Contributions to the numerical approximation of shallow water asymptotics

Habilitation à Diriger des Recherches
Institut de Mathématiques et de Modélisation de Montpellier
Université Montpellier 2

Rapporteurs :

R. Abgrall, Professeur, Université de Zürich
D. Bresch, DR CNRS, Université de Savoie
C. Parès, Professeur, Université de Malaga

présenté le 11 décembre 2014, par Fabien Marche, devant le jury:

R. Abgrall, Professeur, Université de Zürich
F. Bouchut, DR CNRS, Université Paris-Est
D. Bresch, DR CNRS, Université de Savoie
P. Fabrie, Professeur, Université de Bordeaux
D. Lannes, DR CNRS, Université de Bordeaux
D. Le Roux, Professeur, Université Lyon 1
B. Mohammadi, Professeur, Université Montpellier 2

Contents

Introduction

Ce mémoire constitue une synthèse des travaux effectués depuis mon arrivée à l'Institut de Mathématiques et de Modélisation de Montpellier (I3M), en septembre 2007. Son objectif est avant tout de décrire et d'apporter une vision cohérente, à défaut d'exhaustive, des modèles et méthodes numériques introduits dans les articles de recherche, actes de conférence, ou preprints que j'ai publiés avec mes collaborateurs, ou récemment soumis à évaluation.

Le fil conducteur reliant l'ensemble de ces travaux est la nécessité de pouvoir réaliser des simulations numériques précises, à échelle réaliste, de phénomènes complexes en hydrodynamique littorale, en incluant les principaux phénomènes et transformations s'y déroulant, tels que le redressement des vagues, le déferlement, les interactions vagues-courant ou encore les phénomènes de submersion pendant des épisodes de tempêtes (voir Fig. 0.1). Comme nous allons le voir, cela nécessite d'aborder simultanément les domaines de la modélisation, de l'analyse numérique et du calcul scientifique.

Ce mémoire est structuré en 3 parties. La première partie a pour objectif d'introduire brièvement le lecteur aux cadres généraux (modélisation et méthodes numériques) dans lesquels se situent les travaux présentés dans les deux parties suivantes. C'est aussi l'occasion d'introduire les principales notations qui seront utilisées tout au long de ce mémoire.

La deuxième partie est dédiée à la construction de divers schémas numériques pour les équations de *Saint-Venant*, éventuellement avec termes sources (eux même éventuellement raides). Nous aborderons au passage des thématiques telles que modèles et schémas de relaxation, solveurs de *Riemann* approchés, schémas préservant les états d'équilibre ou encore modèles asymptotiques et schémas préservant ces asymptotiques.

La troisième partie est consacrée à l'étude de toute une hiérarchie de nouveaux modèles de type *Green-Naghdi*, et à leur approximation numérique. Ces modèles, caractérisés par la présence de termes dispersifs non-linéaires, nous permettent *a priori* de sortir du cadre de l'étude des systèmes hyperboliques. Toutefois, les premières amours étant tenaces, nous verrons que ces termes dispersifs peuvent être interprétés comme termes sources dans les équations de *Saint-Venant*, nous permettant de développer des méthodes d'approximations pertinentes.

J'ai tenté, d'une manière générale, de réduire les détails et passages techniques, en me limitant à une présentation des modèles et schémas et mettant en évidence leurs

principales propriétés. J'ai ressenti le besoin, dans certaines parties, de modifier la présentation telle qu'introduite dans les articles concernés, parfois en étendant légèrement le cadre initial, parfois au contraire en le restreignant. J'ai également voulu, dans un souci de lisibilité, faire un effort d'uniformisation des notations. Pour toutes ces raisons, que le lecteur curieux qui prendrait le temps de consulter les articles originaux ne s'étonne donc pas de trouver certaines dissemblances.

Certains travaux présentés dans la suite ont été effectués dans le cadre des thèses de Doctorat de *M.Tissier* [T1], effectuée à l'Université Bordeaux 1, et d' *A.Duran* [T3] à l'Université Montpellier 2, que j'ai respectivement co-encadrée et encadrée. Enfin, dans un souci de cohérence, j'ai choisi de ne pas détailler dans ce manuscript certains articles ([A4, A5, A7, A13, A17, A21]), traitant d'applications ou de sujets ne s'inscrivant pas directement dans la présentation qui va suivre.

Les modèles et schémas numériques décris dans ce manuscript ont été implémentés et validés au sein de divers codes de recherche et bibliothèques, rassemblés aujourd'hui dans la plateforme de calcul WaveBox [S1], qui je l'espère, continuera à être activement développée.

Figure 0.1: Exemple d'application du module Volume-Fini de WaveBox (anciennement SURF_WB), couplé à un modèle de houle à phase moyennée, effectuée dans le cadre du projet ANR MISEEVA (2009-2012): submersion de Palavas-les-Flots lors de la tempête de novembre 1982 ; Figure issue de [82].

1

A little background

In this first Chapter, we begin with a brief recall of the derivation of the asymptotic models in the shallow water regime, starting from the *Euler* equations with free surface conditions. A detailed derivation, together with the justification of the models, can be found in [62]. We then introduce the general discretization frameworks in which the following numerical methods will take place.

1.1 The models

1.1.1 Water waves equations

In an incompressible, homogeneous, irrotational and inviscid fluid under the influence of gravity, the propagation of surface waves is governed by the *Euler* equations with nonlinear boundary conditions at the surface and at the bottom. In the following, d refers to the spacial dimension of the surface of the fluid and takes the value 1 or 2.The water waves problem consists in describing the motion of the fluid in terms of the evolution of the free surface elevation above a reference level, parameterized by $\zeta(t,\mathbf{x}) \in \mathbb{R}$, and the velocity $\mathbf{V}(t,\mathbf{x},z) \in \mathbb{R}^{d+1}$ (with $\mathbf{x} = (x,y)$ if $d = 2$ and $\mathbf{x} = x$ if $d = 1$). The fluid domain is given by

$$\Omega(\zeta,b) = \{(\mathbf{x},z) \in \mathbb{R}^{d+1},\ -h_0 + b(\mathbf{x}) \leq z \leq \zeta(t,\mathbf{x})\},$$

where h_0 is a reference depth and $-h_0 + b(\mathbf{x})$ is a parameterization of the bottom, see Fig. 1.1.

The free surface *Euler* equations can be written as follows, in terms of ζ and $\mathbf{V}$:

$$\begin{cases} \partial_t \mathbf{V} + \mathbf{V}\cdot\nabla\mathbf{V} = -\mathbf{g} - \nabla P, \quad (\mathbf{x},z) \in \Omega,\ t \geq 0, & (1.1)\\ \nabla\cdot\mathbf{V} = 0, \quad (\mathbf{x},z) \in \Omega,\ t \geq 0 & (1.2)\\ \nabla\times\mathbf{V} = 0, \quad (\mathbf{x},z) \in \Omega,\ t \geq 0 & (1.3)\\ \partial_t\zeta - \sqrt{1+|\nabla_{\mathbf{x}}\zeta|^2}\mathbf{n}\cdot\mathbf{V}_{|z=\zeta(t,\mathbf{x})} = 0, \quad t \geq 0, & (1.4)\\ P_{|z=\zeta(t,\mathbf{x})} = P_{atm}, \quad t \geq 0, & (1.5)\\ \mathbf{n}\cdot\mathbf{V}_{|z=-h_0+b(\mathbf{x})} = 0, \quad t \geq 0, & (1.6) \end{cases}$$

where P stands for the pressure, $\mathbf{g}$ for the gravitational acceleration and $\mathbf{n}$ the normal vector pointing upwards the fluid domain.

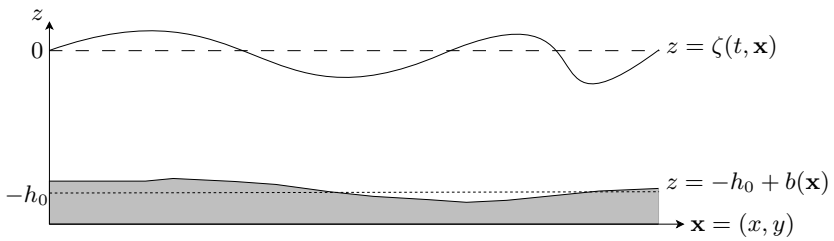

Figure 1.1: Sketch of the domain

1.1.2 Bernouilli's formulation

Using the incompressibility and irrotationality conditions, one can write the water wave equations under *Bernouilli*'s formulation in terms of a potential velocity ϕ (the velocity field is thus given by $\mathbf{V} = \nabla_{\mathbf{x},z}\phi$), for $t \geq 0$:

$$\begin{cases} \Delta_{\mathbf{x}}\phi + \partial_z^2\phi = 0, \quad (\mathbf{x}, z) \in \Omega, & (1.7) \\ \left[\partial_t\phi + \frac{1}{2}(|\nabla_{\mathbf{x}}\phi| + \partial_z\phi^2) + gz\right]_{|z=\zeta(t,\mathbf{x})} = 0, & (1.8) \\ \left[\partial_t\zeta + \nabla_{\mathbf{x}}\zeta \cdot \nabla_{\mathbf{x}}\phi - \partial_z\phi\right]_{|z=\zeta(t,\mathbf{x})} = 0, & (1.9) \\ \partial_{\mathbf{n}}\phi_{|z=-h_0+b(\mathbf{x})} = 0, & (1.10) \end{cases}$$

where $\partial_{\mathbf{n}}$ stands for the outward normal derivative at the bottom boundary.

Considering the four main length scales involved in this problem, namely the characteristic water depth h_0, the characteristic horizontal scale ℓ_0, the order of free surface amplitude a_0, and the order of bottom variations b_0, three independent dimensionless parameters are formed:

$$\text{the nonlinearity parameter:} \quad \varepsilon := \frac{a_0}{h_0}, \tag{1.11}$$

$$\text{the shallowness parameter:} \quad \mu := \left(\frac{h_0}{\ell_0}\right)^2, \tag{1.12}$$

$$\text{the bottom amplitude parameter:} \quad \beta := \frac{b_0}{h_0}. \tag{1.13}$$

All variables are normalized with the general nondimensionalization, which applies to any wave regime:

$$\mathbf{x} = \ell_0\mathbf{x}', \quad z = h_0\nu z', \quad t = \frac{\ell_0}{\sqrt{g\,h_0\,\nu}}t', \quad \zeta = a_0\zeta', \quad \phi = \frac{a_0\ell_0}{h_0}\left(\frac{gh_0}{\nu}\right)^{\frac{1}{2}}\phi', \quad b = b_0b',$$

where $\sqrt{g\,h_0\,\nu}$ and $\frac{a_0\ell_0}{h_0}\left(\frac{gh_0}{\nu}\right)^{\frac{1}{2}}$ are respectively the characteristic scales for the wave celerity and the potential velocity, with $\nu = \tanh(\mu^{\frac{1}{2}})/\mu^{\frac{1}{2}}$.

There are basically two main categories of asymptotic regimes: shallow water ($\mu \ll 1$) and deep water ($\mu \approx 1$). Within each of these categories, various sub-regimes can be identified depending on the assumptions made on the nonlinearity parameter ε, the topography parameter β. In the following, we will only focus on the *shallow*

water scaling, namely $\mu \ll 1$, leading to $\nu \simeq 1$. This regime, which is of particular interest in coastal oceanography, corresponds to the configuration where the wave length ℓ_0 of the flow is large compared to the typical depth h_0.
Anticipating on this assumption, we set $\nu = 1$. Injecting these nondimensionalized variables into equations (1.7-1.10), and omitting the primes, we obtain the nondimensionalized equations:

$$\begin{cases} \mu\Delta_{\mathbf{x}}\phi + \partial_z^2\phi = 0, \quad -1+\beta b < z < \varepsilon\zeta, & (1.14)\\ [\partial_t\phi + \frac{1}{2}(\varepsilon|\nabla_{\mathbf{x}}\phi|^2 + \frac{\varepsilon}{\mu}\partial_z\phi^2) + z]_{|z=\varepsilon\zeta(t,\mathbf{x})} = 0, & (1.15)\\ [\partial_t\zeta + \varepsilon\nabla_{\mathbf{x}}\zeta\cdot\nabla_{\mathbf{x}}\phi - \frac{1}{\mu}\partial_z\phi]_{|z=\varepsilon\zeta(t,\mathbf{x})} = 0, & (1.16)\\ [\partial_z\phi - \mu\beta\nabla_{\mathbf{x}}b\cdot\nabla_{\mathbf{x}}\phi]_{|z=-1+\beta b(\mathbf{x})} = 0. & (1.17) \end{cases}$$

1.1.3 *Zakharov-Craig-Sulem* formulation

Following [109], we introduce the trace of the velocity potential at the free surface $\psi = \phi_{|z=\varepsilon\zeta}$, and the *Dirichlet-Neumann* operator $\mathcal{G}_\mu[\varepsilon\zeta, \beta b]$ such that

$$\mathcal{G}_\mu[\varepsilon\zeta,\beta b]\psi = [-\mu\varepsilon\nabla_{\mathbf{x}}\zeta\cdot\nabla_{\mathbf{x}}\phi + \partial_z\phi]_{|z=\varepsilon\zeta}, \tag{1.18}$$

with ϕ the solution of the following boundary values problem

$$\begin{cases} \mu\Delta_{\mathbf{x}}\phi + \partial_z^2\phi = 0, \quad -1+\beta b \le z \le \varepsilon\zeta, & (1.19)\\ \phi_{|z=\varepsilon\zeta} = \psi, & (1.20)\\ \partial_{\mathbf{n}}\phi_{|z=-1+\beta b} = 0. & (1.21) \end{cases}$$

In particular, the *Dirichlet-Neumann* operator can be expressed as

$$\mathcal{G}_\mu[\varepsilon\zeta,\beta b]\psi = \sqrt{1+\varepsilon^2|\nabla_{\mathbf{x}}\zeta|^2}\,\partial_{\mathbf{n}}\phi_{|z=\varepsilon\zeta}.$$

The system (1.14)-(1.17) can be recasted as a system of two scalar evolution equations in terms of ζ and ψ:

$$\begin{cases} \partial_t\zeta - \frac{1}{\mu}\mathcal{G}_\mu[\varepsilon\zeta,\beta b]\psi = 0, & (1.22)\\ \partial\psi + \zeta + \frac{\varepsilon}{2}|\nabla_{\mathbf{x}}\psi|^2 - \varepsilon\mu\frac{\left(\frac{1}{\mu}\mathcal{G}_\mu[\varepsilon\zeta,\beta b]\psi + \varepsilon\nabla_{\mathbf{x}}\zeta\cdot\nabla_{\mathbf{x}}\psi\right)^2}{2(1+\varepsilon^2\mu|\nabla_{\mathbf{x}}\zeta|^2)} = 0. & (1.23) \end{cases}$$

It remains now to express the key relation between $\mathcal{G}_\mu[\varepsilon\zeta,\beta b]\psi$ and the depth averaged horizontal velocity defined as follows:

$$\mathbf{v}(t,\mathbf{x}) = \frac{1}{1+\varepsilon\zeta-\beta b}\int_{-1+\beta b}^{\varepsilon\zeta}\nabla_{\mathbf{x}}\phi\,dz, \tag{1.24}$$

leading to

$$-\frac{1}{\mu}\mathcal{G}_\mu[\varepsilon\zeta,\beta b]\psi = \nabla_{\mathbf{x}}\cdot(h\mathbf{v}), \tag{1.25}$$

where $h = 1+\varepsilon\zeta-\beta b$ denotes the water height. Follows the usual formulation of the first equation (the subscript $\mathbf{x}$ will be omitted from now on the differential operator ∇ and Δ):

$$\begin{cases} \partial_t \zeta + \nabla \cdot (h\mathbf{v}) = 0, & (1.26) \\ \partial_t \psi + \zeta + \frac{\varepsilon}{2}|\nabla\psi|^2 - \varepsilon\mu \frac{\left(\frac{1}{\mu}\mathcal{G}_\mu[\varepsilon\zeta, \beta b]\psi + \varepsilon\nabla\zeta\cdot\nabla\psi\right)^2}{2(1+\varepsilon^2\mu|\nabla\zeta|^2)} = 0. & (1.27) \end{cases}$$

In the following, these equations will be referred to as *water waves* equations. The (linear) dispersive properties of the water waves equations (1.26)-(1.27) depend on the set of all $(\omega, \mathbf{k}) \in \mathbb{R}\times\mathbb{R}^d$ such that there exists a plane wave solution $(\zeta_0, \psi_0)e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}$ to the linearization of (1.26)-(1.27) around the rest state $(\zeta = \zeta_0, \nabla\psi = 0)$, and for flat bottoms. The frequency dispersion is consequently ruled by the following dispersion relation:

$$\omega(\mathbf{k})_\pm = \pm\sqrt{|\mathbf{k}|^2 \frac{\tanh(\sqrt{\mu}|\mathbf{k}|)}{\sqrt{\mu}|\mathbf{k}|}}. \tag{1.28}$$

1.1.4 Shallow water asymptotics

Let us now focus on the *shallow water* assumption, corresponding to the scaling $\mu \ll 1$. We seek for an asymptotic expansion of ϕ:

$$\phi(t, \mathbf{x}, z) = \phi_0 + \mu\phi_1 + \mathcal{O}(\mu^2). \tag{1.29}$$

Plugging this expression into (1.19) and canceling the residual up to the order $\mathcal{O}(\mu^2)$, we obtain

$$\phi_0 = \psi \tag{1.30}$$

$$\phi_1 = (z-\varepsilon\zeta)\left(-\frac{1}{2}(z+\varepsilon\zeta) - 1 + \beta b\right)\Delta\psi + \beta(z-\varepsilon\zeta)\nabla b\cdot\nabla\psi. \tag{1.31}$$

Coming back to (1.24), and injecting the asymptotic approximation (1.29), we obtain a relation between $\mathbf{v}$ and $\nabla\psi$ of order $\mathcal{O}(\mu^2)$:

$$\nabla\psi = (I + \mu\mathcal{T}[h, b])\mathbf{v} + \mathcal{O}(\mu^2), \tag{1.32}$$

where the differential operator $\mathcal{T}[h, b](\cdot)$ is defined for all smooth enough R^d-valued function $\mathbf{u}$ by

$$\mathcal{T}[h, b]\mathbf{u} = \mathcal{R}_1[h, b](\nabla\cdot\mathbf{u}) + \beta\mathcal{R}_2[h, b](\nabla b\cdot\mathbf{u}), \tag{1.33}$$

with, for all smooth enough scalar-valued function w:

$$\mathcal{R}_1[h, b]w = -\frac{1}{3h}\nabla(h^3 w) - \beta\frac{h}{2}w\nabla b, \tag{1.34}$$

$$\mathcal{R}_2[h, b]w = \frac{1}{2h}\nabla(h^2 w) + \beta w\nabla b. \tag{1.35}$$

1.1.5 $\mathcal{O}(\mu^2)$ accuracy: the *Green-Naghdi* equations

The *Green-Naghdi* (GN) equations are now obtained gathering equations (1.25)-(1.26)-(1.27)-(1.32) and neglecting the $\mathcal{O}(\mu^2)$ terms:

$$\begin{cases} \partial_t \zeta + \nabla \cdot (h\mathbf{v}) = 0, & (1.36) \\ [I + \mu \mathcal{T}[h,b]]\partial_t \mathbf{v} + \nabla \zeta + \varepsilon (\mathbf{v} \cdot \nabla)\mathbf{v} + \varepsilon \mu \mathcal{Q}[h,b](\mathbf{v}) = 0, & (1.37) \end{cases}$$

where the nonlinear differential operator $Q[h,b](\cdot)$ is defined by

$$\mathcal{Q}[h,b](\mathbf{v}) = \mathcal{R}_1[h,b](\nabla \cdot (\mathbf{v}\nabla \cdot \mathbf{v}) - 2(\nabla \cdot \mathbf{v})^2) + \beta \mathcal{R}_2[h,b]((\mathbf{v} \cdot \nabla)^2 b). \quad (1.38)$$

These equations are obtained without any additional assumption on the bottom variations or the nonlinearity parameter ε and we still are in the *large amplitude regime*, also referred to as *fully nonlinear* regime $\varepsilon = O(1)$.
When no confusion is possible, we use in the following the simplified notations $\mathcal{T}$ and $\mathcal{Q}$ instead of $\mathcal{T}[h,b]$ and $\mathcal{Q}[h,b]$.

The corresponding equations (1.36-1.37) have been derived first by *Serre* [93] and then by *Su* and *Gardner* [97], *Miles* and *Salmon* [78], *Seabra-Santos et. al.* [92] and *Green* and *Naghdi* [49]. These equations are also called *fully nonlinear Boussinesq* equations. A rigorous mathematical justification of these models is given in [3].

An additional *weak nonlinearity* assumption $\varepsilon = \mathcal{O}(\mu)$ can be performed to simplify the previous equations. Without any assumption on the bottom variations (*i.e.* $\beta = \mathcal{O}(1)$), we recover the *Boussinesq* equations, as derived by *Peregrine*:

$$\begin{cases} \partial_t \zeta + \nabla \cdot (h\mathbf{v}) = 0, & (1.39) \\ (I + \mu \mathcal{T}[h_b, \beta b])\partial_t \mathbf{v} + \nabla \zeta + \varepsilon (\mathbf{v} \cdot \nabla)\mathbf{v} = 0, & (1.40) \end{cases}$$

with $h_b = 1 - \beta b$. Such assumption implies that $\varepsilon \ll 1$, hence the so-called *small amplitude* regime. Several additional assumptions and transformations can be further performed, allowing to recover some other widely used *Boussinesq*-type (BT) models, like for instance the equations of *Nwogu* [80] or *Madsen* and *Sorensen* [74].

1.1.6 $\mathcal{O}(\mu)$ accuracy: the Saint-Venant equations.

