn_x y_1) \sqrt{1 - \frac{y_1^2}{4}} dy_1,
\end{aligned}$$

and likewise for  $F_{hv}$  we have

$$F_{hv} = \frac{h}{\pi} \int_{y_1=-\frac{\tilde{u}}{c}}^{+\infty} (\tilde{u} + cy_1)(v + cn_y y_1) \sqrt{1 - \frac{y_1^2}{4}} dy_1.$$

It is possible to obtain explicit formula for the expressions of  $F_h$ ,  $F_{hu}$  and  $F_{hv}$  since defining

$$I_1(z) = \int^z (u + cz) \sqrt{1 - \frac{z^2}{4}} dz, \quad I_2(z) = \int^z (u + cz)(v + cnz) \sqrt{1 - \frac{z^2}{4}} dz,$$

we have

$$I_1(z) = -\frac{4c}{3} \left(1 - \frac{z^2}{4}\right)^{3/2} + u \left( \frac{z}{2} \sqrt{1 - \frac{z^2}{4}} + \arcsin\left(\frac{z}{2}\right) \right),$$

and

$$\begin{aligned}
I_2(z) &= -c^2 n z \left(1 - \frac{z^2}{4}\right)^{3/2} + \frac{c^2 n + uv}{2} z \sqrt{1 - \frac{z^2}{4}} + (c^2 n + uv) \arcsin\left(\frac{z}{2}\right) \\
&\quad - \frac{4c(nu + v)}{3} \left(1 - \frac{z^2}{4}\right)^{3/2}.
\end{aligned}$$

Therefore, it comes

$$\begin{aligned}
F_h &= \begin{cases} 0 & \text{if } -\frac{\tilde{u}}{c} \geq 2 \\ \frac{h}{\pi} \tilde{u} \arcsin\left(\frac{\tilde{u}}{2c}\right) + \frac{h\tilde{u}}{2} + \frac{h}{c} \left(\frac{\tilde{u}^2}{6} + \frac{4}{3}c^2\right) \chi_0\left(\frac{\tilde{u}}{c}\right) & \text{if } -2 \leq -\frac{\tilde{u}}{c} \leq 2 \\ h & \text{if } -\frac{\tilde{u}}{c} \leq -2 \end{cases} \\
F_{hu} &= \begin{cases} 0 & \text{if } -\frac{\tilde{u}}{c} \geq 2 \\ \frac{h}{\pi} (c^2 n_x + u\tilde{u}) \arcsin\left(\frac{\tilde{u}}{2c}\right) + \frac{h}{2} (c^2 n_x + u\tilde{u}) \\ \quad + \frac{h}{12c} (2u\tilde{u}^2 - n_x \tilde{u}^3 + 16c^2 u + 10c^2 \tilde{u} n_x) \chi_0\left(\frac{\tilde{u}}{c}\right) & \text{if } -2 \leq -\frac{\tilde{u}}{c} \leq 2 \\ hu\tilde{u} + hc^2 n_x & \text{if } -\frac{\tilde{u}}{c} \leq -2 \end{cases} \\
F_{hv} &= \begin{cases} 0 & \text{if } -\frac{\tilde{u}}{c} \geq 2 \\ \frac{h}{\pi} (c^2 n_y + v\tilde{u}) \arcsin\left(\frac{\tilde{u}}{2c}\right) + \frac{h}{2} (c^2 n_y + v\tilde{u}) \\ \quad + \frac{h}{12c} (2v\tilde{u}^2 - n_y \tilde{u}^3 + 16c^2 v + 10c^2 \tilde{u} n_y) \chi_0\left(\frac{\tilde{u}}{c}\right) & \text{if } -2 \leq -\frac{\tilde{u}}{c} \leq 2 \\ hv\tilde{u} + hc^2 n_y & \text{if } -\frac{\tilde{u}}{c} \leq -2 \end{cases}
\end{aligned}$$