When $\mu \ll 1$ and without assumptions on ε and β, a further simplified model is obtained by neglecting all the $\mathcal{O}(\mu)$ terms and corresponds to the classical *Nonlinear Shallow Water* (NSW) equations or *Saint-Venant* equations:

$$\begin{cases} \partial_t \zeta + \nabla \cdot (h\mathbf{v}) = 0, & (1.41) \\ \partial_t \mathbf{v} + \nabla \zeta + \varepsilon (\mathbf{v} \cdot \nabla)\mathbf{v} = 0. & (1.42) \end{cases}$$

We point out that the asymptotic regime is the same for the GN and NSW equations. The only difference is the precision of the approximation.
Anticipating on the numerical discretization with conservative methods, these equations can be alternatively written in terms of conservative variables $h(t,\mathbf{x})$ and $h(t,\mathbf{x})\mathbf{v}(t,\mathbf{x})$):

$$\begin{cases} \partial_t h + \varepsilon \nabla \cdot (h\mathbf{v}) = 0, & (1.43) \\ \partial_t (h\mathbf{v}) + h\nabla \zeta + \varepsilon \nabla \cdot (h\mathbf{v} \otimes \mathbf{v})\Big) = 0. & (1.44) \end{cases}$$

When $d = 2$, it is convenient to denote

$$\mathbf{W}(t,\mathbf{x}) = (h(t,\mathbf{x}), h(t,\mathbf{x})\mathbf{v}(t,\mathbf{x}))^T, \quad \text{with} \quad \mathbf{v}(t,\mathbf{x}) = (v_1(t,\mathbf{x}), v_2(t,\mathbf{x}))^T$$

or alternatively

$$\mathbf{W}(t,\mathbf{x}) = (h(t,\mathbf{x}), \mathbf{q}(t,\mathbf{x}))^T \quad \text{with} \quad \mathbf{q}(t,\mathbf{x}) = (q_1(t,\mathbf{x}), q_2(t,\mathbf{x}))^T,$$

the state vector of conservative variables. Coming back to dimensionalized variables, and recalling that the water height $h = h_0 - b + \zeta$ is a positive quantity, we introduce the following convex set of *admissible states*:

$$\Theta = \left\{ \mathbf{W} = (h, \mathbf{v})^T \in \mathbb{R}^3;\ h \geq 0, \mathbf{v} \in \mathbb{R}^2 \right\}. \tag{1.45}$$

When there is no bottom variations, or when these are neglected (*i.e.* b=0), we obtain the *homogeneous* system associated with (1.43)-(1.44), which is know to be strictly hyperbolic for $h > 0$. This system can be written in a compact general conservative form as follows:

$$\partial_t \mathbf{W} + \nabla \cdot \mathcal{F}(\mathbf{W}) = 0, \tag{1.46}$$

where $\mathcal{F} : \Omega \to \mathbb{R}^3 \times \mathbb{R}^3$ is the flux function, which reads

$$\mathcal{F}(\mathbf{W}) = \begin{pmatrix} hv_1 & hv_2 \\ hv_1^2 + p(h) & hv_1v_2 \\ hv_1v_2 & hv_2^2 + p(h) \end{pmatrix}, \tag{1.47}$$

with $p(h) = gh^2/2$.

1.2 Numerical methods and notations

We introduce now some notations for the upcoming discrete approximations of the solutions of (1.46) in Chapter 2 and (1.36)-(1.37) in Chapter 3. We assume the reader to be familiar with the basis of the Finite-Volume methods (FVMs) or discontinuous Galerkin methods (DGMs) which are applied in the following. Otherwise, we refer to the reference textbooks [18, 47, 45, 67] for FVMs and to [37] for DGMs.
From now, we consider *boundary value problems* associated with the previous asymptotics, defined on a convex bounded domain Ω in $\mathbb{R}^d$. For the sake of simplicity, in this report, we completely leave the important issues of boundary conditions out. As for the time-discretization, all the following fully discrete schemes are introduced with a 1^{st}-order *Euler* scheme for time discretization. When higher-order space discretizations are investigated, we assume that some *Runge-Kutta* (RK) methods (or Strong Stability Preserving *Runge-Kutta* (SSP-RK) methods, as far as stability is concerned) of consistent order of accuracy are used.

1.2.1 The mesh

When $d = 2$, we assume for the sake of simplicity that the computational domain Ω is a polygonal domain (*i.e.* its boundary $\partial\Omega$ is a piecewise linear polygon), so that

Ω can be exactly partitioned using N_t geometrically-conforming non-overlapping straight-sided triangles:

$$\Omega = \bigcup_{l=1}^{N_t} T^l.$$

The triangulation is denoted $\mathcal{P}_\mathfrak{h}$, $\{S_l\}_{1\leq l\leq N_v}$ refers to the associated N_v vertices, and $\mathfrak{h}$ to a global mesh length-scale. The *primary mesh* exactly coincides with this triangular partition.

In a second time, we build a set of *control-volumes* defined around the vertices of the primary mesh, resulting in the *dual mesh*. For a given vertex S_l, the associated control volume is obtained by connecting the centroids of the surrounding triangles having S_l as a common vertex, see Fig. 1.2. Special treatments are needed for the boundary vertices, not detailed here.

In the following of this manuscript, we study FVMs built on the *dual mesh* and DGMs built on the *primary mesh*. In both cases, $\{\mathcal{C}_l,\ 1 \leq l \leq N_e\}$ refer to the set of *elemental domain*, which identifies either to the set of control-volumes for the FVMs (*i.e.* $N_e = N_v$) or to the triangulation $\mathcal{P}_\mathfrak{h}$ for the DGMs (*i.e.* $N_e = N_t$). The element $\mathcal{C}_l$ has a boundary denoted $\partial\mathcal{C}_l$, a unit outward normal vector $\mathbf{n}_l$, an area $|\mathcal{C}_l|$ and a perimeter $\mathfrak{p}_l$.

Additionally, we define the following notations, for the neighborhood of $\mathcal{C}_l$:

- $\Lambda(l)$: number of adjacent elements ($\Lambda(l) = 3$ for DGMs),
- $K(l) = \{\sigma(k),\ 1 \leq k \leq \Lambda(l)\}$, with $\sigma(k)$ the index of the elements $\mathcal{C}_{\sigma(k)}$ adjacent to $\mathcal{C}_l$,
- $\Gamma_{l\sigma}$: the boundary interface defined by $\mathcal{C}_l$ and $\mathcal{C}_\sigma$, so that $\partial\mathcal{C}_l = \bigcup_{k=1}^{\Lambda(l)} \Gamma_{l\sigma(k)}$,
- $\ell_{l\sigma}$: the length of $\Gamma_{l\sigma}$, so that $\mathfrak{p}_l = \sum_{k=1}^{\Lambda(l)} \ell_{l\sigma(k)}$,
- $\mathbf{n}_{l\sigma}$: the unit normal to $\Gamma_{l\sigma}$, pointing to $\mathcal{C}^\sigma$.
- M_{ijk}: the mass center of the triangle $S_i\, S_j\, S_k$
- m_{ijk}: the middle of the edge $S_i\, M_{ijk}$

Note that the element $\mathcal{C}_l$ can be sub-divided into $\Lambda(l)$ triangles, denoted $T_{l\sigma(k)}$, for $1 \leq k \leq \Lambda(l)$, defined by joining the interfaces $\Gamma_{l\sigma}$ to the vertex S_l.

When $d = 1$, we consider a partition $\mathcal{P}_\mathfrak{h}$ of the computational domain into N_e non-overlapping segments, still referred to as *elements*, which are defined through a sequence of grid points:

$$\ldots < x_{i-\frac{3}{2}} < x_{i-\frac{1}{2}} < x_{i+\frac{1}{2}} < x_{i+\frac{3}{2}} < \ldots$$

The elements will be alternatively denoted by $\{\mathcal{C}_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]\}_{1\leq i\leq N_e}$ for FVMs, and $\{\mathcal{C}_i = [x_{i,l}, x_{i,r}]\}_{1\leq i\leq N_e}$ for DGMs. In both cases, we denote x_i the center of $\mathcal{C}_i$. Note that $|\mathcal{C}_i| = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}} = x_{i,r} - x_{i,l}$, $\partial\mathcal{C}_i$ is reduced to the 2 boundary nodes of $\mathcal{C}_i$ and that $\mathbf{n}_i$ is reduced to $+1$ and -1 respectively at the right and the left boundary.

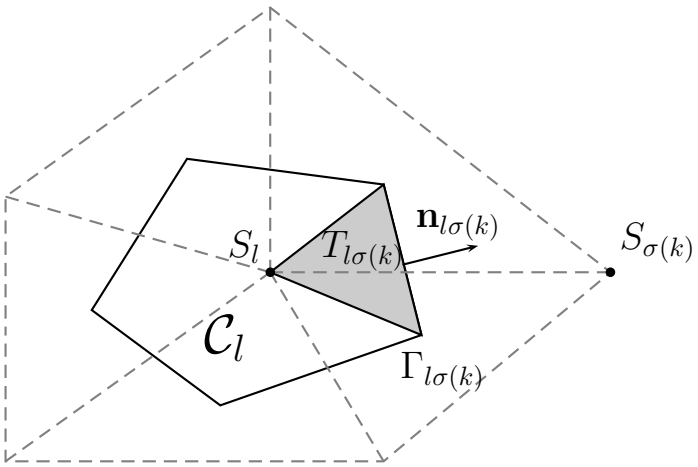

Figure 1.2: A dual cell.

1.2.2 Finite-Volume discretization

2d formulation

At the discrete time t^n, the FV discrete values $\mathbf{W}_i^n$ are approximations of the averages (piecewise constant interpolation) of the exact solutions over the elements:

$$\mathbf{W}_i^n = \frac{1}{|\mathcal{C}_i|}\int_{\mathcal{C}_i} \mathbf{W}(t^n, x)dx,$$

and we can classically obtain a fully discrete FV formulation by integrating (1.46) over $[t^n, t^{n+1}] \times \mathcal{C}_i$:

$$\mathbf{W}_i^{n+1} = \mathbf{W}_i^n - \frac{\Delta t}{|\mathcal{C}_i|}\sum_{k=1}^{\Lambda(i)} \ell_{ij(k)} F_{ij(k)}, \tag{1.48}$$

where Δt the time step and $F_{ij} = F(\mathbf{W}_i^n, \mathbf{W}_j^n, \mathbf{n}_{ij})$ is the normal numerical flux function through the interface between cells $\mathcal{C}_i$ and $\mathcal{C}_j$, *i.e.* a first order approximation of the exchanging term

$$\frac{1}{\ell_{ij}}\int_{\Gamma_{ij}} \mathcal{F}(\mathbf{W}) \cdot \mathbf{n}_{ij}\, ds,$$

computed considering a pseudo 1d scheme along the normal direction $\mathbf{n}_{ij}$.

1d formulation

For 1d approximations, the notations are simplified. The vector state of conservative variable is denoted $\mathbf{w}(t,x) = (h(t,x), q(t,x))^T = (h(t,x), h(t,x)v(t,x))^T$ and the convex set of admissible state is

$$\Theta = \left\{\mathbf{w} = (h,v)^T \in \mathbb{R}^2;\ h \geq 0, v \in \mathbb{R}\right\}. \tag{1.49}$$

Equations (1.46) are now written

$$\partial_t \mathbf{w} + \partial_x \mathcal{F}(\mathbf{w}) = 0, \tag{1.50}$$

with the flux $\mathcal{F} : \Omega \to \mathbb{R}^2$ defined as $\mathcal{F}(\mathbf{w}) = (hv, hv^2 + p(h))^T$.

The conservative formulation of the FV scheme (1.48) reduces to

$$\mathbf{w}_i^{n+1} = \mathbf{w}_i^n - \frac{\Delta t}{\Delta x}(F_{i+\frac{1}{2}} - F_{i-\frac{1}{2}}). \tag{1.51}$$

At 1^{st}-order, we set $F_{i+\frac{1}{2}} = F(\mathbf{w}_i^n, \mathbf{w}_{i+1}^n)$, where the function $(\mathbf{w}_L, \mathbf{w}_R) \mapsto F(\mathbf{w}_L, \mathbf{w}_R)$ is a numerical flux function which determines the scheme. Classically, this numerical flux has to be conservative, Lipschitz continuous, monotone and consistent with the exact flux.

1.2.3 Approximate *Riemann* solvers

There is several ways to build numerical flux formulae for (1.51). In the present work, we consider the formalism introduced by *Harten, Lax* and *Van Leer* (HLL) in [52]. We consider a constant space discretization step denoted by Δx, and we assume that a piecewise constant approximation $\mathbf{w}^{\Delta x}(x, t^n) \in \Omega$ of the solution of (1.50) is known at time t^n:

$$\mathbf{w}^{\Delta x}(x, t^n) = \mathbf{w}_i^n \text{ if } x \in [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}].$$

To evolve this approximation in time for $t > t^n$, we consider an approximate *Riemann* solver stated at each interface $x_{i+\frac{1}{2}}$. This approximate *Riemann* solver, denoted by $\mathbf{w}_{\mathcal{R}}(\frac{x}{t}; \mathbf{w}_L, \mathbf{w}_R)$ has the following structure:

$$\mathbf{w}_{\mathcal{R}}\Big(\frac{x}{t}; \mathbf{w}_L, \mathbf{w}_R\Big) = \begin{cases} \mathbf{w}_L, & \text{if } \frac{x}{t} < \lambda^-, \\ \mathbf{w}^*(\frac{x}{t}; \mathbf{w}_L, \mathbf{w}_R), & \text{if } \lambda^- < \frac{x}{t} < \lambda^+, \\ \mathbf{w}_R, & \text{if } \frac{x}{t} > \lambda^+, \end{cases} \tag{1.52}$$

where λ^- and λ^+ are respectively the minimum and maximum velocity waves involved in the solver. The intermediate state $\mathbf{w}^*$, describes the approximate solution inside the dependence cone characterized by $x = \lambda^- t$ and $x = \lambda^+ t$.
Following [52], we recall that at $t = \Delta t$, the approximate *Riemann* solver (1.52) must satisfy a consistency condition given by:

$$\frac{1}{\Delta x}\int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \mathbf{w}_{\mathcal{R}}\Big(\frac{x}{\Delta t}; \mathbf{w}_L, \mathbf{w}_R\Big)dx = \frac{1}{2}(\mathbf{w}_L + \mathbf{w}_R) - \frac{\Delta t}{\Delta x}\Big(\mathcal{F}(\mathbf{w}_R) - \mathcal{F}(\mathbf{w}_L)\Big), \tag{1.53}$$

where Δt is restricted by the CFL-like condition:

$$\frac{\Delta t}{\Delta x}\max(|\lambda^-|, |\lambda^+|) \leq \frac{1}{2}.$$

Now, assuming that an approximate solver $\mathbf{w}_{\mathcal{R}}(\frac{x - x_{i+\frac{1}{2}}}{t - t^n}; \mathbf{w}_i, \mathbf{w}_{i+1})$ is stated at each cell interface $x_{i+\frac{1}{2}}$, we thus obtain a non-interacting juxtaposition of *Riemann* solvers as soon as the following CFL-like condition is enforced:

$$\frac{\Delta t}{\Delta x}\lambda_o \leq \frac{1}{2}, \tag{1.54}$$

with $\lambda_o = \max_i(|\lambda^-_{i+\frac{1}{2}}|, |\lambda^+_{i+\frac{1}{2}}|)$. We define an approximate solution at time $t^n + t$ for all $t \in (0, \Delta t)$ as follows:

$$\mathbf{w}^{\Delta x}(x, t^n + t) = \mathbf{w}_{\mathcal{R}}\Big(\frac{x - x_{i+\frac{1}{2}}}{t}; \mathbf{w}_i^n, \mathbf{w}_{i+1}^n\Big) \text{ if } x \in [x_i, x_{i+1}]. \tag{1.55}$$

The projection of this solution on the piecewise constant functions gives the expected updated approximation:

$$\mathbf{w}_i^{n+1} = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \mathbf{w}^{\Delta x}(x, t^n + \Delta t) dx. \tag{1.56}$$

Introducing the condition (1.53), the scheme can be rewritten into a standard conservation form:

$$\mathbf{w}_i^{n+1} = \mathbf{w}_i^n - \frac{\Delta x}{\Delta t}\Big(F_{i+\frac{1}{2}} - F_{i-\frac{1}{2}}\Big), \tag{1.57}$$

where we have:

$$F_{i+\frac{1}{2}} = \mathcal{F}(U_i^n) + \frac{\Delta x}{2\Delta t}\mathbf{w}_i^n - \frac{1}{\Delta t}\int_{x_i}^{x_{i+\frac{1}{2}}} \mathbf{w}_{\mathcal{R}}\Big(\frac{x - x_{i+\frac{1}{2}}}{\Delta t}; \mathbf{w}_i^n, \mathbf{w}_{i+1}^n\Big) dx. \tag{1.58}$$

Many choices for (1.52) can be found in the literature. The exact *Godunov* scheme enters the proposed definition with λ^- and λ^+ the exact minimum and maximum characteristic speeds and $\mathbf{w}^*$ is the exact *Riemann* solution into the dependence cone.

Other examples are the one-intermediate state HLL scheme [52], where $\mathbf{w}^*$ is given by the following constant state:

$$\mathbf{w}^*\Big(\frac{x}{t}; \mathbf{w}_L, \mathbf{w}_R\Big) = \frac{\lambda^+\mathbf{w}_R - \lambda^-\mathbf{w}_L}{\lambda^+ - \lambda^-} - \frac{1}{\lambda^+ - \lambda^-}\Big(\mathcal{F}(\mathbf{w}_R) - \mathcal{F}(\mathbf{w}_L)\Big),$$

and $\lambda^\pm$ are estimations of the maximum and minimum wave speeds, the HLLC scheme proposed by [100] or the relaxation schemes, see for instance [18, 55, 68]. The well-known *Lax-Friedrichs* flux (LF in the following) also enters this framework, with the choice $\lambda^\pm = \pm\lambda_o$, leading to the following simple numerical flux, widely used throughout this manuscript:

$$F(\mathbf{w}_L, \mathbf{w}_R) = \frac{1}{2}\left(\mathcal{F}(\mathbf{w}_L) + \mathcal{F}(\mathbf{w}_R)\right) - \frac{\lambda_o}{2}\left(\mathbf{w}_R - \mathbf{w}_L\right). \tag{1.59}$$

One method to obtain such solvers is to use *relaxation models* and construct associated relaxation schemes, see for instance [34, 55, 83, 68] and the next section.

Finally, an interesting feature of the approximate *Riemann* solver framework is that it allows to study the existence of discrete entropy inequalities as well as the preservation of invariant domains. We recall the following result, adapted in the framework of NSW equations:

Proposition 1. *Let $\mathbf{w}_L$ and $\mathbf{w}_R$ be two constant states such that $h_L,\ h_R \geq 0$ and we consider an approximate Riemann solver (1.52), with the additional assumption that $\mathbf{w}^*(\frac{x}{t}; \mathbf{w}_L, \mathbf{w}_R) \in \Theta$ for all $\lambda^- < \frac{x}{t} < \lambda^+$. We consider a FV scheme (1.57) build on this approximate Riemann solver. Then, under the CFL restriction (1.54), we have $h_i^{n+1} \geq 0$ for all i , as soon as $h_i^n \geq 0$ for all i.*

1.2.4 Relaxation models

The relaxation method is a recent approach to approximate the weak solutions of hyperbolic systems of conservation laws, see [55]. There is a large amount of literature concerning the relaxation approach, and relaxation schemes. We refer to some of the reference papers [16, 26, 51, 71] and the reference textbook [19].
The underlying idea is to approximate the weak solutions of hyperbolic system of conservation laws, denoted equilibrium system, by the weak solutions of a relevant first order system with singular perturbations, namely *the relaxation model*. The relaxation schemes, built from the relaxation models, appear as a powerful tool. Indeed, this approach is of great help to construct approximate *Riemann* solvers in the sense of §1.2.3, and therefore to establish discrete entropy inequalities and maximum principles, even for difficult problems, see for instance [11]. However, one of the main difficulties generally arises from the choice of the relaxation model.

Starting from an hyperbolic system under the form (1.46), with $\mathbf{W}(t,\mathbf{x}) \in \mathbb{R}^n$, an associated relaxation system is another hyperbolic system of conservation laws in higher dimension, let say $m > n$, written as

$$\partial_t \mathbf{U}^\mu + \partial_x \mathcal{H}(\mathbf{U}^\mu) = \mu \mathcal{R}(\mathbf{U}^\mu), \tag{1.60}$$

with $\mathbf{U}^\mu \in \Theta_r \subset \mathbb{R}^m$, Θ_r being the convex set of admissible states associated with $\mathbf{U}^\mu$, and $\mathcal{H}(\mathbf{U}^\mu) \in R^m$. We assume the existence of a linear operator $\mathcal{L} : \mathbb{R}^m \to \mathbb{R}^n$ and, for any $\mathbf{W} \in \Theta$, of an equilibrium data $\mathcal{M}(\mathbf{W}) \in \Theta_r$, usually called the *maxwellian equilibrium*, such that

$$\mathcal{L}\mathcal{M}(\mathbf{W}) = \mathbf{W}, \quad \text{and} \quad \mathcal{L}\mathcal{H}(\mathcal{M}(\mathbf{W})) = \mathcal{F}(\mathbf{W}) \quad \forall \mathbf{W} \in \Theta.$$

Additionally, the second member $\mathcal{R}$ is such that

$$\mathcal{L}\mathcal{R}(\mathbf{U}^\mu) = 0, \quad \text{and} \quad \mathcal{R}(\mathbf{U}^\mu) = 0 \Leftrightarrow \mathbf{U}^\mu = \mathcal{M}(\mathbf{W}) \text{ for some } \mathbf{W} \in \Theta.$$

So far, the relaxation approach relies on the idea that $\mathbf{W} = \mathcal{L}\mathbf{U}$ is an approximate solution of (1.46) when $\mathbf{U}$ is a solution of (1.60) close to equilibrium.

As stated above, one of the main interests of relaxation models is the construction of approximate *Riemann* solvers for nonlinear systems, by *relaxing* (some of) the nonlinearities. Starting from a linear (or nonlinear but with only linearly degenerated fields) relaxation model of the form (1.60), we can consider a solution of (1.60) close to equilibrium and show that such a solution is a viscous approximation of the solution of the initial system (1.46) (*i.e.* a solution of (1.46) supplemented by a small diffusion term of order $1/\mu$), under a suitable *stability* condition. We refer to [19] for a review of the stability conditions for relaxation systems.

To build a numerical scheme for (1.60), we update the solution from discrete values at time t^n to time t^{n+1} with the standard splitting approach of [55]:

1. *the relaxation step:* at $t = t^n$, we solve $\partial_t \mathbf{U}^\mu = \mu \mathcal{R}(\mathbf{U}^\mu)$ while taking data at equilibrium (*i.e.*, $\mu \to \infty$) to go to t^{n+1-}.

2. *the transport step:* at $t = t^{n+1-}$, we solve (1.60) with $\mu = 0$ to go to from t^{n+1}.

At last, considering the fully discrete approximation, the *Riemann* solver for (1.60), obtain during the *transport step*, results in the definition of an approximate solver for (1.46). This is formalized for instance in [18], although in a slightly different formalism, where right hand sides are not considered in (1.60).

1.2.5 Discontinuous *Galerkin* approximations of 1^{st}-order systems

The DGMs for systems of hyperbolic conservation laws combine advantageous features commonly associated with Finite-Element methods (FEMs) and FVMs. As in the FVMs, the physics of waves propagation is accounted for by solving the *Riemann* problems that arise from the discontinuous representation of the solution at elements' interfaces. And as in classical FEMs, a high order of accuracy is obtained by means of polynomial approximations of arbitrary order within an element rather than by wide stencils.

The motivations of using of DGMs for the simulation of water waves problems are manifolds. Besides the fact that high-order methods are often desirable when considering potential savings in computational time – especially for large-scale problems involving long-time integration, DGMs can be applied on unstructured grids and handle non-conforming elements, are highly parallelizable (there is little inter-cell communications) and can easily handle adaptive strategies (the order of accuracy can vary from element to element). On top of that, they have several interesting properties with respect to conservation, stability, and convergence.

Discretization

To implement a DG scheme, we aim at computing an approximated solution $\mathbf{W}_\mathfrak{h}$, on the partition $\mathcal{P}_\mathfrak{h}$, in a trial functions space $\mathcal{V}_\mathfrak{h}$ to be defined. In the standard *Galerkin* formulation $\mathbf{W}_\mathfrak{h}$ is sought in a finite-dimensional subspace of the space of compactly supported continuous functions. However, hyperbolic problems generally have solutions in spaces of bounded variation, and the best one can hope for is that solutions are piecewise smooth. Consequently, a more adapted formulation is one where the global trial space contains discontinuous functions. We define the space of piecewise continuous polynomial functions

$$\mathcal{V}_\mathfrak{h} := \{v \in L^2(\Omega) \,|\, \forall \mathcal{C} \in \mathcal{P}_\mathfrak{h}\,,\, v_{|\mathcal{C}} \in \mathbb{P}^p(\mathcal{C})\}\,, \tag{1.61}$$

where $\mathbb{P}^p(\mathcal{C})$ denotes the space of d-variables polynomials in the element $\mathcal{C}$ of degree at most p. As in the standard *Galerkin* formulation, the space of test functions is defined similarly.

We form the following *residual* by injecting the sought approximated solution into (1.46)

$$\mathcal{R}_\mathfrak{h} = \partial_t \mathbf{W}_\mathfrak{h} + \nabla \cdot \mathcal{F}(\mathbf{W}_\mathfrak{h}), \tag{1.62}$$

and we require this residual to be orthogonal to all test functions in $\mathcal{V}_\mathfrak{h}$, resulting in the local statement:

$$\int_{\mathcal{C}_l} \mathcal{R}_\mathfrak{h} \phi_\mathfrak{h}\, dx = 0, \forall \phi_\mathfrak{h} \in \mathcal{V}_\mathfrak{h},\ 1 \leq l \leq N_e.$$

Integrating by part the flux term, we obtain the following discrete weak formulation: *find* $\mathbf{W}_\mathfrak{h} \in (\mathcal{V}_\mathfrak{h})^{d+1}$ *such that,* $\forall \phi_\mathfrak{h} \in \mathcal{V}_\mathfrak{h}$, *and* $\forall\, \mathcal{C}_l \in \mathcal{P}_\mathfrak{h}$ *we have:*

$$\int_{\mathcal{C}_l} \partial_t \mathbf{W}_\mathfrak{h} \phi_\mathfrak{h} dx - \int_{\mathcal{C}_l} \mathcal{F}(\mathbf{W}_\mathfrak{h}) \cdot \nabla \phi_\mathfrak{h} dx + \sum_{k=1}^{3} \int_{\Gamma_{l\sigma(k)}} \mathcal{F}(\mathbf{W}_\mathfrak{h}) \cdot \mathbf{n}_{l\sigma(k)}\, \phi_\mathfrak{h}\, ds = 0. \quad (1.63)$$

To keep the solution within a space of bounded variations and to allow information to propagate between elements, some elemental coupling arises as a result of the way boundary conditions on each element are applied. Boundary conditions are enforced through the normal component of the fluxes $\mathcal{F}(\mathbf{W}_\mathfrak{h}) \cdot \mathbf{n}_{l\sigma(k)}$ appearing in the boundary terms. As the solution may be discontinuous, we have two possible values of $\mathbf{W}_\mathfrak{h}$ at at each interface, denoted by $\mathbf{W}_\mathfrak{h}^-$ and $\mathbf{W}_\mathfrak{h}^+$, which are respectively the restrictions of $\mathbf{W}_{\mathfrak{h}|\mathcal{C}_l}$ and $\mathbf{W}_{\mathfrak{h}|\mathcal{C}_{\sigma(k)}}$ to $\Gamma_{l\sigma(k)}$ (*i.e.* the *interior* and *exterior* traces, with respect to the element $\mathcal{C}_l$).

To generate a stable scheme, upwinding considerations based on the natural propagation of information help to compute this flux and approximate *Riemann* solvers can be used. The normal component of the boundary flux $\mathcal{F}(\mathbf{W}_\mathfrak{h}) \cdot \mathbf{n}_{l\sigma(k)}$ is approximated by $F(\mathbf{W}_\mathfrak{h}^-, \mathbf{W}_\mathfrak{h}^+, \mathbf{n}_{l\sigma(k)})$ where F is a numerical flux function as described in §1.2.3. This approximation is denoted $\widehat{F}_{l\sigma(k)}$ in the following.
The local approximated vector solution $\mathbf{W}_{\mathfrak{h}|\mathcal{C}_l} \in (\mathcal{V}_\mathfrak{h})^{d+1}$ is formulated as a polynomial of order p on each element:

$$\mathbf{W}_{\mathfrak{h}|\mathcal{C}_l}(\mathbf{x}, t) = \sum_{i=1}^{N_p} \widetilde{\mathbf{W}}_i^l(t) \theta_i^l(\mathbf{x}), \quad \forall \mathbf{x} \in \mathcal{C}_l, \forall t \in [0, t_{max}], \quad (1.64)$$

where $\{\theta_i^l\}_{i=1}^{N_p}$ is a polynomial expansion basis for $\mathbb{P}^d(\mathcal{C}_l)$, and $\{\widetilde{\mathbf{W}}_i^l(t)\}_{i=1}^{N_p}$ are the local expansion coefficient vectors, with $N_p = p + 1$ when $d = 1$ and $N_p = (p + 1)(p+2)/2$ when $d = 2$. Many choices are possible for the expansion basis, including orthogonal and interpolant polynomials.
Once, the basis functions are chosen, the discrete weak formulation (1.63) becomes equivalent to the following problem: *find* $\{\widetilde{\mathbf{W}}_i^l(t)\}_{i=1}^{N_p} \in \mathbb{R}^{N_p}$ *such that,* $\forall\, \mathcal{C}_l \in \mathcal{P}_\mathfrak{h}$ and $\forall\ 1 \leq j \leq N_p$ *we have:*

$$\sum_{i=1}^{N_p} \left(\frac{d}{dt} \widetilde{\mathbf{W}}_i^l(t) \int_{\mathcal{C}_l} \theta_i^l \theta_j^l\, d\mathbf{x} \right) - \int_{\mathcal{C}_l} \mathcal{F}(\mathbf{W}_\mathfrak{h}) \cdot \nabla \theta_j^l\, d\mathbf{x} + \sum_{k=1}^{3} \int_{\Gamma_{l\sigma(k)}} \widehat{F}_{l\sigma(k)}\, \theta_j^l\, ds = 0. \quad (1.65)$$

When $d = 1$, (1.65) reduces to: *find* $\{\tilde{\mathbf{w}}_i^l(t)\}_{i=1}^{N_p} \in \mathbb{R}^{N_p}$ *such that,* $\forall\, \mathcal{C}_j \in \mathcal{P}_\mathfrak{h}$ *we have:*

$$\sum_{i=1}^{N_p} \frac{d}{dt} \tilde{\mathbf{w}}_i^j(t) \int_{\mathcal{C}_j} \theta_i^j \theta_k^j dx - \int_{\mathcal{C}_j} \mathcal{F}(\mathbf{w}_\mathfrak{h}) \frac{d}{dx} \theta_k^j dx + [\widehat{F} \theta_k^j]_{x_l^j}^{x_r^j} = 0, \quad 1 \leq k \leq N_p. \quad (1.66)$$

We recall that the occurrence of discontinuities in the propagating waves can induce *Gibbs*-type oscillations in the vicinity of discontinuities, which can lead to numerical instabilities and unbounded solutions. These oscillations can be suppressed by adapting the ideas of slope limiters from FVMs, designed to restrict or suppress oscillations near discontinuities through a non-linear procedure based on comparing elemental solution slopes and curvatures, with those of neighboring elements.

However, classical techniques of flux limiting coming from FVMs are not directly applicable to DGMs because of the presence of integral surface terms in the formulation. Therefore, slope limiters are not integrated in the computation of the residual, but are applied as a post-process. We do not focus on this issue in this manuscript. A survey of the available limiting methods is performed in [A16], as well as a documented implementation strategy. Some additional remarks are done concerning on-going works in the last chapter.

1.2.6 On the DG discretization of higher order derivatives

Although DGMs were initially designed for the approximation of the solutions of hyperbolic PDE's, recent works have led to formulations for the approximations of parabolic, elliptic and even problems with derivatives higher than 2. The pioneering work was certainly [8], in which the DG formulation had to be extended to handle the viscous terms in the *Navier-Stokes* equations. A recent review is performed in [106] and a unified analysis can be found in [4], and [40, 41].
To further illustrate the ideas, let us consider the following 2^{nd}-order elliptic problem with *Dirichlet* boundary conditions:

$$\begin{cases} -\Delta u = f \quad \text{in } \Omega, \\ u = 0, \quad \text{on } \partial\Omega, \end{cases} \tag{1.67}$$

with $f \in L^2(\Omega)$. An auxiliary variable is introduced, allowing to write (1.67) as a first order system:

$$\begin{cases} \mathbf{q} = \nabla u \quad \text{in } \Omega, \\ -\nabla \cdot \mathbf{q} = f \quad \text{in } \Omega, \\ u = 0, \quad \text{on } \partial\Omega. \end{cases} \tag{1.68}$$

Looking for a discretization of (1.68), the discrete weak formulation can be stated as: *find* $(u_\mathfrak{h}, \mathbf{q}_\mathfrak{h}) \in \mathcal{V}_\mathfrak{h} \times (\mathcal{V}_\mathfrak{h})^2$ such that $\forall (\phi_\mathfrak{h}, \pi_\mathfrak{h}) \in \mathcal{V}_\mathfrak{h} \times (\mathcal{V}_\mathfrak{h})^2$, *and* $\forall \mathcal{C}_l \in \mathcal{T}_\mathfrak{h}$ *we have:*

$$\int_{\mathcal{C}_l} \mathbf{q}_\mathfrak{h} \cdot \pi_\mathfrak{h} \, dx + \int_{\mathcal{C}_l} u_\mathfrak{h} \nabla \cdot \pi_\mathfrak{h} \, dx = \sum_{k=1}^{3} \int_{\Gamma_{l\sigma(k)}} \widehat{u}_{l\sigma(k)} (\pi_\mathfrak{h} \cdot \mathbf{n}_{l\sigma(k)}) \, ds, \tag{1.69}$$

$$\int_{\mathcal{C}_l} \mathbf{q}_\mathfrak{h} \cdot \nabla \phi_\mathfrak{h} \, dx - \int_{\mathcal{C}_l} f \phi_\mathfrak{h} = \sum_{k=1}^{3} \int_{\Gamma_{l\sigma(k)}} \phi_\mathfrak{h} (\widehat{\mathbf{q}}_{l\sigma(k)} \cdot \mathbf{n}_{l\sigma(k)}) \, ds, \tag{1.70}$$

where $\widehat{u}_{l\sigma(k)}$ and $\widehat{\mathbf{q}}_{l\sigma(k)}$ are consistent and conservative approximations of the traces of u and $\mathbf{q}$ along the elemental boundary $\Gamma_{l\sigma(k)}$. The construction of these numerical fluxes in terms of $u_\mathfrak{h}$ and $\mathbf{q}_\mathfrak{h}$ defines the DG discretization. Let us introduce the following trace operators, namely the interface averaging and jump difference across the interface $\Gamma_{l\sigma}$ for the scalar quantity $u_\mathfrak{h}$:

$$\{\{u_\mathfrak{h}\}\} = \frac{u_\mathfrak{h}^- + u_\mathfrak{h}^+}{2}, \qquad [\![u_\mathfrak{h}]\!] = (u_\mathfrak{h}^- - u_\mathfrak{h}^+)\mathbf{n}_l,$$

and for the vector quantity $\mathbf{q}_\mathfrak{h}$:

$$\{\{\mathbf{q}_\mathfrak{h}\}\} = \frac{\mathbf{q}_\mathfrak{h}^- + \mathbf{q}_\mathfrak{h}^+}{2}, \qquad [\![\mathbf{q}_\mathfrak{h}]\!] = (\mathbf{q}_\mathfrak{h}^- - \mathbf{q}_\mathfrak{h}^+) \cdot \mathbf{n}_l,$$

where $u_\mathfrak{h}^-$ and $u_\mathfrak{h}^+$ are respectively the traces of $u_{\mathfrak{h}|\mathcal{C}_l}$ and $u_{\mathfrak{h}|\mathcal{C}_\sigma}$ on $\Gamma_{l\sigma}$ (and similar notations for $\mathbf{q}_\mathfrak{h}^\pm$). Among the available methods, let us state for instance, for a given element $\mathcal{C}_l$:

- ♯ *Bassi-Rebay* (BR):
$$\widehat{u} = \{\{u_\mathfrak{h}\}\}, \qquad \widehat{\mathbf{q}} = \{\{\mathbf{q}_\mathfrak{h}\}\}, \tag{1.71}$$
- ♯ Interior Penalty (IP):
$$\widehat{u} = \{\{u_\mathfrak{h}\}\}, \qquad \widehat{\mathbf{q}} = \{\{\nabla u_\mathfrak{h}\}\} - \nu \mathfrak{h}_l^{-1} [\![u_\mathfrak{h}]\!], \tag{1.72}$$
- ♯ Local DG (LDG):
$$\widehat{u} = \{\{u_\mathfrak{h}\}\} - \boldsymbol{\beta} \cdot [\![u_\mathfrak{h}]\!], \qquad \widehat{\mathbf{q}} = \{\{\mathbf{q}_\mathfrak{h}\}\} + \boldsymbol{\beta} \cdot [\![\mathbf{q}_\mathfrak{h}]\!] - \nu \mathfrak{h}_l^{-1} [\![u_\mathfrak{h}]\!], \tag{1.73}$$
- ♯ Non-symmetric IP (NIPG-j):
$$\widehat{u} = \{\{u_\mathfrak{h}\}\} + \mathbf{n}_l \cdot [\![u_\mathfrak{h}]\!], \qquad \widehat{\mathbf{q}} = \{\{\nabla u_\mathfrak{h}\}\} - \nu \mathfrak{h}_l^{-j} [\![u_\mathfrak{h}]\!], \ j = 1 \text{ or } j = 3. \tag{1.74}$$

where $\boldsymbol{\beta} \in (L^2(\partial\mathcal{C}))^2$ is a vector-valued function which is constant on each edge. The term $\nu \mathfrak{h}_l^{-1} [\![u_\mathfrak{h}]\!]$ is usually called a *penalty* term, which is applied on interface jumps to achieve discrete coercivity of the bilinear form associated with the *primal* formulation of the method, see for instance [37]. In this penalty term the parameter $\nu \geq 0$ denotes the stabilization parameter, which indicates how "strongly" the element are coupled, and $\mathfrak{h}_l$ stands for a *local* (elemental) mesh length scale.

These various formulae lead to as many DGMs, which differ in a variety of factors such as the stencil width, stability, eigenspectrum or even $\mathfrak{h}$-convergence and p-convergence properties. In [A19] and §3.2.2, we mainly focus on the BR and LGL fluxes. While the BR formulae lead to centered fluxes for both u and $\mathbf{q}$, the idea of the LDG approach is to introduce some *upwinding* in the discretization, even if this does not appear as natural for elliptic operators. Additionally, (1.73) judiciously alternate the upwind direction, allowing to reach stability, see [32]. The BR flux is known to degenerate to suboptimal convergence rates for odd polynomial orders. Furthermore, it gives rise to a 10 elements stencil. The LDG flux, with $\boldsymbol{\beta} = \mathbf{n}_l/2$, generally has optimal convergence and a smaller stencil. On the other hand, the BR flux allows for the largest time step and is the conceptually easiest. Note that the LDG flux can be seen as a generalization of the BR flux, which is obtained by setting $\boldsymbol{\beta} = 0$ as well as $\mathbf{u} = 0$. If we only set $\boldsymbol{\beta} = 0$, we obtain a stabilized *Bassi-Rebay* (sBR) flux.

The DG discretization of higher order derivatives (greater than 2) and more particularly of dispersive wave equations has be studied in [108, 107, 69, 105] for the scalar linear and nonlinear cases.

2

Saint-Venant equations

In this first part, we focus on the NSW equations, possibly supplemented by source terms, accounting for the bottom variations and friction. For the sake of simplicity, we begin with 1d partitions with a constant discretization step Δx. We mainly focus on the important issues of stability and well-balancing. The study of stability is investigated through the preservation of the set of invariant domains, that will be referred to as *robustness* property in the following: we aim at keeping the water height positive, eventually under some additional restrictions on the time step Δt. The issue of discrete entropy inequalities is left out.

In §2.1, we introduce a relaxation interpretation of the well-known VFRoe-ncv scheme for the homogeneous NSW equations in 1d, together with a new class of related relaxation schemes. This relaxation interpretation allows to place the VFRoe-ncv scheme in the class of approximate *Riemann* solver in the sense of §1.2.3. This enables to use the argument of non-negativity of *Riemann* solution intermediate states as a sufficient condition to ensure the robustness of the scheme, which was not yet established.

In §2.2, we consider the NSW equations supplemented by a quadratic friction source term of *Chezy* type, and we investigate the construction of a new approximate *Riemann* solver that directly accounts for this source term. This allow us, again, to establish some robustness results under a regular CFL conditions and overcome the difficulties associated with the occurrence of dry areas when such friction terms are added. The numerical preservation of the asymptotic parabolic regime associated with the model is also investigated.

Lastly, in §2.3, we focus on the discretization of the topography source terms and study some variations around the well-known *hydrostatic reconstruction* for the preservation of the motionless steady states, based on an alternative formulation of the NSW equations. Again, some general robustness results are established, including for higher-order schemes on unstructured meshes. A robust DG extension of the method is also detailed, allowing to reach an arbitrary order of accuracy.

2.1 A relaxation interpretation of the VFRoe-ncv scheme

The VFRoe and VFRoe-ncv schemes are approximate *Godunov* type schemes first introduced in [76, 21] and [46] for the NSW equations, applicable to conservative systems of the general form (1.50), which implementation is very easy. Indeed, they

rely on the resolution of the *Riemann* problems by linearization in the spirit of the method of *Roe*. However, the numerical flux is defined directly as the value of the physical flux at the approximate interface value. The VFRoe-ncv scheme has been extensively used during my Ph.D, see for instance [A2] and the communications [C4, C5].
Let us recall the main ideas of the VFRoe-ncv schemes. First, we adopt an admissible change of variable, $\mathbf{u} = \mathcal{U}(\mathbf{w})$ (hence the suffix "*ncv*"). With some abuse in the notations, we set $\mathcal{W}(\mathbf{u}) = \mathbf{w}$ the inverse function of $\mathcal{U}$. Considering smooth enough solutions, the system (1.50) writes as follows:

$$\partial_t \mathbf{u} + B(\mathbf{u})\partial_x \mathbf{u} = 0, \tag{2.1}$$

where $B(\mathbf{u}) = (\nabla_{\mathbf{u}}\mathcal{W}(\mathbf{u}))^{-1} A(\mathcal{W}(\mathbf{u}))\nabla_{\mathbf{u}}\mathcal{W}(\mathbf{u})$ and $A(\mathbf{w})$ is the Jacobian matrix associated with the flux function $\mathcal{F}$. Next, the following linearized *Riemann* problem is considered:

$$\partial_t \mathbf{u} + B(\tilde{\mathbf{u}})\partial_x \mathbf{u} = 0, \tag{2.2}$$

$$\mathbf{u}(x,0) = \begin{cases} \mathbf{u}_L = \mathcal{U}(\mathbf{w}_L) & \text{if} \quad x < 0, \\ \mathbf{u}_R = \mathcal{U}(\mathbf{w}_R) & \text{if} \quad x > 0, \end{cases} \tag{2.3}$$

where $\tilde{\mathbf{u}} := \tilde{\mathbf{u}}(\mathbf{u}_L, \mathbf{u}_R)$ represents any averaging of the variable $\mathbf{u}_L$ and $\mathbf{u}_R$. Since the above problem is linear, the exact solution $\mathbf{u}^*(x/t; \mathbf{u}_L, \mathbf{u}_R)$ is easily obtained. A numerical flux function is then defined as follows:

$$F(\mathbf{w}_L, \mathbf{w}_R) = \mathcal{F}\Big(\mathcal{W}(\mathbf{u}^*(0; \mathcal{U}(\mathbf{w}_L), \mathcal{U}(\mathbf{w}_R)))\Big) \tag{2.4}$$

Hence the explicit form of the VFRoe-ncv scheme is given by:

$$\mathbf{w}_i^{n+1} = \mathbf{w}_i^n - \frac{\Delta t}{\Delta x}\Big(\mathcal{F}\Big(\mathcal{W}(\mathbf{u}^*(0; \mathcal{U}(\mathbf{w}_i^n), \mathcal{U}(\mathbf{w}_{i+1}^n)))\Big) - \mathcal{F}\Big(\mathcal{W}(\mathbf{u}^*(0; \mathcal{U}(\mathbf{w}_{i-1}^n), \mathcal{U}(\mathbf{w}_i^n)))\Big)\Big). \tag{2.5}$$

The numerical flux function is obviously consistent with the exact flux function, and the definition (2.4) ensures conservation for any linearization matrix, in contrast with the classical *Roe* solver.

In [46], motivated by the form of the *Riemann* invariants, the change of variable $\mathcal{U}(\mathbf{w}) = {}^t(2c, v)$ is proposed, with $c = \sqrt{gh}$, and the averaged state $\tilde{\mathbf{u}} = (\mathbf{u}_L + \mathbf{u}_R)/2$ is used for the linearization involved in $B(\tilde{\mathbf{u}})$ and thus for the characterization of $\tilde{v}$ and $\tilde{c}$. The exact solution of the linearized *Riemann* problem (2.2) is

$$\mathcal{W}(\mathbf{u}^*(0; \mathcal{U}(\mathbf{w}_i^n), \mathcal{U}(\mathbf{w}_{i+1}^n))) = \begin{cases} \mathbf{w}_i^n & \text{if} \quad (\lambda^-)_{i+\frac{1}{2}} > 0, \\ \mathcal{W}(\mathbf{u}_{i+\frac{1}{2}}) & \text{if} \quad (\lambda^-)_{i+\frac{1}{2}} < 0 < (\lambda^+)_{i+\frac{1}{2}}, \\ \mathbf{w}_{i+1}^n & \text{if} \quad (\lambda^+)_{i+\frac{1}{2}} < 0, \end{cases} \tag{2.6}$$

where $(\lambda^\pm)_{i+\frac{1}{2}} = \tilde{v}_{i+\frac{1}{2}} \pm \tilde{c}_{i+\frac{1}{2}}$ and $\mathbf{u}_{i+\frac{1}{2}} = {}^t(2c_{i+\frac{1}{2}}, v_{i+\frac{1}{2}})$ with

$$c_{i+\frac{1}{2}} = \frac{1}{2}(c_i + c_{i+1}) - \frac{1}{4}(v_{i+1} - v_i), \qquad v_{i+\frac{1}{2}} = \frac{1}{2}(v_i + v_{i+1}) - (c_{i+1} - c_i). \tag{2.7}$$

To conclude this brief recall, we emphasize that as introduced, the VFRoe method was not recognized to belong to the class of the approximate *Riemann* solvers §1.2.3. Besides, the lack of robustness property was mistakenly admitted as a shortcoming,

even if some efforts were made to secure the playground, see [53]. We show in the following that this issue should have been imputed rather to an inappropriate choice of the linearization states. Note that recently, a relaxation interpretation of the VFRoe method (not VFRoe-ncv) has been proposed in [83].

2.1.1 A new relaxation model

In [A3], a new relaxation method is proposed for the numerical approximation of the weak solutions of system (1.46). We establish the robustness of the VFRoe-ncv scheme with non-conservative variables by identifying it with the relaxation solver. Indeed, the relaxation solver allows some flexibility in the specification of the wave structure through the choice of the relaxation parameters. This freedom is then exploited to robustly handle the occurrence of dry areas.

More precisely, the approach proposed in [A3] differs from the classical relaxation scheme developed by [55] and studied in [79] in the scalar case, since we do not relax every nonlinearity. Indeed, we suggest a choice of relaxation variables that is motivated by the form of the *Riemann* invariants. We formulate a relaxation system with linear degeneracy in all the characteristic fields (which makes the *Riemann* problem easily solvable), a property obtained by a decoupling of the linear equations governing the relaxation variables from the remaining nonlinear equations of the relaxation model. These new variables are governed by linear equations with coefficients that determine the eigenstructure of the relaxation model. Then, we show that the associated relaxation solver is formally equivalent to the VFRoe-ncv solver with nonconservative variables recalled above. This relaxation interpretation allows to place the VFRoe-ncv scheme in the class of approximate *Riemann* solver in the sense of §1.2.3 and to use the argument of non-negativity of the intermediate states as a sufficient condition to reach robustness.

We suggest to approximate the celerity c and velocity v by the new variables Σ and V, intended to be relaxed to c and v at equilibrium. These two new variables are governed by the following equations

$$\begin{cases} \partial_t \Sigma + \bar{v}\partial_x \Sigma + \dfrac{\bar{c}}{2}\partial_x V = \mu(c - \Sigma), \\ \partial_t V + 2\bar{c}\partial_x \Sigma + \bar{v}\partial_x V = \mu(v - V), \end{cases}$$

where $\bar{c}$ and $\bar{v}$ are relaxation parameters to be defined and μ is a parameter intended to tend to infinity. The following first order system with singular perturbations

$$\begin{cases} \partial_t h + \partial_x(\dfrac{\Sigma^2}{g}V) = 0, \quad t > 0,\ x \in \mathbb{R}, \\ \partial_t(hv) + \partial_x\Big(\dfrac{\Sigma^2}{2g}(2V^2 + \Sigma^2)\Big) = 0, \\ \partial_t \Sigma + \bar{v}\partial_x \Sigma + \dfrac{\bar{c}}{2}\partial_x V = \mu(c - \Sigma), \\ \partial_t V + 2\bar{c}\partial_x \Sigma + \bar{v}\partial_x V = \mu(v - V), \end{cases} \tag{2.8}$$

is considered to approximate the weak solutions of (1.50). Indeed, in the formal limit of μ to infinity, the relaxation system (2.8) aims to restore the initial equilibrium

system (1.50). This limit will be referred to as the *equilibrium limit*, defined by $\Sigma = c$ and $V = v$. The conservations of the water height and momentum of the relaxation model (2.8) give those of the equilibrium system (1.50). For the sake of simplicity, we use the general abstract form (1.60) for the relaxation system (2.8):

$$\partial_t \mathbf{U} + \partial_x \mathcal{H}(\mathbf{U}) = \mu \mathcal{R}(\mathbf{U}), \tag{2.9}$$

where we have set $\mathbf{U} = {}^t(h, hv, \Sigma, V)$ defined over the following convex set:

$$\Theta_r = \left\{ \mathbf{U} \in \mathbb{R}^4;\ h \geq 0 \right\}.$$

2.1.2 The *Riemann* problem

Let us consider some relaxation parameters $\bar{c} > 0$ and $\bar{v} \in \mathbb{R}$ such that $\bar{v} \pm \bar{c} \neq 0$ and assume $\mu = 0$, so that the first order system $(2.8)_{\mu=0}$ is hyperbolic for all $\mathbf{U} \in \Theta_r$. Indeed, it admits $\lambda_1^0 = \lambda_2^0 = 0$ and $\lambda^{\pm} = \bar{v} \pm \bar{c}$ as eigenvalues and the associated fields are linearly degenerated.

Let $\mathbf{U}_L$ and $\mathbf{U}_R$ be constant states in Θ_r and define

$$\mathbf{U}_0(x) = \begin{cases} \mathbf{U}_L & \text{if } x < 0, \\ \mathbf{U}_R & \text{if } x > 0, \end{cases} \tag{2.10}$$

as the initial data of the *Riemann* problem for the system $(2.8)_{\mu=0}$. Let us set

$$\Sigma^* = \frac{\Sigma_L + \Sigma_R}{2} - \frac{1}{4}(V_R - V_L), \tag{2.11}$$

$$V^* = \frac{V_L + V_R}{2} - (\Sigma_R - \Sigma_L). \tag{2.12}$$

and define the function $\mathcal{I} : \mathbb{R}^2 \to \mathbb{R}$ $\mathcal{I}(\Sigma, V) = \frac{\Sigma^2}{2g}(\Sigma^2 + 2V^2)$. The first interest of (2.8) is that, as stated above, considering the linear degeneracy property satisfied by all the fields, the *Riemann* problem associated with system (2.8) turns out to be easy to solve:

Proposition 2. *The weak solution of the system* $(2.8)_{\mu=0}$ *with the initial data (2.10) is given by*

- *If* $\lambda^- < 0 < \lambda^+$*:*

$$\mathbf{U}(x,t) = \begin{cases} \mathbf{U}_L & \text{if} \quad & x/t & < \lambda^- \\ \mathbf{U}_L^* & \text{if} \quad \lambda^- < & x/t & < 0, \\ \mathbf{U}_R^* & \text{if} \quad 0 < & x/t & < \lambda^+, \\ \mathbf{U}_R & \text{if} \quad \lambda^+ < & x/t & \end{cases}$$

where

$$\mathbf{U}_L^* = {}^t(h_L^*, q_L^*, \Sigma^*, V^*), \quad \mathbf{U}_R^* = {}^t(h_R^*, q_R^*, \Sigma^*, V^*),$$

and the values of (h_L^*, q_L^*) *and* (h_R^*, q_R^*) *are given by*

$$h_L^* = h_L + \frac{\Sigma^{*2} V^* - \Sigma_L^2 V_L}{g\lambda^-}, \quad q_L^* = (hv)_L - \frac{\mathcal{I}(\Sigma^*, V^*) - \mathcal{I}(\Sigma_L, V_L)}{\lambda^-}, \tag{2.13}$$

$$h_R^* = h_R + \frac{\Sigma^{*2} V^* - \Sigma_R^2 V_R}{g\lambda^+}, \quad q_R^* = (hv)_R - \frac{\mathcal{I}(\Sigma^*, V^*) - \mathcal{I}(\Sigma_R, V_R)}{\lambda^+}. \tag{2.14}$$

- *If* $0 < \lambda^- < \lambda^+$*:*

$$\mathbf{U}(x,t) = \begin{cases} \mathbf{U}_L & \text{if} \quad & x/t & < 0 \\ \mathbf{U}_L^* & \text{if} \quad 0 < & x/t & < \lambda^-, \\ \mathbf{U}_R^* & \text{if} \quad \lambda^- < & x/t & < \lambda^+, \\ \mathbf{U}_R & \text{if} \quad \lambda^+ < & x/t, \end{cases}$$

where

$$\mathbf{U}_L^* = {}^t(h_L^*,\, q_L^*,\, \Sigma_L,\, U_L), \quad \mathbf{U}_R^* = {}^t(h_R^*,\, q_R^*,\, \Sigma^*,\, U^*),$$

and the values of (h_L^*, q_L^*) *and* (h_R^*, q_R^*) *are given by*

$$h_L^* = h_R^* + \frac{\Sigma_L^2 V_L - \Sigma^{*2} V^*}{g\lambda^-}, \quad q_L^* = q_R^* - \frac{\mathcal{I}(\Sigma_L, V_L) - \mathcal{I}(\Sigma^*, V^*)}{\lambda^-}, \tag{2.15}$$

$$h_R^* = h_R + \frac{\Sigma^{*2} V^* - \Sigma_R^2 V_R}{g\lambda^+}, \quad q_R^* = (hv)_R - \frac{\mathcal{I}(\Sigma^*, V^*) - \mathcal{I}(\Sigma_R, V_R)}{\lambda^+}. \tag{2.16}$$

- *If* $\lambda^- < \lambda^+ < 0$*:*

$$\mathbf{U}(x,t) = \begin{cases} \mathbf{U}_L & \text{if} \quad & x/t & < \lambda^- \\ \mathbf{U}_L^* & \text{if} \quad \lambda^- < & x/t & < \lambda^+, \\ \mathbf{U}_R^* & \text{if} \quad \lambda^+ < & x/t & < 0, \\ \mathbf{U}_R & \text{if} \quad 0 < & x/t, \end{cases}$$

where

$$\mathbf{U}_L^* = {}^t(h_L^*,\, q_L^*,\, \Sigma^*,\, V^*), \quad \mathbf{U}_R^* = {}^t(h_R^*,\, q_R^*,\, \Sigma_R,\, V_R),$$

and the values of (h_L^*, q_L^*) *and* (h_R^*, q_R^*) *are given by*

$$h_L^* = h_L + \frac{\Sigma^{*2} V^* - \Sigma_L^2 V_L}{g\lambda^-}, \quad q_L^* = (hv)_L - \frac{\mathcal{I}(\Sigma^*, V^*) - \mathcal{I}(\Sigma_L, V_L)}{\lambda^-}, \tag{2.17}$$

$$h_R^* = h_L^* + \frac{\Sigma_R^2 V_R - \Sigma^{*2} V^*}{g\lambda^+}, \quad q_R^* = q_L^* - \frac{\mathcal{I}(\Sigma_R, V_R) - \mathcal{I}(\Sigma^*, V^*)}{\lambda^+}. \tag{2.18}$$

This analysis of the *Riemann* problem solutions can be supplemented by a preliminary study of the non-negativity of the water depth. Let us focus our attention on h_L^*. Since Σ^* and V^* do not depend on the eigenvalues $\lambda^\pm$, while $\lambda^\pm$ does not depend on the unknowns but just on the fixed parameters $\bar{v}$ and $\bar{c}$, with $h_L > 0$ and $h_R > 0$ it is clear that h_L^* remains non-negative as soon as $|\lambda^\pm|$ is larger enough. Now, assume $h_L = 0$ and $h_R > 0$, to write with $V_L = 0$:

$$h_L^* = \begin{cases} \frac{1}{\lambda^- g}(\Sigma^*)^2 V^* & \text{if } \lambda^- < 0, \\ h_R + \frac{1}{\lambda^+ g}\left((\Sigma^*)^2 V^* - \Sigma_R^2 V_R\right) - \frac{1}{\lambda^- g}(\Sigma^*)^2 V^* & \text{if } \lambda^- > 0. \end{cases}$$

Two cases must be distinguished. If $V^* < 0$, we have just to set $\lambda^- < 0$. Reversely, if we have $V^* > 0$, the parameters $\bar{v}$ and $\bar{c}$ must be chosen to enforce $0 < \lambda^- < \lambda^+$ large enough to satisfy $h_L^* > 0$. Finally, in the case of $h_L = h_R = 0$, with the convention $V_L = V_R = 0$, we obtain $h_L^* = 0$. Involving the same analysis with h_R^*, we have easily established the following result:

Lemma 1. *Assume* $h_L \geq 0$ *and* $h_R \geq 0$ *with the convention* $V_L = 0$ *if* $h_L = 0$ *and* $V_R = 0$ *if* $h_R = 0$*. Then we can find suitable parameters* $\bar{v}$ *and* $\bar{c}$ *such that the functions* h_L^* *and* h_R^**, defined in Proposition 2, are non-negative.*

2.1.3 The relaxation scheme

To approximate the solution at time $t^{n+1} = t^n + \Delta t$, the usual splitting technique [55] is adopted. In a first step, we solve the relaxation model (2.9) omitting the relaxation source terms which are considered in a second step. We assume that a piecewise constant approximate equilibrium solution $\mathbf{w}^h(x, t^n) \in \Omega$ is known at time t^n, defined by

$$\mathbf{w}^h(x, t^n) = \mathbf{w}_i^n = {}^t\left(h_i^n, (hv)_i^n\right), \quad x \in [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}].$$

1. We evolve in time a relevant approximation of the relaxation model (2.9). To access such an issue, we introduce $\mathbf{U}^h \in \Theta_r$ such that for all $0 < t < \Delta t$, the function $\mathbf{U}^h(x, t^n + t)$ is the weak solution of the Cauchy problem for the relaxation system $(2.9)_{\mu=0}$:

$$\partial_t \mathbf{U} + \partial_x \mathcal{H}(\mathbf{U}) = 0, \tag{2.19}$$

supplemented by the following initial equilibrium data:

$$\begin{aligned} \mathbf{U}^h(x, t^n) &= \mathbf{U}_i^n \\ &= {}^t\left(h_i^n, (hv)_i^n, \Sigma_i^n, V_i^n\right), \quad x \in [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}], \end{aligned}$$

where the equilibrium state is defined by $\Sigma_i^n = \sqrt{g\, h_i^n}$ and $V_i^n = (hv)_i^n / h_i^n$. Under the CFL like condition

$$\frac{\Delta t}{\Delta x} \max_i \left(|\lambda^-_{i+\frac{1}{2}}|, |\lambda^+_{i+\frac{1}{2}}| \right) \leq \frac{1}{2}, \tag{2.20}$$

the solution $\mathbf{U}^h$ at time $t^n + \Delta t$ is made of the juxtaposition of the non-interacting *Riemann* problem solution set at the element interfaces $x_{i+\frac{1}{2}}$. Next, the projection of this solution on the piecewise constant functions reads:

$$\mathbf{U}_i^{n+1,-} = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \mathbf{U}^h(x, t^n + \Delta t) dx.$$

Note that according to Lemma 1, we are committed to use a local definition of the parameter $\bar{v}$ and $\bar{c}$ at each interface $x_{i+\frac{1}{2}}$ to ensure the robustness. At each interface $x_{i+\frac{1}{2}}$, we set $\mathbf{U}_L = \mathbf{U}_i^n$ and $\mathbf{U}_R = \mathbf{U}_{i+1}^n$ to define the parameters $\bar{v}_{i+\frac{1}{2}} := \bar{v}(\mathbf{w}_i^n, \mathbf{w}_{i+1}^n)$ and $\bar{c}_{i+\frac{1}{2}} := \bar{c}(\mathbf{w}_i^n, \mathbf{w}_{i+1}^n)$ according to the non-negativity condition. Assuming the CFL restriction (2.20), the relaxation parameters may therefore vary from one interface to another.

2. the second step of the scheme is devoted to the relaxation procedure.We solve the system

$$\partial_t \mathbf{U} = \mu \mathcal{R}(\mathbf{U}),$$

with the piecewise constant approximation $\mathbf{U}_i^{n+1,-}$ as initial data, while μ tends to infinity.

Practically, at time $t = t^n + \Delta t$, this leads to define the updated approximate equilibrium solution $\mathbf{w}^{n+1}(x)$ as follows:

$$\mathbf{w}^{n+1}(x) = \left(h_i^{n+1,-}, (hv)_i^{n+1,-}\right)^T, \quad x \in [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}], \tag{2.21}$$

and to set $\Sigma_i^{n+1} = \sqrt{g h_i^{n+1}}$ and $V_i^{n+1} = (hv)_i^{n+1} / h_i^{n+1}$.

To conclude, once all the algebraic computations done, we observe in [A3] that the numerical relaxation flux function exactly coincides with the VFRoe-ncv flux given by (2.5)-(2.6). Indeed, the relaxation scheme summarizes as follows:

$$\mathbf{w}_i^{n+1} = \mathbf{w}_i^n - \frac{\Delta t}{\Delta x}\left(F_{i+\frac{1}{2}}^n - F_{i-\frac{1}{2}}^n\right), \tag{2.22}$$

where the numerical flux function is defined by $F_{i+\frac{1}{2}}^n = F(\mathbf{w}_i^n, \mathbf{w}_{i+1}^n)$ and

$$F(\mathbf{w}_L, \mathbf{w}_R) = \begin{pmatrix} (c^*)^2 v^*/g \\ (c^*)^2(2(v^*)^2 + (c^*)^2)/(2g) \end{pmatrix}, \tag{2.23}$$

where

$$c^* = \begin{cases} c_L & \text{if } \lambda^- > 0, \\ \dfrac{c_L + c_R}{2} - \dfrac{1}{4}(v_R - v_L) & \text{if } \lambda^- < 0 < \lambda^+, \\ c_R & \text{if } \lambda^+ < 0, \end{cases}$$

$$v^* = \begin{cases} v_L & \text{if } \lambda^- > 0, \\ \dfrac{v_L + v_R}{2} - (c_R - c_L) & \text{if } \lambda^- < 0 < \lambda^+, \\ v_R & \text{if } \lambda^+ < 0. \end{cases}$$

Actually, the two schemes may differ in the evaluation of the eigenvalues $\lambda^\pm$. In this sense, the VFRoe scheme (2.5)-(2.6) is closed when enforcing the linearization $\tilde{v} = \bar{v}$ and $\tilde{c} = \bar{c}$. We can obtain the following expected robustness result:

Theorem 1. *Assume that $\mathbf{w}_i^n \in \Theta$ for all i and assume that the eigenvalues $\lambda_{i+\frac{1}{2}}^\pm$ are evaluated according to the depth non-negativity Lemma 1. Under the CFL condition (2.20), the relaxation scheme (2.22)-(2.23), or equivalently the VFRoe scheme (2.5)-(2.6), preserves the non-negativity of h: $h_i^{n+1} \geq 0$ for all i.*

2.1.4 A class of relaxation schemes

Following the same strategy, many other relaxation models, involving easy linearization, can be considered, with associated schemes which are easy to implement, like the VFRoe-ncv method. For instance, let us consider the following relaxation family of models:

$$\begin{cases} \partial_t h + \partial_x(\dfrac{\Sigma^2}{g}V) = 0, \quad t > 0,\ x \in \mathbb{R}, \\ \partial_t(hv) + \partial_x\Big(\dfrac{\Sigma^2}{2g}(2V^2 + \Sigma^2)\Big) = 0, \\ \partial_t \Sigma + \dfrac{\lambda^+ - \lambda^-}{\alpha + \beta}\partial_x V + \dfrac{\alpha\lambda^- + \beta\lambda^+}{\alpha + \beta}\partial_x \Sigma = \mu(c - \Sigma), \\ \partial_t V + \dfrac{\alpha\lambda^+ + \beta\lambda^-}{\alpha + \beta}\partial_x V - \dfrac{\alpha\beta}{\alpha + \beta}(\lambda^+ - \lambda^-)\partial_x \Sigma = \mu(v - V), \end{cases} \tag{2.24}$$

where α and β are positive parameters to be fixed. This model is nothing but an extension of (2.8). Indeed, (2.24) coincides with (2.8) as soon as $\alpha = \beta = 2$. Considering physical applications, a suitable choice of the parameters α and β should give more accurate simulations. We skip the algebra analysis of this system and

after some computations, the resulting scheme reads in the form (2.22) where the numerical flux function is given by (2.23). Only the definition of (v^*, c^*) has changed and now are given by:

$$c^* = \begin{cases} c_L & \text{if } \lambda^- > 0, \\ \dfrac{\alpha c_L + \beta c_R}{\alpha + \beta} - \dfrac{1}{\alpha + \beta}(v_R - v_L) & \text{if } \lambda^- < 0 < \lambda^+, \\ c_R & \text{if } \lambda^+ < 0, \end{cases}$$

$$v^* = \begin{cases} v_L & \text{if } \lambda^- > 0, \\ \dfrac{\alpha v_L + \beta v_R}{\alpha + \beta} - \dfrac{\alpha\beta}{\alpha + \beta}(c_R - c_L) & \text{if } \lambda^- < 0 < \lambda^+, \\ v_R & \text{if } \lambda^+ < 0. \end{cases}$$

Once again, we can establish that the obtained updated depth h_i^{n+1} remains non-negative as soon as the relaxation wave speeds λ^- and λ^+ are judiciously chosen.

In [A3], a 4^{th}-order extension of the previous relaxation scheme is also performed and numerically validated. A robustness result adapted from [15] is obtained, relying on the introduction of a suitable limiter.

2.2 The quadratic friction source term

In [A8], we introduce a new approximate *Riemann* solver for the NSW equations supplemented by a resistive friction source term. Indeed in many practical applications, for instance in hydraulic or coastal engineering, the bottom roughness and irregularities entail some energy losses, usually modeled as empirical formulae and introducing additional source terms to the right hand side of the momentum NSW equations, see [A1] for a formal derivation of NSW with friction terms coming from bottom wall-laws.
Plenty of formulations have been proposed in the specialized literature for the friction source terms, possibly including complex parameterizations. We refer to [36] for an extensive documentation and comparison of common laws and parameterizations.
Accounting for such a diversity, we focus on the following general formulation, which encompasses the famous *Chezy* and *Darcy-Weisbach* laws:

$$\partial_t \mathbf{W} + \nabla \cdot \mathcal{F}(\mathbf{W}) = \mathcal{S}_f(\mathbf{W}), \tag{2.25}$$

and

$$\mathcal{S}_f(\mathbf{W}) = -g\frac{\kappa}{h^\gamma}\|q\|(0, q_1, q_2)^T \tag{2.26}$$

with $\|q\| = \sqrt{q_1^2 + q_2^2}$, κ is a roughness dimensionalized parameter and γ a rational number greater than 1.

The bottom friction term is usually discretized with a centred explicit or semi-implicit schemes, and as the friction term is not *a priori* bounded for small water heights, it is only applied in regions where the water depth is below a calibrated threshold, as in *GeoClaw* [65] for instance. The reason why the centred approach is usually adequate is that in most of the practical applications, when the water height is not vanishing, the leading terms in the momentum conservation equations are the inertial and hydrostatic pressure forces. Moreover, the friction source term may be more affected by intrinsic modeling errors than by discretization methods. However, in situations in which the water depth is very small, the bed roughness is high and the flow is quasi-uniform, the bottom friction may become the leading terms.

2.2.1 The 1d case

In [A8], we focus on the 1d case:

$$\partial_t \mathbf{w} + \partial_x \mathcal{F}(\mathbf{w}) = \mathcal{S}_f(\mathbf{w}), \tag{2.27}$$

with

$$\mathcal{S}_f(\mathbf{w}, f) = -g\frac{\kappa}{h^\gamma}(0, |q|q)^T.$$

The proposed numerical procedure comes from a relevant and simple correction of a given robust approximate *Riemann* solver for the solutions of the homogeneous hyperbolic system of conservation laws. The adopted correction gives a discretization of the source term which preserves the robustness and does not change the CFL condition.

To do so, the idea is to modify the intermediate state $\mathbf{w}^* = (h^\star, (hu)^\star)^T$ of the considered solver by introducing a new intermediate state vector $\widetilde{\mathbf{w}}^\star$ defined as the following convex combination:

$$\widetilde{\mathbf{w}}^\star = \underline{\alpha}\mathbf{w}^* + (I - \underline{\alpha})\widehat{\mathbf{w}}^*,$$

with $\widehat{\mathbf{w}}^*$ a correction accounting for the source term, defined as follows:

$$\widehat{\mathbf{w}}^*(\frac{x}{t}; \mathbf{w}_L, \mathbf{w}_R) = \begin{cases} \mathbf{w}_L + \dfrac{h_L^\gamma}{g\kappa}\mathcal{S}_f(\mathbf{w}_L), & \text{if } \dfrac{x}{t} < 0, \\ \mathbf{w}_R + \dfrac{h_R^\gamma}{g\kappa}\mathcal{S}_f(\mathbf{w}_R), & \text{if } \dfrac{x}{t} > 0, \end{cases} \tag{2.28}$$

with $\underline{\alpha}$ a 2×2 real diagonal matrix, defined as $\underline{\alpha} = \begin{pmatrix} 1 & 0 \\ 0 & \alpha \end{pmatrix}$, and $\alpha \in [0, 1]$ is an interface dependent parameter that will be defined later.

With this correction of the intermediate state aboard, the modified approximate solver is now given by:

$$\tilde{\mathbf{w}}_{\mathcal{R}}(\frac{x}{t}; \mathbf{w}_L, \mathbf{w}_R) = \underline{\alpha}\mathbf{w}_{\mathcal{R}}(\frac{x}{t}; \mathbf{w}_L, \mathbf{w}_R) + (I - \underline{\alpha})\hat{\mathbf{w}}_{\mathcal{R}}(\frac{x}{t}; \mathbf{w}_L, \mathbf{w}_R)$$

with

$$\hat{\mathbf{w}}_{\mathcal{R}}(\frac{x}{t}; \mathbf{w}_L, \mathbf{w}_R) = \begin{cases} \mathbf{w}_L & \text{if } \dfrac{x}{t} < \lambda^-, \\ \mathbf{w}_L + \dfrac{h_L^\gamma}{g\kappa}\mathcal{S}_f(\mathbf{w}_L) & \text{if } \min(0, \lambda^-) < \dfrac{x}{t} < \min(0, \lambda^+), \\ \mathbf{w}_R + \dfrac{h_R^\gamma}{g\kappa}\mathcal{S}_f(\mathbf{w}_R) & \text{if } \max(0, \lambda^-) < \dfrac{x}{t} < \max(0, \lambda^+), \\ \mathbf{w}_R & \text{if } \dfrac{x}{t} > \lambda^+. \end{cases} \tag{2.29}$$

The particular form of this modified solver can be compared with [14], where long-time asymptotic preserving schemes are proposed for a general class of hyperbolic system of conservation laws with source terms that can be written under a "relaxation-like" form:

$$\partial_t \mathbf{w} + \partial_x \mathcal{F}(\mathbf{w}) = \sigma(\mathbf{w})(\mathcal{R}(\mathbf{w}) - \mathbf{w}), \tag{2.30}$$

with $\sigma(\mathbf{w}) \geq 0$. The NSW equations with a friction source term (2.27) directly fit this framework, with

$$\sigma(\mathbf{w}) = g\frac{\kappa}{h^\gamma}, \tag{2.31}$$

and

$$\mathcal{R}(\mathbf{w}) = \mathbf{w} + \frac{h^\gamma}{g\kappa}\mathcal{S}_f(\mathbf{w}). \tag{2.32}$$

However, our main goal in [A8] is to provide a simple and robust scheme for the discretization of the friction source term, without focusing on the restoration of the associated parabolic asymptotic regime. Therefore, the simple choice of $\underline{\alpha}$ proposed

above has the benefit of keeping the original (homogeneous) discretization of the mass equation untouched, while providing a robust and consistent way to deal with the source term occurring in the momentum equation. A possible asymptotic preserving correction is investigated in §2.2.3.

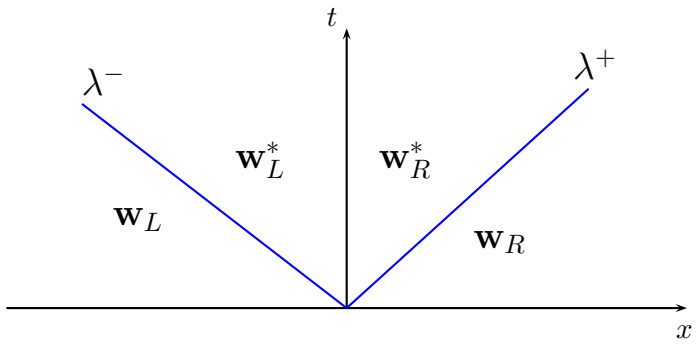

Figure 2.1: Modified approximate *Riemann* solver for the friction scheme with $\mathbf{w}_L^* = \underline{\alpha}\mathbf{w}^* + (I - \underline{\alpha})\mathcal{R}(\mathbf{w}_L)$ and $\mathbf{w}_R^* = \underline{\alpha}\mathbf{w}^* + (I - \underline{\alpha})\mathcal{R}(\mathbf{w}_R)$

Now, using the new solver (2.29), we consider, as usual, the juxtaposition of the *Riemann* problems stated at each interface. Such a juxtaposition is non-interacting as long as Δt satisfies the initial CFL condition (2.20). Then, at time $t^n + t$ for all $t \in (0, \Delta t)$, the juxtaposition of non-interacting approximate solutions reads as follows:

$$\tilde{\mathbf{w}}^{\Delta x}(x, t^n + t) = \tilde{\mathbf{w}}_{\mathcal{R}}\Big(\frac{x - x_{i+\frac{1}{2}}}{t}; \mathbf{w}_i^n, \mathbf{w}_{i+1}^n\Big), \quad \text{if} \quad x \in (x_i, x_{i+1}),$$

and we obtain the updated states are mean values of the solution at time $t + \Delta t$ inside each element by projection:

$$\mathbf{w}_i^{n+1} = \frac{1}{\Delta x}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \tilde{\mathbf{w}}^{\Delta x}(x, t^n + \Delta t)dx. \tag{2.33}$$

Plugging (2.29) into (2.33) we obtain the following scheme:

$$\begin{aligned} h_i^{n+1} &= h_i^n - \frac{\Delta t}{\Delta x}\Big(F^h_{i+\frac{1}{2}} - F^h_{i-\frac{1}{2}}\Big), \qquad (2.34)\\ q_i^{n+1} &= q_i^n - \frac{\Delta t}{\Delta x}\Big(\alpha_{i+\frac{1}{2}}F^q_{i+\frac{1}{2}} - \alpha_{i-\frac{1}{2}}F^q_{i-\frac{1}{2}}\Big) + \Delta t\Big(\frac{1-\alpha_{i-\frac{1}{2}}}{\Delta x}s^+_{i-\frac{1}{2}} + \frac{1-\alpha_{i+\frac{1}{2}}}{\Delta x}s^-_{i+\frac{1}{2}}\Big), \end{aligned}$$

where $F_{i\pm\frac{1}{2}} = (F^h_{i\pm\frac{1}{2}}, F^q_{i\pm\frac{1}{2}})^T$ are defined by (1.58) and

$$\begin{aligned} s^+_{i-\frac{1}{2}} &= \max(0, \lambda^-_{i-\frac{1}{2}})(q^n_{i-1} - q^n_i + |q^n_i|q^n_i) - \max(0, \lambda^+_{i-\frac{1}{2}})|q^n_i|q^n_i + \mathcal{F}^q(\mathbf{w}^n_i), \qquad (2.35)\\ s^-_{i+\frac{1}{2}} &= \min(0, \lambda^-_{i+\frac{1}{2}})|q^n_i|q^n_i + \min(0, \lambda^+_{i+\frac{1}{2}})(q^n_i - q^n_{i+1} - |q^n_i|q^n_i) - \mathcal{F}^q(\mathbf{w}^n_i). \end{aligned}$$

For the particular case $\lambda^-_{i-\frac{1}{2}} < 0 < \lambda^+_{i-\frac{1}{2}}$ and $\lambda^-_{i+\frac{1}{2}} < 0 < \lambda^+_{i+\frac{1}{2}}$, which is the case documented in [A8], the resulting scheme reads:

$$\begin{aligned} h_i^{n+1} &= h_i^n - \frac{\Delta t}{\Delta x}\Big(F^h_{i+\frac{1}{2}} - F^h_{i-\frac{1}{2}}\Big), \\ q_i^{n+1} &= q_i^n - \frac{\Delta t}{\Delta x}\Big(\alpha_{i+\frac{1}{2}} F^q_{i+\frac{1}{2}} - \alpha_{i-\frac{1}{2}} F^q_{i-\frac{1}{2}}\Big) \\ &\quad + \Delta t\Big(\frac{1-\alpha_{i+\frac{1}{2}}}{\Delta x}\lambda^-_{i+\frac{1}{2}} - \frac{1-\alpha_{i-\frac{1}{2}}}{\Delta x}\lambda^+_{i-\frac{1}{2}}\Big)|q_i^n| q_i^n + \frac{\Delta t}{\Delta x}(\alpha_{i+\frac{1}{2}} - \alpha_{i-\frac{1}{2}})\mathcal{F}^q(\mathbf{w}_i^n). \end{aligned} \tag{2.36}$$

To conclude the definition of this numerical scheme, we need to define the parameters $\alpha_{i\pm\frac{1}{2}}$ and we consider consistency conditions in order to do so. In particular, $\alpha_{i+\frac{1}{2}}$ shall fulfill the following requirements:

- ♯ No friction limit: $\alpha_{i+\frac{1}{2}} = 1$ whenever $\kappa = 0$,
- ♯ High friction limit: $\alpha_{i+\frac{1}{2}} \to 0$ in the limit of $\kappa \to \infty$,
- ♯ Source term consistency: $\frac{1-\alpha_{i+\frac{1}{2}}}{\Delta x} = g\frac{\kappa}{h^\gamma} + o(1)$,
- ♯ Dry state conditions: $\alpha_{i+\frac{1}{2}}$ must be defined even if $h = 0$,
- ♯ $\alpha_{i+\frac{1}{2}} \in [0, 1]$.

Among the admissible choices of $\alpha_{i+\frac{1}{2}}$, we choose to consider the following one:

$$\alpha_{i+\frac{1}{2}} = \frac{(h^n_{i+\frac{1}{2}})^\gamma(\lambda^+_{i+\frac{1}{2}} - \lambda^-_{i+\frac{1}{2}})}{(h^n_{i+\frac{1}{2}})^\gamma(\lambda^+_{i+\frac{1}{2}} - \lambda^-_{i+\frac{1}{2}}) + g\kappa\Delta x}, \text{ with } h^n_{i+\frac{1}{2}} = \frac{h_i^n + h_{i+1}^n}{2}. \tag{2.37}$$

The interesting properties of this scheme are summarized as follows:

Theorem 2. *The scheme (2.34)-(2.35)-(2.37) is consistent with (2.27) and assuming that*

1. $\mathbf{w}_i^n \in \Theta$ *for all* i
2. $\mathbf{w}_{\mathcal{R}}(\frac{x - x_{i+\frac{1}{2}}}{\Delta t}; \mathbf{w}_i^n, \mathbf{w}_{i+1}^n) \in \Theta$ *for all* i,

then $\mathbf{w}_i^{n+1} \in \Theta$ *for all* i *under the following CFL condition*

$$\frac{\Delta t}{\Delta x}\max_i\left(|\lambda^-_{i+\frac{1}{2}}|, |\lambda^+_{i+\frac{1}{2}}|\right) \le \frac{1}{2}. \tag{2.38}$$

Additionally, a formally 2^{nd}-order accuracy MUSCL extension is proposed [A8], and we show that the robustness property is extended to the resulting scheme for conservative reconstructions under a classical half CFL-like condition, providing that the MUSCL reconstruction preserves the positivity of the water height.

2.2.2 A 2d extension

In the communication [C12] and then in [A18], a 2d extension on unstructured meshes is proposed. We consider now a vertex-centered 1^{st}-order FV approximation of the solutions of (2.25) on the dual mesh described in §1.2. In order to simplify the analysis, we consider in the following a simplified approximate solver: for a given interface between two elements $\mathcal{C}_i$ and $\mathcal{C}_j$, we define the minimum and maximum wave speeds as follows

$$\lambda_{ij}^{\pm} = \pm\lambda_{ij}, \quad \text{with} \quad \lambda_{ij} = \max(|u_i| + c_i, |u_j| + c_j).$$

We consider the following scheme:

$$\begin{aligned}
h_i^{n+1} &= h_i^n - \frac{\Delta t}{|\mathcal{C}_i|} \sum_{j\in K(i)} \ell_{ij} F^h(\mathbf{W}_i^n, \mathbf{W}_j^n, \mathbf{n}_{ij}), \\
(q_1)_i^{n+1} &= (q_1)_i^n - \frac{\Delta t}{|\mathcal{C}_i|} \sum_{j\in K(i)} \ell_{ij}\alpha_{ij} \Big(F^{q_1}(\mathbf{W}_i^n, \mathbf{W}_j^n, \mathbf{n}_{ij}) - \mathcal{F}^{q_1}(\mathbf{W}_i^n)\cdot \mathbf{n}_{ij} \Big) \\
&\quad - \frac{\Delta t}{|\mathcal{C}_i|} \sum_{j\in K(i)} \ell_{ij}(1-\alpha_{ij})\lambda_{ij} |(q_1)_i^n| (q_1)_i^n, \qquad (2.39) \\
(q_2)_i^{n+1} &= (q_2)_i^n - \frac{\Delta t}{|\mathcal{C}_i|} \sum_{j\in K(i)} \ell_{ij}\alpha_{ij} \Big(F^{q_2}(\mathbf{W}_i^n, \mathbf{W}_j^n, \mathbf{n}_{ij}) - \mathcal{F}^{q_2}(\mathbf{W}_i^n)\cdot \mathbf{n}_{ij} \Big) \\
&\quad - \frac{\Delta t}{|\mathcal{C}_i|} \sum_{j\in K(i)} \ell_{ij}(1-\alpha_{ij})\lambda_{ij} |(q_2)_i^n| (q_2)_i^n,
\end{aligned}$$

where, for a given interface, α_{ij} is defined as

$$\alpha_{ij} = \frac{2h_{ij}^{\gamma}\Lambda(i)\lambda_{ij}}{2h_{ij}^{\gamma}\Lambda(i)\lambda_{ij} + g\kappa\frac{|\mathcal{C}_i|}{\ell_{ij}}}, \quad \text{and} \quad h_{ij} = \frac{h_i + h_j}{2}. \qquad (2.40)$$

The quantities α_{ij} and h_{ij} are of course evaluated at time $t = t^n$. We show in [A18] that the following consistency and robustness properties are satisfied:

Theorem 3. *1. The scheme (2.39)-(2.40) is consistent with (2.27).*

2. The scheme (2.39)-(2.40) preserves the set of admissible states Θ as soon as the underlying approximate Riemann solver does, under the following CFL condition:

$$\Delta t \max_{i,j\in K(i)} \lambda_{ij}\frac{\ell_{ij}}{|T_{ij}|} \leq \frac{1}{2}. \qquad (2.41)$$

The main idea that leads to this robustness results is to show that (2.39) can be

regarded as a convex combination of 1d schemes of the form (2.36):

$$\widetilde{h}_{ij}^{n+1} = h_i^n - \frac{\Delta t}{\delta_{ij}}\Big(F^h(\mathbf{W}_i^n, \mathbf{W}_j^n, \vec{n}_{ij}) - F^h(\mathbf{W}_i^n, \mathbf{W}_i^n, \mathbf{n}_{ij})\Big),$$

$$\begin{aligned}(\widetilde{q_1})_{ij}^{n+1} = (q_1)_i^n &- \frac{\Delta t}{\delta_{ij}}\Big(\alpha_{ij}F^{q_1}(\mathbf{W}_i^n, \mathbf{W}_j^n, \mathbf{n}_{ij}) - \alpha_{ii}F^{q_1}(\mathbf{W}_i^n, \mathbf{W}_i^n, \mathbf{n}_{ij})\Big)\\ &- \frac{\Delta t}{\delta_{ij}}\lambda_{ij}\Big((1-\alpha_{ij}) + (1-\alpha_{ii})\Big)|(q_1)_i^n|(q_1)_i^n + \frac{\Delta t}{\delta_{ij}}(\alpha_{ij}-\alpha_{ii})\mathcal{F}^{q_1}(\mathbf{W}_i)\cdot\mathbf{n}_{ij},\end{aligned}$$

$$\begin{aligned}(\widetilde{q_2})_{ij}^{n+1} = (q_2)_i^n &- \frac{\Delta t}{\delta_{ij}}\Big(\alpha_{ij}F^{q_2}(\mathbf{W}_i^n, \mathbf{W}_j^n, \mathbf{n}_{ij}) - \alpha_{ii}F^{q_2}(\mathbf{W}_i^n, \mathbf{W}_i^n, \mathbf{n}_{ij})\Big)\\ &- \frac{\Delta t}{\delta_{ij}}\lambda_{ij}\Big((1-\alpha_{ij}) + (1-\alpha_{ii})\Big)|(q_2)_i^n|(q_2)_i^n + \frac{\Delta t}{\delta_{ij}}(\alpha_{ij}-\alpha_{ii})\mathcal{F}^{q_2}(\mathbf{W}_i)\cdot\mathbf{n}_{ij},\end{aligned}$$

with $\delta_{ij} = |T_{ij}|/\ell_{ij}$, $\alpha_{ii} = 1$ and α_{ij} defined according to (2.40). These intermediate states are evolved with the 1d scheme (2.36), with a discretization step δ_{ij}. We can then recast (2.39) as follows:

$$\mathbf{W}_i^{n+1} = \sum_{j\in K(i)} \omega_{ij}\tilde{\mathbf{W}}_{ij}^{n+1}, \tag{2.42}$$

with

$$\omega_{ij} = \frac{|T_{ij}|}{|\mathcal{C}_i|} \quad \text{and} \quad \sum_{j\in K(i)} \omega_{ij} = 1.$$

Of course one can easily derive a more general scheme, relying on any robust approximate *Riemann* solver, starting with 1d contributions built from (2.34). Note that in [A18], the 2d extension is introduced in a more general framework as it also accounts for the topography source term in a well-balanced way and it induces a modification of the FV scheme for the mass conservation equation to restore the parabolic asymptotic regime associated with (2.25)-(2.26).

2.2.3 An asymptotic preserving extension

In [A18], we also investigate the *late-time* asymptotic behavior of solutions of the 2d equations (2.25) and suggest a FV discretization that aims at restoring this limit regime on unstructured meshes. Note that this study enters the generalized analysis proposed in [13], that encompasses systems with strongly nonlinear stiff source terms and involves several relaxation time-scales. Such a system, with the correct rescaling, may be written as

$$\epsilon\partial_t\mathbf{W}^\epsilon + \nabla\cdot\mathcal{F}(\mathbf{W}^\epsilon) = \frac{\sigma(\mathbf{W}^\epsilon)}{\epsilon^m}(R(\mathbf{W}^\epsilon) - \mathbf{W}^\epsilon), \tag{2.43}$$

where $m \geq 1$, introducing an additional time scale in the problem. Indeed, it is shown in [13] that the study of the 1d NSW equations (2.27) with a quadratic friction source term, in late-time regime or when the friction is assumed to dominate over the transport (*i.e.* when $t\,\sigma(\mathbf{w}) \to \infty$), relies on the following re-scaling

$$t \leftarrow \frac{t}{\epsilon}, \qquad \sigma(\mathbf{w}) \leftarrow \frac{\sigma(\mathbf{w}^\epsilon)}{\epsilon^2}, \tag{2.44}$$

and, unlike the *Euler* equations with friction term or the *M1* model for radiative transfer, leads to a *nonlinear* diffusive asymptotic regime.

We first derive the singular limit regime of (2.25) in the 2d case, based on the rescaling (2.44). While the topography variations are accounted for in [A18], we choose to consider the flat bottom case in the following, to keep this presentation coherent and simple. The re-scaled equations read

$$\epsilon\partial_t \mathbf{W}^\epsilon + \nabla\cdot\mathcal{F}(\mathbf{W}^\epsilon) = \frac{1}{\epsilon^2}\mathcal{S}_f(\mathbf{W}^\epsilon), \tag{2.45}$$

and in the spirit of *Chapman-Enskog* expansions, see for instance [16], we perform a 1^{st}-order formal expansion of each component of the variable vector:

$$\mathbf{W}^\epsilon = \mathbf{W}^0 + \epsilon\mathbf{W}^1 + O(\epsilon^2)\,. \tag{2.46}$$

Considering that equations (2.45) have to be relevant whenever ϵ tends to zero, matching together the terms of same order in magnitude in ϵ and some algebraic manipulations lead to the following *nonlinear* diffusive regime:

$$\partial_t h - \nabla\cdot\Big(\frac{h^{\frac{\gamma+1}{2}}}{\sqrt{g\kappa\|\nabla h\|}}\nabla h\Big) = 0\,, \tag{2.47}$$

which is the 2d generalization of the equation obtained in [13]. To the author knowledge, the convergence of the solution of (2.45) toward the solution of (2.47) has not been studied yet.

Next, a FV discretization is suggested in [A18], that aims at degenerating accordingly to this diffusive limit regime, providing that a "good" scheme is available for the discretization of the limit equation (2.47). In other words, the limit of the scheme when $t\sigma(\mathbf{W}) \to \infty$ shall be a consistent approximation of the limit diffusion equation. This property is generally not fulfilled by usual numerical schemes and the design of *asymptotic-preserving* schemes has been an important issue during the last decade, see for instance [20, 25, 48].
Achieving such a property on unstructured meshes for a targeted linear parabolic regime is already difficult, since the classical FV schemes for 2^{nd}-order operators are generally not-consistent [45], and do not naturally provide a discrete *maximum principle* (see the Ph.D. of *M.Cathala* on this topic [24]). As far as nonlinear regimes are concerned, the conception of a consistent FV discretization for (2.47) promises to be an interesting and difficult work by itself, not even mentioning the issue of robustness. In [A18], we therefore assume that such a method is available, eventually with complementary assumptions on the mesh.

Reformulating (2.25) as:

$$\partial_t \mathbf{W} + \nabla\cdot\mathcal{F}(\mathbf{W}) = (\sigma(\mathbf{W}) + \overline{\sigma})(\overline{\mathcal{R}}(\mathbf{W}) - \mathbf{W}), \tag{2.48}$$

where

$$\overline{\mathcal{R}}(\mathbf{W}) = \frac{\sigma(\mathbf{W})\mathcal{R}(\mathbf{W}) + \overline{\sigma}\mathbf{W}}{\sigma(\mathbf{W}) + \overline{\sigma}},$$

and $\overline{\sigma}$ is a free parameter, we consider the 2d scheme (2.39) for the computation of the weak solutions of the formulation (2.48) with the adapted definitions

$$\alpha_{ij} = \frac{2\Lambda(i)\lambda_{ij}}{2\Lambda(i)\lambda_{ij} + (\sigma_{ij} + \overline{\sigma}_{ij})\frac{|\mathcal{C}_i|}{\ell_{ij}}}, \qquad \sigma_{ij} = g\frac{2\kappa}{h_i^\gamma + h_j^\gamma}, \tag{2.49}$$

and $\overline{\sigma}_{ij}$ to be defined later. As for (2.40), it is possible to show that (2.39)-(2.49) is consistent with (2.48) and is robust under the CFL (2.41), providing that $\sigma_{ij} + \overline{\sigma}_{ij} \geq 0$.

Focusing now on the asymptotic behavior of (2.39) – (2.49) under the rescaling (2.44), we have

$$\alpha_{ij}^{\epsilon} = \frac{2\epsilon\Lambda(i)\lambda_{ij}}{2\epsilon\Lambda(i)\lambda_{ij} + (\sigma_{ij} + \overline{\sigma}_{ij})\frac{|\mathcal{C}_i|}{\ell_{ij}}}, \tag{2.50}$$

and

$$1 - \alpha_{ij}^{\epsilon} = \frac{(\sigma_{ij} + \overline{\sigma}_{ij})\frac{|\mathcal{C}_i|}{\ell_{ij}}}{2\epsilon\Lambda(i)\lambda_{ij} + (\sigma_{ij} + \overline{\sigma}_{ij})\frac{|\mathcal{C}_i|}{\ell_{ij}}}. \tag{2.51}$$

Injecting into (2.39) and matching the terms of same magnitude in ϵ, we get that in the limit regime, we necessarily recover the equilibrium $\mathcal{R}(\mathbf{W}_i^n) = \mathbf{W}_i^n$, together with the following limit scheme for the water height:

$$h_i^{n+1} = h_i^n - \frac{\Delta t}{|\mathcal{C}_i|} \sum_{j \in K(i)} \ell_{ij} \frac{2\lambda_{ij}\Lambda(i)}{(\sigma_{ij} + \overline{\sigma}_{ij})\frac{|\mathcal{C}_i|}{\ell_{ij}}} \Big(F^h(\mathbf{W}_i^n, \mathbf{W}_j^n, \mathbf{n}_{ij}) - \mathcal{F}^h(\mathbf{W}_i^n)\cdot\mathbf{n}_{ij} \Big)_{\mathcal{R}(\mathbf{W}_i^n)=\mathbf{W}_i^n},$$

which, using for instance a local LF flux (1.59), reduces to

$$h_i^{n+1} = h_i^n + \frac{\Delta t}{|\mathcal{C}_i|} \sum_{j \in K(i)} \ell_{ij} \frac{\lambda_{ij}^2\Lambda(i)}{(\sigma_{ij} + \overline{\sigma}_{ij})\frac{|\mathcal{C}_i|}{\ell_{ij}}} \Big(h_j - h_i \Big). \tag{2.52}$$

Now, let us assume that a consistent FV discretization of the targeted nonlinear diffusive regime (2.47) is available:

$$h_i^{n+1} = h_i^n + \frac{\Delta t}{|\mathcal{C}_i|} \sum_{j \in K(i)} \int_{\Gamma_{ij}} F(h, \nabla h)\nabla h \cdot \mathbf{n}_{ij}\,, \tag{2.53}$$

where $F : (h, \nabla h) \longmapsto \dfrac{h^{\frac{\gamma+1}{2}}}{\sqrt{g\kappa\|\nabla h\|}}$, to obtain :

$$h_i^{n+1} = h_i^n + \frac{\Delta t}{|\mathcal{C}_i|} \sum_{j \in K(i)} \ell_{ij} F_{ij}\, (G_{ij} \cdot \mathbf{n}_{ij})(h_j - h_i)\,, \tag{2.54}$$

where F_{ij} is a discretization of $F(h, \nabla\eta)_{|\Gamma_{ij}}$ and $G_{ij} \cdot \mathbf{n}_{ij}(h_j - h_i)$ an upwind discretization of $\nabla h_{|\Gamma_{ij}}$. We also assume that $G_{ij} \cdot \mathbf{n}_{ij}$ is positive. Then, defining $\overline{\sigma}_{ij}$ such that the following equality holds:

$$\sigma_{ij} + \overline{\sigma}_{ij} = \frac{\lambda_{ij}^2\Lambda(i)\ell_{ij}}{|\mathcal{C}_i| F_{ij}\, (G_{ij} \cdot \mathbf{n}_{ij})}, \tag{2.55}$$

the limit scheme (2.52) for the water height coincide with the targeted FV discretization (2.54) for the limit diffusive regime (2.47).

We recall that in [A18] and [C12], the scheme is introduced in a more general framework, and is both asymptotic preserving and well-balanced. We show on Fig. 2.2-2.3 an example of the modifications induced by the correction, which can have a noticeable impact on the wet/dry interface celerity. Further validations and comparisons can be found in the references above and in the Ph.D. thesis of *A.Duran* [T3].

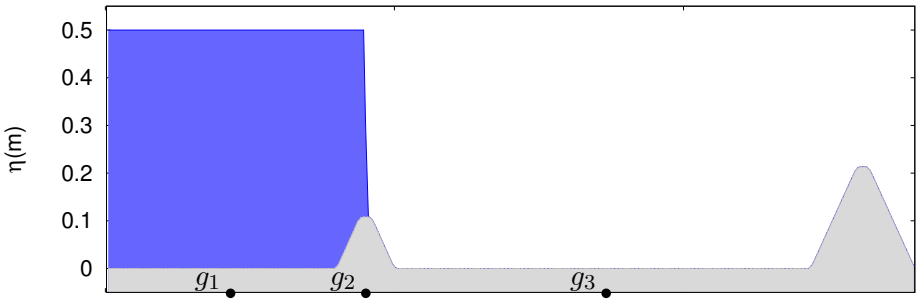

Figure 2.2: Dam-break flow over two frictional humps: basin and initial condition.

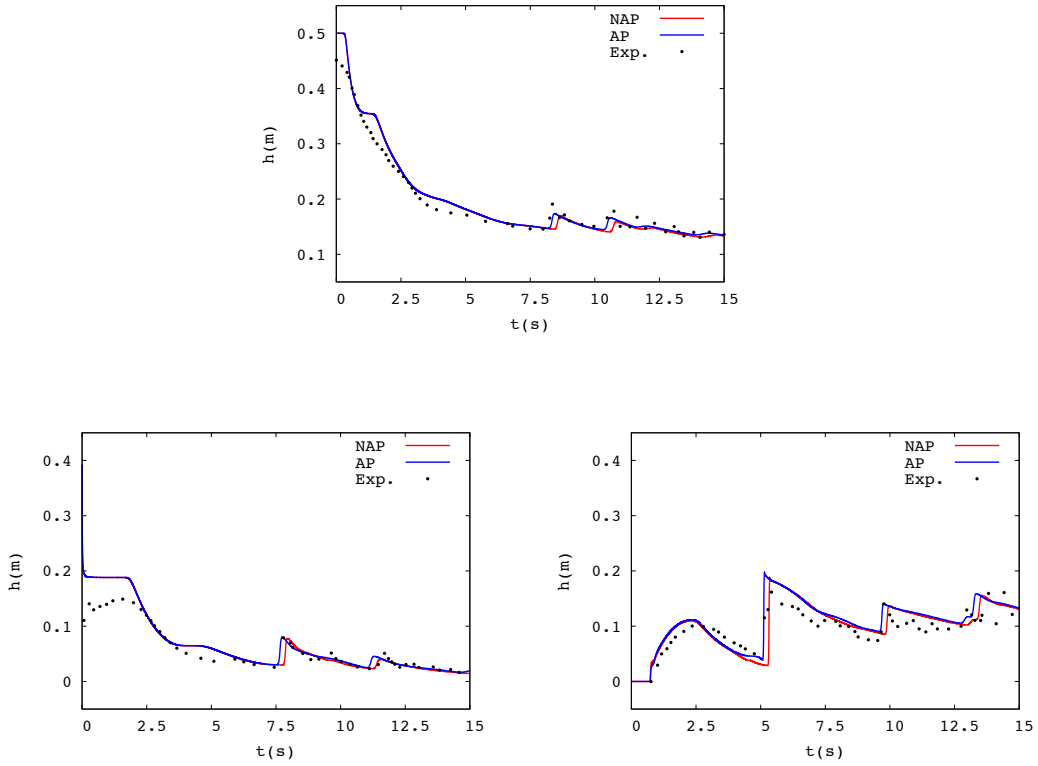

Figure 2.3: Time series of the free surface at gauges g1, g2 and g3. NAP stands for the non asymptotic preserving scheme (2.36) and AP for the modified scheme (2.49)-(2.55)

2.3 The topography source term

In this section, we focus on the discretization of the NSW equations supplemented by the topography source term. The issue of finding suitable discretizations that allow to preserve some important classes of steady states has received a lot of attention in the last 20 years. We say in the following that a numerical FV scheme for (1.43)-(1.44) exactly preserves the motionless steady-states for wet bed if, starting from discrete values $h_i^n + b_i = \eta_e$ and $v_i^n = 0$ the scheme produces updated values at time t^{n+1}, verifying $h_i^{n+1} + b_i = \eta_e$ and $v_i^{n+1} = 0$.

There is a huge of available literature, let us just mention the few works [50, 66, 10, 23] and refer the interested reader to [A15, A16] for a wide list of additional references. One interesting step is certainly the introduction of the so-called *hydrostatic reconstruction* [6], which allows to obtain a robust and well-balanced scheme that preserves motionless steady states, providing that the underlying 1^{st}-order scheme for the homogeneous equations is robust.

In [A6] , we propose an interesting variation on the *hydrostatic reconstruction*, that relies on an alternative form of the NSW equations, namely the *pre-balanced* NSW equations. This approach leads to a very compact formulation, which has became quite popular among the coastal engineering community, see for instance [87, 94]. A 2d extension is proposed in [A15], and the method is recently adapted to the framework of arbitrary order DGMs in [A16].

Concerning the discretization of the topography source term, we also point out the communication [C7], in which we propose a relevant approximate *Riemann* solver interpretation of the VFRoe-ncv scheme for non-homogeneous equations. This interpretation allows method to consider the VFRoe-ncv scheme within the usual Finite Volume formalism. and to derive suitable corrections to enforce the required robustness properties. This work is not detailed in the following.

2.3.1 Pre-balanced hydrostatic reconstruction

Let us again begin with the 1d case, and come back to equations (1.43)-(1.44), written in their dimensionalized 1d form:

$$\left\{ \begin{array}{l} \partial_t \zeta + \partial_x (hv) = 0, \\ \partial_t (hv) + \partial_x \Big(hv^2 \Big) + gh\partial_x \zeta = 0, \end{array} \right. \tag{2.56}$$

and recall that the water height can be written as $h = h_s + \zeta$, where $h_s = h_0 - b$ is the still water depth. In [90], an alternative formulation of the shallow water equations is introduced, with the purpose of naturally balancing flux gradient and topography source term, using the approximate solver of *Roe*. Instead of using the common decomposition

$$gh\partial_x \zeta = g\partial_x (\frac{h^2}{2}) + ghb', \tag{2.57}$$

they suggest to use the alternative splitting and distribution:

$$gh\partial_x \zeta = \frac{1}{2} g\partial_x \Big(\zeta^2 + 2\zeta h_s \Big) - g\zeta h_s', \tag{2.58}$$

giving a deviatoric form of the NSW equations which is properly balanced when solving within a FV *Godunov*-type framework. Additionally, this formulation still retains the hyperbolic nature of the NSW equations, see [89].

However, there is some additional difficulties with the occurrence of dry areas, as the quantities h_s and ζ are not defined. To overcome this, we suggest in [A6] the following redistribution of the $gh\partial_x\zeta$ term:

$$gh\partial_x\zeta = \frac{1}{2}g\partial_x\Big(\eta^2 - 2\eta\, b\Big) + g\eta\, b', \tag{2.59}$$

defined in terms of the *total free surface elevation* above the datum $\eta = h+b$, leading to the following formulation of the NSW equations:

$$\mathbf{u}_t + \partial_x\mathcal{H}(\mathbf{u}, b) = \mathcal{S}_b(\mathbf{u}, b)\,, \tag{2.60}$$

with :

$$\mathbf{u} = \begin{pmatrix}\eta\\ q\end{pmatrix}\ ,\ \mathcal{S}_b(\mathbf{u}, b) = \begin{pmatrix}0\\ -g\eta b'\end{pmatrix}\ ,\quad \mathcal{H}(\mathbf{u}, b) = \begin{pmatrix}q\\ \frac{q^2}{\eta-b} + p(\eta, b)\end{pmatrix}, \tag{2.61}$$

and $p(\eta, b) = \frac{1}{2}g\eta(\eta-2b)$. These equations will be referred to as PB-NSW equations in the following.

This formulation leads to a motionless steady-states preserving scheme, with a centered discretization of the source term:

$$\mathbf{u}_i^{n+1} = \mathbf{u}_i^n - \frac{\Delta t}{\Delta x}(H^n_{i+\frac{1}{2}} - H^n_{i-\frac{1}{2}}) + S^n_{b,i}, \tag{2.62}$$

with

$$S^n_{b,i} = \frac{1}{\Delta x}\begin{pmatrix}0\\ -g\eta_i(b_{i+\frac{1}{2}} - b_{i-\frac{1}{2}})\end{pmatrix}, \tag{2.63}$$

using any consistent numerical flux $H^n_{i+\frac{1}{2}}$. Indeed, it is easy to check that, when a motionless steady state $(\eta_e, 0)$ is reached, the associated discrete fluxes reduces to

$$H_{i+\frac{1}{2}} = p(\eta_e, b_{i+\frac{1}{2}}), \qquad H_{i-\frac{1}{2}} = p(\eta_e, b_{i-\frac{1}{2}}),$$

hence the discrete balance, comparing with (2.63).

This scheme is all the more simple, but with the assumption that the free surface η was constant over the whole domain, we obviously exclude the possibility of the occurrence of dry areas and the previous scheme is unable to preserve motionless steady states involving dry areas. To overcome this drawback, and inspired by the *hydrostatic reconstruction*, we suggest in [A6] the following scheme:

$$\mathbf{u}_i^{n+1} = \mathbf{u}_i^n - \frac{\Delta t}{\Delta x}(H_{i+\frac{1}{2}} - H_{i-\frac{1}{2}}) + \frac{\Delta t}{\Delta x}(S_{b,i} + \widehat{S}_{b,i}), \tag{2.64}$$

with

$$H_{i+\frac{1}{2}} = H(\mathbf{u}^-_{i+\frac{1}{2}}, \mathbf{u}^+_{i+\frac{1}{2}}, b^*_{i+\frac{1}{2}}, b^*_{i+\frac{1}{2}}), \quad H_{i-\frac{1}{2}} = H(\mathbf{u}^-_{i-\frac{1}{2}}, \mathbf{u}^+_{i-\frac{1}{2}}, b^*_{i-\frac{1}{2}}, b^*_{i-\frac{1}{2}}), \tag{2.65}$$

$$\widetilde{S}^n_{b,i} = \begin{pmatrix} 0 \\ -g\eta_i(b^*_{i+\frac{1}{2}} - b^*_{i-\frac{1}{2}}) \end{pmatrix}, \tag{2.66}$$

$$\widehat{S}^n_{b,i} = \frac{g}{2}\begin{pmatrix} 0 \\ \Delta_{i,l}\Big(b^*_{i+\frac{1}{2}} - (b^*_{i-\frac{1}{2}} - \Delta_{i,l})\Big) + \Delta_{i,r}\Big((b^*_{i+\frac{1}{2}} - \Delta_{i,r}) - b^*_{i-\frac{1}{2}}\Big) \end{pmatrix}, \tag{2.67}$$

$$b^*_{i+\frac{1}{2}} = \max(b_i, b_{i+1}), \tag{2.68}$$

$$\Delta_{i,l} = \max(0, b^*_{i-\frac{1}{2}} - \eta_i), \qquad \Delta_{i,r} = \max(0, b^*_{i+\frac{1}{2}} - \eta_i), \tag{2.69}$$

and

$$\begin{aligned} h^-_{i+\frac{1}{2}} &= \max(\eta_i - b^*_{i+\frac{1}{2}}, 0), & h^+_{i+\frac{1}{2}} &= \max(\eta_{i+1} - b^*_{i+\frac{1}{2}}, 0), \\ \eta^-_{i+\frac{1}{2}} &= h^-_{i+\frac{1}{2}} + b^*_{i+\frac{1}{2}}, & \eta^+_{i+\frac{1}{2}} &= h^+_{i+\frac{1}{2}} + b^*_{i+\frac{1}{2}}, \\ \mathbf{u}^-_{i+\frac{1}{2}} &= (\eta^-_{i+\frac{1}{2}}, h^-_{i+\frac{1}{2}} v_i)^T, & \mathbf{u}^+_{i+\frac{1}{2}} &= (\eta^+_{i+\frac{1}{2}}, h^+_{i+\frac{1}{2}} v_{i+1})^T. \end{aligned} \tag{2.70}$$

and $(\mathbf{u}^-, \mathbf{u}^+, b^-, b^+) \mapsto H(\mathbf{u}^-, \mathbf{u}^+, b^-, b^+)$ corresponds for instance to the LF flux (1.59) computed from the variable state $\mathbf{u}$ and the physical flux $\mathcal{H}$. For the sake of simplicity, the scheme (2.64)→(2.70) will be simply referred to as (2.64) in the following.

The new terms $\Delta_{i,l}$ and $\Delta_{i,r}$ are *dry element sensors* that activate the correction $\widehat{S}^n_{b,i}$ only when a dry neighbor is detected. Let us have a closer look, for instance, on the equilibrium depicted on Fig. 2.4:

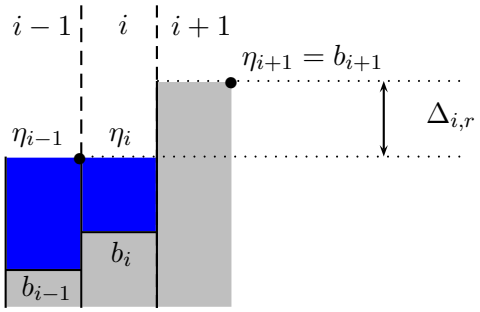

Figure 2.4: Wet/dry interface.

- interface $i+\frac{1}{2}$: $h_{i+1} = 0$, $b^*_{i+\frac{1}{2}} = b_{i+1} = \eta_{i+1}$ and $\Delta_{i,r} = b_{i+1} - \eta_i$. Consequently

$$h^-_{i+\frac{1}{2}} = h^+_{i+\frac{1}{2}} = 0 \quad \text{and} \quad \eta^-_{i+\frac{1}{2}} = \eta^+_{i+\frac{1}{2}} = b_{i+1}.$$

- interface $i-\frac{1}{2}$: $\Delta_{i,l} = 0$, $b^*_{i-\frac{1}{2}} = b_i$ and $\eta^-_{i-\frac{1}{2}} = \eta^+_{i-\frac{1}{2}} = \eta_i = \eta_{i-1} = \eta^e$.

Injecting these values into (2.64), we see that the correction $\widehat{S}_{b,i}$ helps to restore the perfect balance. A study of the other possible configurations leads to the same

conclusion.

Interestingly, an alternative and equivalent formulation to (2.64) reads as follows:

$$\mathbf{u}_i^{n+1} = \mathbf{u}_i^n - \frac{\Delta t}{\Delta x}(H^-_{i+\frac{1}{2}} - H^+_{i-\frac{1}{2}}) + S_{b,i}, \tag{2.71}$$

with

$$H^-_{i+\frac{1}{2}} = H(\mathbf{u}^-_{i,r}, \mathbf{u}^+_{i,r}, b^*_{i,r}, b^*_{i,r}), \qquad H^+_{i-\frac{1}{2}} = H(\mathbf{u}^-_{i,l}, \mathbf{u}^+_{i,l}, b^*_{i,l}, b^*_{i,l}), \tag{2.72}$$

$$S^n_{b,i} = \begin{pmatrix} 0 \\ -g\eta_i(b^*_{i,r} - b^*_{i,l}) \end{pmatrix}, \tag{2.73}$$

and

$$\begin{aligned}
b^*_{i,l} &= b^*_{i-\frac{1}{2}} + \Delta_{i,l}, & b^*_{i,r} &= b^*_{i+\frac{1}{2}} + \Delta_{i,r}, \\
\eta^\pm_{i,r} &= h^\pm_{i+\frac{1}{2}} + b^*_{i,r}, & \eta^\pm_{i,l} &= h^\pm_{i-\frac{1}{2}} + b^*_{i,l}, \\
\mathbf{u}^-_{i,r} &= (\eta^-_{i,r}, h^-_{i+\frac{1}{2}} v_i)^T, & \mathbf{u}^+_{i,r} &= (\eta^+_{i,r}, h^+_{i+\frac{1}{2}} v_{i+1})^T, \\
\mathbf{u}^-_{i,l} &= (\eta^-_{i,l}, h^-_{i-\frac{1}{2}} v_{i-1})^T, & \mathbf{u}^+_{i,l} &= (\eta^+_{i,l}, h^+_{i-\frac{1}{2}} v_i)^T,
\end{aligned} \tag{2.74}$$

while $b^*_{i\pm\frac{1}{2}}$, $h^\pm_{i\pm\frac{1}{2}}$, $\Delta_{i,r}$ and $\Delta_{i,l}$ remain unchanged. In the following, the formulation (2.71)→(2.74) will be simply referred to as (2.71). We have the following result:

Proposition 3. *The scheme (2.71) (or alternatively (2.64)) is consistent with (2.56) and preserves the motionless steady states, even with the possible occurrence of dry areas.*

We clearly see that the introduction of the modified quantities $b^*_{i,r}$ and $b^*_{i,l}$ and the subsequent variables modifications (2.74) do not modify the scheme (2.64) when there is no dry area. When dry elements have to be taken into account in the computation of the numerical fluxes (2.72), then the induced modifications perfectly restore the motionless steady state at the wet/dry interface.
Additionally, we can check that the modifications entailed by the use of the reconstructed values (2.74) are of order $\mathcal{O}(\Delta x)$, as

$$\mathbf{u}^\pm_{i,r} = \mathbf{u}(x_{i+\frac{1}{2}}) + \mathcal{O}(\Delta x), \quad \text{and} \quad \mathbf{u}^\pm_{i,l} = \mathbf{u}(x_{i-\frac{1}{2}}) + \mathcal{O}(\Delta x).$$

2.3.2 Robustness

So far, we have focused on the preservation of the motionless equilibrium. Let now investigate the robustness property. Indeed, within the PB-NSW formulation, we work with the *free surface* as a variable in the mass conservation equation instead of the water height, obviously modifying the numerical viscosity of the associated numerical schemes when the bottom is not flat.
However, the analysis can be easily performed is some cases, with the use of the formulation (2.71) rather than (2.64). More precisely, the robustness result can be easily obtained in the case when the numerical flux formula involves only a difference of free surface element values. Let us be more specific with the example of the global LF flux (1.59):

Lemma 2. *We consider the following first-order scheme for the update of the free surface value:*

$$\eta_i^{n+1} = \eta_i^n - \frac{\Delta t}{\Delta x}(H^{-,\eta}_{i+\frac{1}{2}} - H^{+,\eta}_{i-\frac{1}{2}}), \tag{2.75}$$

where $H^{-,\eta}_{i+\frac{1}{2}}$ *and* $H^{+,\eta}_{i-\frac{1}{2}}$ *are the first component of the well-balanced fluxes (2.72), defined according to the reconstructions (2.74) and LF flux (1.59). If we have* $h_i^n \geq 0\ \forall i$*, then* $h_i^{n+1} \geq 0\ \forall i$ *under the condition* $\frac{\Delta t}{\Delta x}\lambda_o \leq 1$.

Proof. The idea is to write (2.75) as a linear combination of positive values. We have:

$$\eta_i^{n+1} = \eta_i^n - \frac{\Delta t}{2\Delta x}\Big((h^-_{i+\frac{1}{2}} v_i^n + h^+_{i+\frac{1}{2}} v_{i+1}^n) - \lambda_o(\eta^+_{i,r} - \eta^-_{i,r}) \\ - (h^-_{i-\frac{1}{2}} v_{i-1}^n + h^+_{i-\frac{1}{2}} v_i^n) + \lambda_o(\eta^+_{i,l} - \eta^-_{i,l})\Big),$$

Denoting that $\eta^+_{i,r} - \eta^-_{i,r} = h^+_{i+\frac{1}{2}} - h^-_{i+\frac{1}{2}}$ and $\eta^+_{i,l} - \eta^-_{i,l} = h^+_{i-\frac{1}{2}} - h^-_{i-\frac{1}{2}}$, and subtracting b_i at both side of the equality, we can conclude, writing (2.75) as a linear combination of positive values:

$$\eta_i^{n+1} - b_i = \Big(1 - \frac{\Delta t}{2\Delta x}(\lambda_o + v_i^n)\frac{h^-_{i+\frac{1}{2}}}{h_i^n} - \frac{\Delta t}{2\Delta x}(\lambda_o - v_i^n)\frac{h_{i-\frac{1}{2}^+}}{h_i^n}\Big)h_i^n \\ + \Big(\frac{\Delta t}{2\Delta x}(\lambda_o + v_{i-1}^n)\frac{h^-_{i-\frac{1}{2}}}{h_{i-1}^n}\Big)h_{i-1}^n + \Big(\frac{\Delta t}{2\Delta x}(\lambda_o - v_{i+1}^n)\frac{h^+_{i+\frac{1}{2}}}{h_{i+1}^n}\Big)h_{i+1}^n, \tag{2.76}$$

since $h^-_{i+\frac{1}{2}} < h_i^n$ and $h^+_{i-\frac{1}{2}} < h_i^n$ by construction. □

A similar result can be obtained with the *Rusanov* or the *HLL* fluxes, although with more restricted time steps.

2.3.3 Higher order of accuracy

The formulation (2.71) has the advantage of being very compact and is virtually unmodified when 2^{nd}-order interface values are used. For a given element $\mathcal{C}_i$, let us replace the piecewise constant values η_i, h_i and v_i respectively by $\eta_{i,r}$, $h_{i,r}$, $v_{i,r}$ and $\eta_{i,l}$, $h_{i,l}$, $v_{i,l}$, which are some 2^{nd}-order reconstructed values respectively at the right and left interfaces, obtained for instance by a MUSCL reconstruction [64]. To simplify the analysis, we restrict the study to the case of *conservative* reconstructions:

$$\eta_i = (\eta_{i,r} + \eta_{i,l})/2, \quad h_i = (h_{i,r} + h_{i,l})/2, \quad \text{and } v_i = (v_{i,r} + v_{i,l})/2. \tag{2.77}$$

As in [6], we also deduce some 2^{nd}-order reconstructed interface values for the topography, let say $b_{i,r} = \eta_{i,r} - h_{i,r}$ and $b_{i,l} = \eta_{i,l} - h_{i,l}$, to ensure that the free surface at rest won't be disturbed by the high-order reconstruction process.
Adapting now the reconstructions (2.74) and (2.73), see [A6] for details, and injecting these new values into (2.72), we obtain a 2^{nd}-order scheme which preserves the motionless steady states. Additionally, if the underlying 1^{st}-order scheme is robust (in the sense of Lemma 2) under a given CFL condition, we can classically prove that the 2^{nd}-order scheme is robust under a half CFL condition.

The robustness result can be extended to non-conservative or even higher-order reconstructions, with a more restricted CFL condition and eventually an additional limitation, with the use of an intermediate ghost state, in the spirit of [15]. Another road to enforce the robustness with high-order reconstruction is to use the quadrature-based limiter of [111], providing that the reconstructed polynomial is available. This approach is detailed in §3.2

2.3.4 The 2D case

In [A15], we extend the scheme to the 2d case on unstructured meshes. The PB-NSW equations read as follows

$$\mathbf{U}_t + \nabla \cdot \mathcal{H}(\mathbf{U}, b) = \mathcal{S}_b(\mathbf{U}, b)\,, \tag{2.78}$$

with :

$$\mathbf{U} = \begin{pmatrix} \eta \\ hv_1 \\ hv_2 \end{pmatrix} \;,\; \mathcal{S}_b(\mathbf{U}, b) = -g\eta \begin{pmatrix} 0 \\ \partial_x b \\ \partial_y b \end{pmatrix} \;,\quad \mathcal{H}(\mathbf{U}, b) = \begin{pmatrix} q_1 & q_2 \\ \frac{q_1^2}{\eta - b} + p(\eta, b) & \frac{q_1 q_2}{\eta - b} \\ \frac{q_1 q_2}{\eta - b} & \frac{q_2^2}{\eta - b} + p(\eta, b) \end{pmatrix}.$$

For a given element $\mathcal{C}_i$, we suggest the following scheme:

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \frac{\Delta t}{|\mathcal{C}_i|} \sum_{k=1}^{\Lambda(i)} \ell_{ij(k)} H_s(\mathbf{U}_{ij(k)}^-, \mathbf{U}_{ij(k)}^+, b_i, b_j, \mathbf{n}_{ij(k)}), \tag{2.79}$$

with

$$H_s(\mathbf{U}_{ij}^-, \mathbf{U}_{ij}^+, b_i, b_j, \mathbf{n}_{ij}) = H(\mathbf{U}_{ij}^-, \mathbf{U}_{ij}^+, \check{b}_{ij}, \check{b}_{ij}, \mathbf{n}_{ij}) - S_{c,ij}, \tag{2.80}$$

$$S_{c,i} = \sum_{k=1}^{\Lambda(i)} \ell_{ij(k)} S_{c,ij(k)} = \sum_{k=1}^{\Lambda(i)} \ell_{ij(k)} \begin{pmatrix} 0 \\ g\hat{\eta}_{ij(k)}(b_i - \check{b}_{ij(k)})\mathbf{n}_{ij(k)} \end{pmatrix}, \tag{2.81}$$

where $\hat{\eta}_{ij(k)}$ is a consistent approximation of η at the interface $\Gamma_{ij(k)}$ and for a given interface Γ_{ij}, the left and right *Riemann* states are obtained through the following reconstructions:

$$b_{ij}^* = \max(b_i, b_j), \quad \Delta_{ij} = \max(0, b_{ij}^* - \eta_i), \quad \check{b}_{ij} = b_{ij}^* - \Delta_{ij}, \tag{2.82}$$

$$h_{ij}^* = \max(0, \eta_i - b_{ij}^*), \qquad h_{ji}^* = \max(0, \eta_j - b_{ij}^*), \tag{2.83}$$

$$\eta_{ij}^- = h_{ij}^* + \check{b}_{ij}, \qquad \eta_{ij}^+ = h_{ji}^* + \check{b}_{ij}, \tag{2.84}$$

leading to the new edge values:

$$\mathbf{U}_{ij}^- = (\eta_{ij}^-, h_{ij}^* \mathbf{v}_i), \quad \mathbf{U}_{ij}^+ = (\eta_{ij}^+, h_{ji}^* \mathbf{v}_j). \tag{2.85}$$

The simpler choice for $\hat{\eta}_{ij(k)}$ is certainly η_i, leading to the simplified source term:

$$S_{c,i} = -g\eta_i \sum_{k=1}^{\Lambda(i)} \ell_{ij(k)} \begin{pmatrix} 0 \\ \check{b}_{ij(k)} \mathbf{n}_{ij(k)} \end{pmatrix}. \tag{2.86}$$

We have the following properties:

Theorem 4. *The scheme (2.79)→(2.86) preserves the motionless steady states, even with the occurrence of dry states. Additionally, assuming that :*

1. $h_i^n \geq 0$ *and* $h_{j(k)}^n \geq 0,\ 1 \leq k \leq \Lambda(i)$,
2. *the numerical fluxes (2.80) are computed with the LF flux (1.59).*

Then we have $h_i^{n+1} \geq 0$ *under the CFL*

$$\Delta t \lambda_o \frac{\ell_{ij}}{|T_{ij}|} \leq 1, \quad \forall i, \forall j. \tag{2.87}$$

The key argument to achieve the robustness is that again, we can write (2.79) as a convex combination of 1^{st}-order schemes, of the form studied in Lemma2, that reads:

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \frac{\ell_{ij}\Delta t}{|T_{ij}|}\Big(H(\mathbf{U}_{ij(k)}^-, \mathbf{U}_{ij(k)}^+, \check{b}_{ij}, \check{b}_{ij}, \mathbf{n}_{ij(k)}) - \mathcal{H}(\mathbf{U}_i, b_i)\cdot\mathbf{n}_{ij(k)})\Big) + \frac{\ell_{ij}\Delta t}{|T_{ij}|} S_{c,ij(k)}.$$

A more stringent CFL was obtained in [A15], with the choice $\hat{\eta}_{ij(k)} = \frac{\eta_{ij}^- + \eta_{ij}^+}{2}$ and following another path to ensure the robustness.

2.3.5 Accuracy improvement

A 2^{nd}-order scheme is also proposed and studied in [A15]. As in the 1d case, the order of space-accuracy is improved by substituting to the original values $\mathbf{U}_i$ and $\mathbf{U}_j$ in the computation of (2.82)→(2.85) some "better" interpolated/extrapolated vector states $\mathbf{U}_{ij}$ and $\mathbf{U}_{ji}$ at the interfaces Γ_{ij}.
There is several way to gain a better convergence and accuracy on unstructured meshes, from the simple *Green-Gauss* reconstructions to more complex and wider stencils ENO/WENO approaches [1]. During the Ph.D of *A.Duran* [T3], several MUSCL reconstructions have been implemented, compared and validated, including the ones proposed in [7] and the V6 schemes introduced in [22], which seems to provide less diffusive results and better convergence rates on many test cases. This reconstructtion mainly relies on the judicious construction of a weighted combination of several reconstructed gradients, issued from P^1 interpolations, centered difference formulae, and a particular *higher-order* gradient which gives a numerical dissipation built as a sixth-order spatial derivative for the linear advection equation.
We show in [A15] that the resulting scheme preserve the motionless steady states, even with the occurrence of dry elements, and is robust under a suitable time step restriction. The study of the robustness is more involved than in the 1^{st}-order case. Indeed, we can find some robustness results in the literature for 2d schemes on unstructured meshes, like in [84, 85]. However, the stability analysis mainly relies on some conservative assumptions for the reconstructed values, which don't hold in our case. An interesting way of extending this analysis to general reconstructions is introduced in [12] and adapted to our study.
A sub-mesh is introduced, as shown on Fig.2.5, together with new intermediate states that allow to interpret the reconstructed states coming from (2.85) as piecewise constant states on sub-elements. Considering a element $\mathcal{C}_i$, we split the edges $S_i\, M_{ij(k)j(k+1)}$ into two segments, separated by the vertex $m_{ij(k)j(k+1)}$. Joining the

vertices $m_{i\,j(k)\,j(k+1)}$, we split the dual element $\mathcal{C}_i$ into a sub-element $\mathcal{C}_i^*$ and $\Lambda(i)$ sub-elements $\mathcal{C}_{ij(k)}$ obtained with joining the vertexes $M_{ij(k-1)j(k)}$, $M_{ij(k)j(k+1)}$, $m_{ij(k)j(k+1)}$ and $m_{ij(k-1)j(k)}$, leading to:

$$\mathcal{C}_i = \mathcal{C}_i^* \cup \left(\bigcup_{k=1}^{\Lambda(i)} \mathcal{C}_{ij(k)} \right).$$

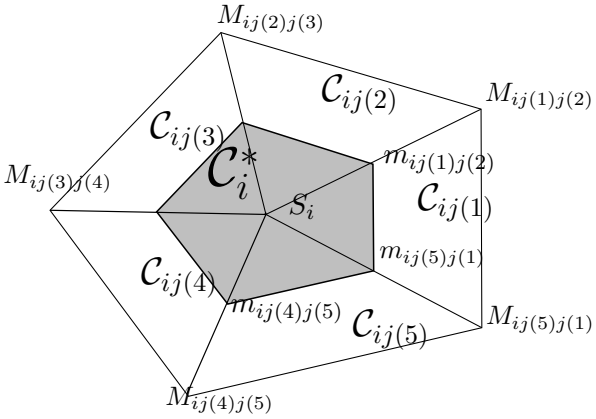

Figure 2.5: Sub-element decomposition of $\mathcal{C}_i$

Let us also denote, for $1 \leq k \leq \Lambda(i)$:

- $\Gamma^*_{ij(k)}$ the segment separating $\mathcal{C}_i^*$ and $\mathcal{C}_{ij(k)}$,
- $\ell^*_{ij(k)}$ the length of $\Gamma^*_{ij(k)}$
- $\mathbf{n}^*_{ij(k)}$ the outer unit normal to $\Gamma^*_{ij(k)}$,
- $\Gamma^p_{ij(k)}$, for $p \in \{1, \ldots, 4\}$, the boundaries of $\mathcal{C}_{ij(k)}$ (so that $\bigcup_{p=1}^{4} \Gamma^p_{ij(k)} = \partial \mathcal{C}_{ij(k)}$),
- $\ell^p_{ij(k)}$ the length of $\Gamma^p_{ij(k)}$,
- $\mathbf{n}^p_{ij(k)}$ the outer unit normal to $\Gamma^p_{ij(k)}$,
- $\mathbf{U}^p_{ij(k)}$, for $p \in \{1, \ldots, 4\}$, the value of $\mathbf{U}$ in the neighboring elements of $\mathcal{C}_{ij(k)}$.

Let us finally assume that all the sub-elements described above are the disjoint union of triangles:

$$\mathcal{C}_{ij(k)} = \bigcup_{p=1}^{4} T^p_{ij(k)} \quad \text{and} \quad \mathcal{C}_i^* = \bigcup_{k=1}^{\Lambda(i)} T^*_{ij(k)}. \tag{2.88}$$

We associate the inner reconstructed vector state $\mathbf{U}^-_{ij(k)}$ to the element $\mathcal{C}_{ij(k)}$ and we define a new intermediate state $\mathbf{U}_i^*$, associated with the sub-element $\mathcal{C}_i^*$, defined as follows:

$$\frac{|\mathcal{C}_i^*|}{|\mathcal{C}_i|}\mathbf{U}_i^* + \sum_{k=1}^{\Lambda(i)} \frac{|\mathcal{C}_{ij(k)}|}{|\mathcal{C}_i|}\mathbf{U}^-_{ij(k)} = \mathbf{U}_i^n. \tag{2.89}$$

We also define a new intermediate state h_i^* in the same way :

$$\frac{|\mathcal{C}_i^*|}{|\mathcal{C}_i|} h_i^* + \sum_{k=1}^{\Lambda(i)} \frac{|\mathcal{C}_{ij(k)}|}{|\mathcal{C}_i|} h_{ij(k)}^* = h_i^n, \tag{2.90}$$

and deduce intermediate topography values b_i^* attached to $\mathcal{C}_i^*$. Then, assuming that h_i^n, h_i^* and all the interface values $h_{ij(k)}$ coming from the high-order reconstruction are positive, and that the LF flux is used (or any consistent and conservative numerical flux satisfying a robustness property similar to Lemma 2, the resulting high-order scheme is shown to be robust under the following CFL conditions:

$$\Delta t \frac{\ell_{ij(k)}^p}{|T_{ij(k)}^p|} \max \left| \lambda^{\pm}(\mathbf{U}_{ij(k)}^-, \mathbf{U}_{ij(k)}^p, \mathbf{n}_{ij(k)}^p) \right| \leq 1, \, \forall \, \mathcal{C}_{ij(k)}, \, T_{ij(k)}^p, \, 1 \leq p \leq 4, \tag{2.91}$$

$$\Delta t \frac{\ell_{ij(k)}^*}{|T_{ij(k)}^*|} \max \left| \lambda^{\pm}(\mathbf{U}_i^*, \mathbf{U}_{ij(k)}^-, \mathbf{n}_{ij(k)}^*) \right| \leq 1, \, \forall \mathcal{C}_i^*, \, T_{ij(k)}^*, \, 1 \leq k \leq \Lambda(i). \tag{2.92}$$

This result is again obtained by interpreting the high-order scheme as a convex combination of 1^{st}-order schemes. We evolve independently in time both intermediate states η_i^* and reconstructed states $\eta_{ij(k)}^-$, considered as constant states in their respective sub-elements, with the robust 1^{st}-order scheme (2.79)→(2.85), and recover the updated averaged value η_i^{n+1} from (2.89).

Obviously, for an arbitrary high-order reconstruction procedure, we have to ensure that the high-order reconstructions h_{ij}, h_{ji} and the intermediate state h_i^* are positive, and an additional limitation procedure may be necessary to enforce this. An example of computation is shown on Fig. 2.6.

2.3.6 An arbitrary order accuracy well-balanced DGM

The development of well-balanced DG schemes for the NSW equations is recent, and there is very few existing works, especially when considering the 2d case on unstructured grids. In the 1d case, we can refer to [103] for a general (but complex) discretization of a class of conservation laws with separable source terms, leading to high-order DGMs with the well-balanced property for motionless equilibrium. The key ingredient is a suitable decomposition of the integral of the source terms into a sum of several terms, each of which is discretized independently in a consistent way with the discretization of the corresponding flux derivative terms. An easier approach is subsequently introduced by the same authors in [102], based on an adaptation of the methods introduced in [6]. This second approach is then extended in 2d in [104]. Note that similar ideas are used in [42], in which the well-balancing is ensured for polynomial expansions of arbitrary orders and on unstructured meshes.

In [A16], we extend the previous ideas to an arbitrary order accuracy DG discretization of the PB-NSW equations on unstructured meshes. Additionally, we obtain an accuracy preserving robustness property for the mean value of the water height, adapting the ideas of [112]. Indeed, recent efforts were made in [59, 60] to adapt the FV scheme [A6] to the DG framework but only in 1d for P1 approximations (and on 2d cartesian grids in [58]).

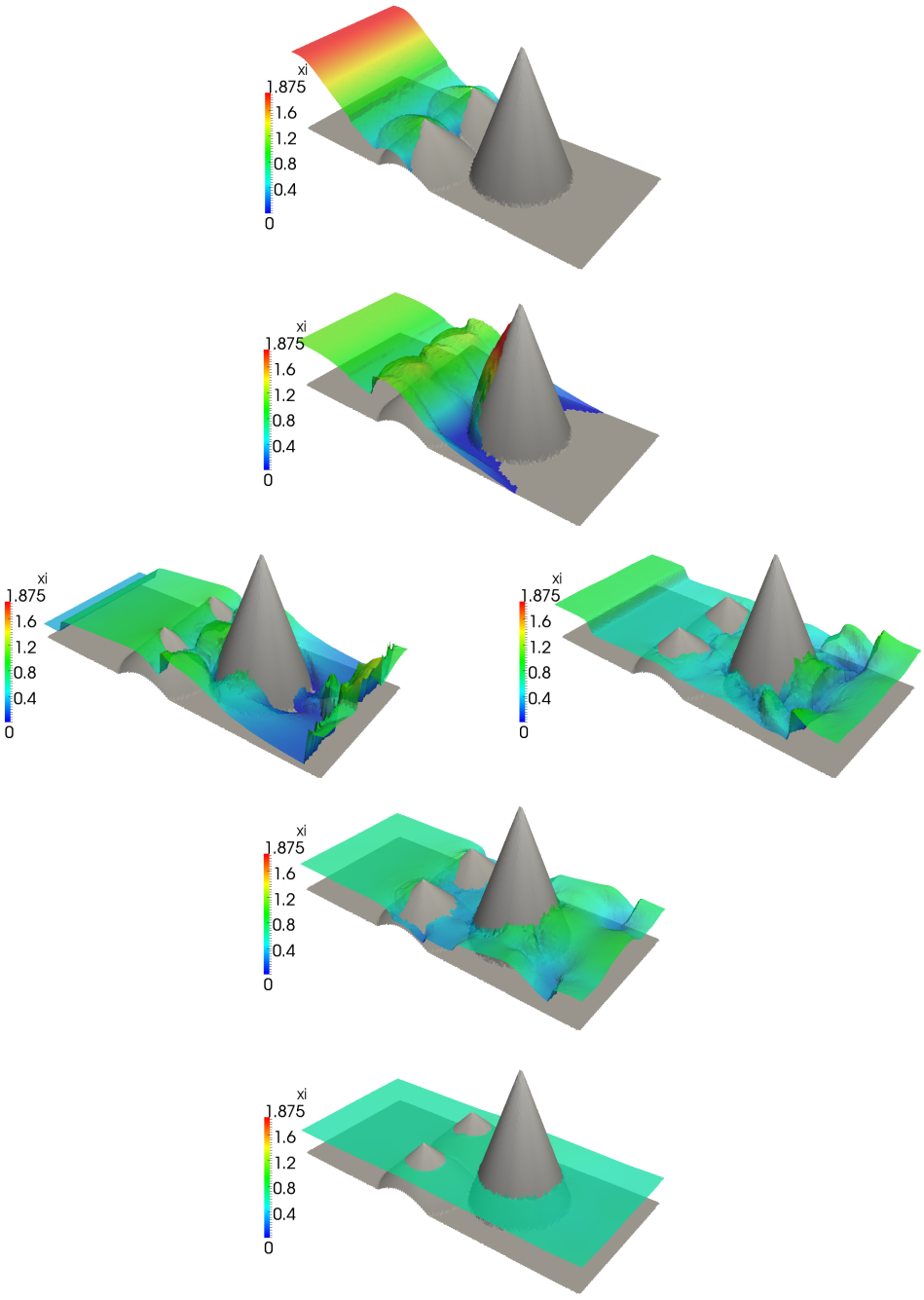

Figure 2.6: Dam break over 3 mounds: free surface at times t=2.5s, 5.5s, 16s, 23s, and 300s.

Discrete formulation and polynomial basis

The discrete weak formulation associated with (2.78) is obtained by seeking an approximation $\mathbf{U}_\mathfrak{h} \in (\mathcal{V}_\mathfrak{h})^3$ and reads for all test function $\phi_\mathfrak{h} \in \mathcal{V}_\mathfrak{h}$:

$$\int_{\mathcal{C}_l} \partial_t \mathbf{U}_\mathfrak{h} \phi_\mathfrak{h} \, d\mathbf{x} - \int_{\mathcal{C}_l} \mathcal{H}(\mathbf{U}_\mathfrak{h}, b_\mathfrak{h}) \cdot \nabla \phi_\mathfrak{h} \, d\mathbf{x} + \int_{\partial \mathcal{C}_l} \widehat{H}_l \, \phi_\mathfrak{h} \, ds = \int_{\mathcal{C}_l} \mathcal{S}(\mathbf{U}_\mathfrak{h}, b_\mathfrak{h}) \phi_\mathfrak{h} \, d\mathbf{x}. \quad (2.93)$$

In [A16], the polynomial basis $\{\theta_i^l\}_{i=1}^{N_p}$ is chosen to be an interpolant Lagrangian basis. To define the interpolant basis, we need a distribution of nodes and we choose to work with the *electrostatic points* of [54]. The nodes locations are obtained from a 2d electrostatic energy minimization problem in the reference triangle, and allow to minimize the *Lebesgue* constant of the associated interpolation problem (*i.e.* control the oscillatory nature of the interpolant polynomials for higher orders of approximation).
Looking for an approximate solution expressed as a polynomial of order p on each $\mathcal{C}_l$, the local expansion coefficients $\{\widetilde{\mathbf{U}}_i^l(t)\}_{i=1}^{N_p} = \left\{\left(\tilde{\eta}_i^l, (\tilde{q_1})_i^l, (\tilde{q_2})_i^l\right)^T\right\}_{i=1}^{N_p}$ simply are the (approximated) *nodal values* of the solution at the interpolation nodes.
To compute the last term in (2.93), we also need an expansion of the topography:

$$b_{\mathfrak{h}|\mathcal{C}_l}(\mathbf{x}) = \sum_{i=1}^{N_p} \widetilde{b}_i^l \theta_i^l(\mathbf{x}), \quad \forall \mathbf{x} \in \mathcal{C}_l, \quad (2.94)$$

where the expansion coefficients are obtained by "reading" the value of the topography at the interpolation nodes. The resulting semi-discrete formulation reads for $1 \leq j \leq N_p$:

$$\begin{aligned} \sum_{i=1}^{N_p} \left(\frac{d}{dt} \widetilde{\mathbf{U}}_i^l(t) \int_{\mathcal{C}_l} \theta_i^l \theta_j^l \, d\mathbf{x} \right) - \int_{\mathcal{C}_l} \mathcal{H}(\mathbf{U}_\mathfrak{h}, b_\mathfrak{h}) \cdot \nabla \theta_j^l \, d\mathbf{x} + \sum_{k=1}^{3} \int_{\Gamma_{l\sigma(k)}} \widehat{H}_{l\sigma(k)} \, \theta_j^l \, ds \\ = \int_{\mathcal{C}_l} \mathcal{S}_b(\mathbf{U}_\mathfrak{h}, b_\mathfrak{h}) \theta_j^l \, d\mathbf{x}. \end{aligned} \quad (2.95)$$

Numerical fluxes and well-balancing

We introduce in [A16] a new method to compute arbitrary order interface fluxes $\widehat{H}_{l\sigma(k)}$, inspired from the FV scheme of [A15]. We show that this approach leads to a well-balanced scheme that exactly preserves motionless steady states.

We define, for a given interface $\Gamma_{l\sigma(k)}$, $\mathbf{U}_k^-$ and $\mathbf{U}_k^+$ as the respective traces of $\mathbf{U}_{\mathfrak{h}|\mathcal{C}_l}$ and $\mathbf{U}_{\mathfrak{h}|\mathcal{C}_{\sigma(k)}}$ on $\Gamma_{l\sigma(k)}$. Similarly, b_k^- and b_k^+ stand for the interior and exterior values of $b_\mathfrak{h}$ on $\Gamma_{l\sigma(k)}$. For each interface, we define new interfaces values:

$$b_k^* = \max(b_k^-, b_k^+), \qquad \check{b}_k = b_k^* - \max(0, b_k^* - \eta_k^-) \quad (2.96)$$

and

$$\check{h}_k^- = \max(0, \eta_k^- - b_k^*), \qquad \check{h}_k^+ = \max(0, \eta_k^+ - b_k^*), \quad (2.97)$$

$$\check{\eta}_k^- = \check{h}_k^- + \check{b}_k, \qquad \check{\eta}_k^+ = \check{h}_k^+ + \check{b}_k, \quad (2.98)$$

leading to the new traces:

$$\check{\mathbf{U}}_k^- = (\check{\eta}_k^-, \frac{\check{h}_k^-}{\eta_k^- - b_k^-} \mathbf{q}_\mathbf{k}^-)^T, \quad \check{\mathbf{U}}_k^+ = (\check{\eta}_k^+, \frac{\check{h}_k^+}{\eta_k^+ - b_k^+} \mathbf{q}_\mathbf{k}^+)^T. \quad (2.99)$$

Now we set

$$H_{l\sigma(k)} = H(\check{\mathbf{U}}_k^-, \check{\mathbf{U}}_k^+, \check{b}_k, \check{b}_k, \mathbf{n}_{l\sigma(k)}) + \check{H}_{l\sigma(k)}, \tag{2.100}$$

as the numerical normal flux through the interface between $\mathcal{C}_l$ and $\mathcal{C}_{\sigma(k)}$, where:

1. the numerical flux H is computed from the LF flux (1.59), anticipating on the requirements for robustness. Of course any numerical flux satisfying a robustness property analogous to Lemma 2 would be suitable as well.

2. $\check{H}_{l\sigma(k)}$ is a correction term needed to ensure flux balancing at motionless steady states, defined as follows:

$$\check{H}_{l\sigma(k)} = \begin{pmatrix} 0 & 0 \\ g\eta_k^-(\check{b}_k - b_k^-) & 0 \\ 0 & g\eta_k^-(\check{b}_k - b_k^-) \end{pmatrix} \cdot \mathbf{n}_{l\sigma(k)}. \tag{2.101}$$

Note that the modified interface fluxes (2.100) induce perturbations of order $p+1$ when compared to the traditional interface fluxes. Equipped with this modified flux, we prove in [A16] the following result:

Proposition 4. *We consider the scheme (2.95), together with the interface fluxes discretization (2.100), and we assume that all the integrals involved in (2.95) are computed exactly for motionless equilibrium. Then (2.95)→(2.100) exactly preserves the motionless steady states.*

Note that taking polynomials of order 0 (*i.e.* piecewise constant functions), this well-balanced DGM coincides with the 1^{st}-order well-balanced FVM of §2.3.4.

On numerical integration

There is plenty of way to compute the surface and boundary integrals involved in (2.95). Even if some quadrature-free approaches are available (see [5] for instance), we choose in [A16] to use cubature rules for the surface integrals [33], and *Gauss* quadrature rules for the line integrals. Focusing for instance on the boundary integrals, we know that for linear problems, the use of a $(p+1)$-point *Gauss* rule is enough to compute the boundary integrals up to the chosen machine accuracy.
However for nonlinear problems like (2.95), even if more quadrature nodes are theoretically needed to ensure that the inner products of polynomials are "accurately" computed, it is a common practice to still use only $p+1$ nodes, mainly motivated by computational savings. Such *under-integration* is known to induce some aliasing of the high-modes energy of the nonlinear terms into the lower modes, but this can generally be cured with simple filtering techniques. Additionally, we are anyway committed to introduce some aliasing errors in the approximation of the NSW equations, as the momentum fluxes are computed from primitive variables, which are themselves rational functions of the conservative variables. Similar considerations also hold for the surface integrals computation.
The noticeable point is that fluxes in the momentum equations become polynomials when a motionless equilibrium is reached and the surface integrals $\int_{\mathcal{C}_l} \mathcal{H}(\mathbf{U}_\mathfrak{h}, b_\mathfrak{h}) \cdot \nabla\theta_j^l \, d\mathbf{x}$ and boundary integrals $\int_{\Gamma_{l\sigma(k)}} \widehat{H}_{l\sigma(k)} \, \theta_j^l \, ds$ can therefore be exactly computed. Such an exact computation is actually mandatory to ensure the exact preservation of the steady state, following Proposition 4. We highlight here that, within our

pre-balanced approach, the fluxes at equilibrium only involve polynomials of order p, as the free surface is constant. As a consequence, the use of $p+1$ quadrature nodes leads to an exact computation of the face integrals at equilibrium, which is an improvement when compared with the usual formulation.

Robustness

The issue of robustness for high-order schemes is a difficult one. In [111], a general procedure is documented for both WENO and DG schemes, that allows to enforce an averaged values maximum principle for general conservation laws of the form (1.46). This approach is extended to 2d on unstructured meshes in [112], and adapted to the NSW equations in [104]. In [A16], we show that the overall procedure of [112] can be successfully adapted to the PB-NSW equations. Note that when we submitted [A16], the 2d study on unstructured meshes [104] was not available yet.

The main idea of [111] to achieve the robustness is, somehow, a generalization of the ideas used for the FV schemes. But instead of using sub-grids, the idea is to decompose the element averages as a convex combination of point values of the DG polynomials on a judiciously chosen distribution of quadrature nodes. Without entering too much into details, here are the main steps to ensure robustness for the mean water height:

♯ 1 the first step is to ensure that the 1^{st}-order FV scheme associated with the mean values of η preserves the positivity of the water height with the modified fluxes (2.100). We already know that is true for the LF flux (1.59). Indeed, starting from (2.95) and using an overline $\overline{\ }$ for element averaging, such a scheme reads

$$\overline{\eta}^{n+1}_{\mathfrak{h}|\mathcal{C}_l} = \overline{\eta}^{n}_{\mathfrak{h}|\mathcal{C}_l} - \frac{\Delta t}{|\mathcal{C}_l|}\sum_{k=1}^{3}\int_{\Gamma_{l\sigma(k)}} H^{\eta}_{l\sigma(k)}\, ds, \tag{2.102}$$

with

$$H^{\eta}_{l\sigma(k)} = H^{\eta}(\check{\mathbf{U}}^{-}_{k}, \check{\mathbf{U}}^{+}_{k}, \check{b}_k, \check{b}_k, \mathbf{n}_{l\sigma(k)}), \tag{2.103}$$

and $\check{\mathbf{U}}^{-}_{k}, \check{\mathbf{U}}^{+}_{k}, \check{b}_k$ defined according to (2.96)-(2.99) and taking the piecewise constant values

$$\mathbf{U}^{-}_{k} = \overline{\mathbf{U}}^{n}_{\mathfrak{h}|\mathcal{C}_l}, \qquad \mathbf{U}^{+}_{k} = \overline{\mathbf{U}}^{n}_{\mathfrak{h}|\mathcal{C}_{\sigma(k)}}, \tag{2.104}$$

$$b^{-}_{k} = \overline{b}_{\mathfrak{h}|\mathcal{C}_l}, \qquad b^{+}_{k} = \overline{b}_{\mathfrak{h}|\mathcal{C}_{\sigma(k)}}, \tag{2.105}$$

as interior and exterior values, with respect to the element $\mathcal{C}_l$. This scheme has already been shown to be robust in Proposition 4 under the CFL condition (2.87) on the *dual mesh*. A similar result can be obtained for the *primal mesh*, see [A16].

♯ 2 the second step is to build the quadrature rule used for the convex decomposition. As shown in [112], this distribution of nodes must

a) lead to an exact computation of the mean values for a DG polynomial of order p,

b) have positive weights,

c) include, for a given element $\mathcal{C}_l$ the "boundary nodes", *i.e.* the nodes used to compute the boundary exchanging terms.

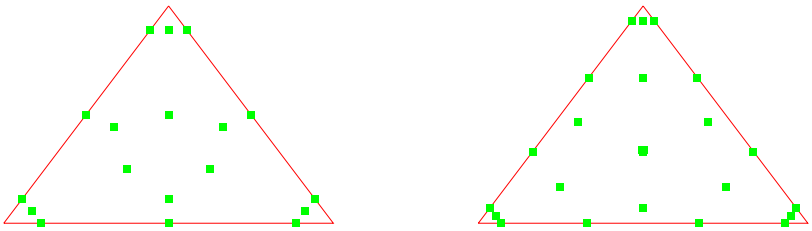

Figure 2.7: Nodes distributions for the *Zhang-Shu* quadrature - P^2 and P^3 cases.

In 1d, the quadrature rule must include the 2 extremities x_l^j and x_r^j and the LGL quadrature rule suits this task. In 2d, things are a bit more sophisticated. As we use *Gauss* rules for the boundary terms, a new quadrature rule build from a tensor product of a β-points *Gauss-Lobatto* quadrature rule and the α-points *Gauss* quadrature rule used to compute the boundary integrals, with α, β chosen accordingly to the polynomial degrees, is projected onto the simplex and is shown to fulfill all the previous desired properties. The corresponding nodes are shown on Fig. 2.7 on a given element, for 2^{nd} and 3^{rd}-order polynomials. In the following, let us denote $S_{\mathcal{C}_l}^p$ the set of points of this new quadrature rule on $\mathcal{C}_l$ and $\hat{\omega}_1^\beta$ the weight associated with the first node of the β-point *Gauss-Lobatto* quadrature.

♯ 3 equipped with this new quadrature rule, the last step is to enforce that $\forall l,\ \eta^n_{\mathfrak{h}|\mathcal{C}_l}(\mathbf{x}) - b^n_{\mathfrak{h}|\mathcal{C}_l}(\mathbf{x}) \geq 0,\ \forall \mathbf{x} \in S^p_{\mathcal{C}_l}$. This can be achieved using for instance the accuracy-preserving linear scaling suggested in [111]. Details on the implementation are given in [A16]. Note that step ♯3 has to be done at each time step or substep if a RK method is used.

Then, the following robustness result holds:

Theorem 5. *We consider the (p+1)-th order scheme (2.95), together with the interface fluxes (2.100) and a first order Euler time discretization. We assume that* $\overline{h}^n_{\mathfrak{h}|\mathcal{C}_l} \geq 0$ *and that* $\eta^n_{\mathfrak{h}|\mathcal{C}_l}(\mathbf{x}) - b^n_{\mathfrak{h}|\mathcal{C}_l}(\mathbf{x}) \geq 0,\ \forall \mathbf{x} \in S^d_{\mathcal{C}_l}, \forall l$. *Then we have* $\overline{h}^{n+1}_{\mathfrak{h}|\mathcal{C}_l} \geq 0$ *under the condition*

$$\lambda_o \frac{\mathfrak{p}_l}{|\mathcal{C}_l|} \Delta t \leq \frac{2}{3} \hat{\omega}_1^\beta, \quad \forall l. \tag{2.106}$$

It is interesting to see that this robustness result does not prevent $h_\mathfrak{h}$ to locally take negative values, on a given element $\mathcal{C}_l$. However, step ♯3 ensures a robust computation in the following sense:

1. the averaged values $\overline{h}^n_{\mathfrak{h}|\mathcal{C}_l} \geq 0$ remain positive,
2. the point-wise values of the traces $h_{\mathfrak{h}|\partial\mathcal{C}_l}$ at all the *Gauss* quadrature nodes, on all element edges, are positive. We recall that a *Riemann* problem has to be solved at each edge's quadrature node to compute the boundary integrals and that the corresponding values of h are used to compute the primitive variables.

3

Green-Naghdi equations

In this 3^{rd} chapter, we turn on considering the GN equations (1.36)-(1.37). As an introduction, I briefly draw the main lines of the physical motivations behind our efforts.

We know that wave propagation in shallow water, and associated processes such as shoaling, wave-breaking and run-up, play an important role in the nearshore hydro-dynamics. The classical method of describing slowly evolving wave-induced circulation in the nearshore is based on the phase-averaged approach, in which the depth-integrated mass and momentum equations are time-averaged over a wave period [72]. However, a detailed knowledge of instantaneous nonlinear wave characteristics is required in place of wave-averaged quantities for important unsteady processes, such as wave run-up in the swash zone, coastal flooding during storms, tsunami and tidal bore propagation or suspended sediment transport. For coastal applications, these *phase-resolving* models are based on NSW, BT and GN equations. NSW equations give a good description of the nonlinear non-dispersive transformations of breaking waves, represented as shocks, in the inner surf and swash zones [17]. However, due to the absence of frequency dispersion, the NSW equations can not be applied to wave propagation before breaking. On the other hand, BT and GN equations incorporate frequency dispersion and can be applied to wave shoaling.

Over the last 40 years, BT equations, and related *weakly nonlinear* models based on a *small amplitude* assumption $\varepsilon = \mathcal{O}(\mu)$, have focused increasing attention. A large number of numerical methods have been developed in the past few years for BT equations. Let us mention for instance some FDMs [80, 101], FEMs [70, 86], FVMs [38, 57], spectral methods [43, 27], hybrid FVM-FDM [39, 99] or DGMs [44]. However, the *small amplitude* assumption generally severely restrict applicability to real nearshore applications. Indeed, in the final stages of shoaling or in the surf zone, the wave dynamics is strongly nonlinear: $\varepsilon = \mathcal{O}(1)$. For instance, ε is close to 0.4 just before breaking and can be larger than 1 in the swash zone.

For these reasons, it is now commonly acknowledged that the GN equations are the relevant system to model highly nonlinear weakly dispersive waves propagating in shallow water. However, from a numerical point of view, the GN equations have received far less attention. When we started working on the numerical approximations of the GN equations, the operational models were mostly based on FDMs, see for instance [56]. However, an increasing interest for the development of modern methods was emerging, with for instance the compact FVM discretization proposed for the 1d *Serre* equations in [30], the pseudo-spectral approach of [81] or the FVM

for a particular 2d flat bottom system derived from a hamiltonian formulation in [63].

To complete the picture, we also have to bring the important issues of *wave breaking*, *moving shoreline* and *dispersive properties* into the discussion. Indeed:

- the (linear) dispersive properties of the BT and GN equations depend on their dispersion relations, and associated *phase* and *group* velocities. For a given value of μ, a model will be valid much further seaward if its dispersive properties are kept close to the properties of the *water waves* equations for the largest values of $\mathbf{k}$. It also gives better results when high frequency are released, for instance when a wave propagates above an obstacle. There is consequently a need for improving the dispersive properties of the GN equations to extend their range of validity seaward. This issue is investigated in §3.1.3 and §3.1.6.
- unlike the NSW equations, BT and GN equations incorporate frequency dispersion and can be used to model the wave shoaling But they do not naturally account for wave breaking. Consequently, to extend the range of validity of the GN equations to the surf zone, significant efforts have to be devoted to numerically handle wave breaking mechanism. This issue is investigated in §3.3.
- accounting for the possibility of a vanishing depth in GN and BT equations while keeping the dispersive effects is mandatory to finally extend the range of validity to the swash zone and obtain a complete operational nearshore model accounting for waves run-up, overtopping and submersion. Some approaches were proposed using for instance Lagrangian techniques [110], porous-seabed technique [98] or linear extrapolations [73] but they suffer from either loss of mass conservation or from instabilities and lack of robustness. We refer to §3.2 for appropriate answers.

Facing these issues, we derive and discretize in a series of papers [A10, A11, A19, A20], see also the review paper [A9], a complete new hierarchy of models, all belonging to the GN family in the sense of a $\mathcal{O}(\mu^2)$ consistency with the *water waves* equations (1.26)-(1.27) . These new models are mostly motivated by numerical requirements, keeping in mind the underlying targeted objectives:

♯1 reaching a high order of accuracy for a proper description of dispersive waves propagation and transformations,

♯2 ensuring the robustness and stability of the numerical method,

♯3 ensuring the preservation of motionless steady states,

♯4 including a flexible and robust wave breaking mechanism,

♯5 improving the dispersive properties to extend the range of validity of the model seaward,

♯6 handling complex geometries and mesh refinement (adaptivity),

♯7 keeping the computational cost as decent as possible, targeting long time integrations of large scale problems.

Of course, meeting all these features concurrently is all about trade-offs. The remainder of this Chapter is organized as follows. We introduce and motivate, in a comprehensive way, our hierarchy of GN models in §3.1. In §3.2, we briefly detail the proposed numerical strategies to achieve the listed objectives, while §3.3 is devoted to the description of some efficient wave breaking mechanism.

3.1 A hierarchy of models

3.1.1 Factorized equations

Let us start with the nondimensionalized GN equations as recalled in §1.1.5:

$$\left\{\begin{array}{l} \partial_t\zeta + \nabla\cdot(h\mathbf{v}) = 0, \\ (I+\mu\mathcal{T}[h,b])\partial_t\mathbf{v} + \nabla\zeta + \varepsilon(\mathbf{v}\cdot\nabla)\mathbf{v} + \varepsilon\mu\mathcal{Q}[h,b](\mathbf{v}) = 0, \end{array}\right. \tag{3.1}$$

with the nonlinear operator $\mathcal{Q}$ defined in (1.38). We begin by performing a suitable factorization of the left hand side of (3.1). Introducing the notation $\mathbf{v}^\perp = (-v_2, v_1)^T$, we show in [A11] that (3.1) can be reformulated as follows

$$\left\{\begin{array}{l} \partial_t\zeta + \nabla\cdot(h\mathbf{v}) = 0, \\ [I+\mu\mathcal{T}]\partial_t\mathbf{v} + \varepsilon[I+\mu\mathcal{T}](\mathbf{v}\cdot\nabla)\mathbf{v} + \nabla\zeta + \varepsilon\mu\mathcal{Q}_1[h,b](\mathbf{v}) = 0, \end{array}\right. \tag{3.2}$$

with

$$\mathcal{Q}_1[h,b](\mathbf{v}) \quad = \quad -2\mathcal{R}_1(\partial_1\mathbf{v}\cdot\partial_2\mathbf{v}^\perp + (\nabla\cdot\mathbf{v})^2) + \beta\mathcal{R}_2(\mathbf{v}\cdot(\mathbf{v}\cdot\nabla)\nabla b).$$

Note that (3.2) is exactly equivalent to (3.1). The advantages of (3.2) are twofolds:

1. it allows to recover a *transport part* corresponding to the NSW equations: the dispersive terms can now be regarded as *source terms* for the hyperbolic equations, as $I+\mu\mathcal{T}$ is invertible thanks to a coercivity property and the *Lax-Milgram* theorem, see [62].

2. there is no 3^{rd}-order derivatives in $\mathcal{Q}_1[h,b]$, which is an interesting feature for both theoretical analysis and numerical stability.

The GN equations (3.2) are stated as two evolution equations for ζ and $\mathbf{v}$. It is of course possible to give an equivalent formulation as a system of two evolution equations on h and $h\mathbf{v}$. With the notation $\mathbf{T} = h\mathcal{T}\frac{1}{h}$, we obtain

$$(\mathcal{G})\left\{\begin{array}{l} \partial_t h + \varepsilon\nabla\cdot(h\mathbf{v}) = 0, \\ [I+\mu\mathbf{T}]\Big(\partial_t(h\mathbf{v}) + \varepsilon\nabla\cdot(h\mathbf{v}\otimes\mathbf{v})\Big) + h\nabla\zeta + \varepsilon\mu h\mathcal{Q}_1(\mathbf{v}) = 0. \end{array}\right. \tag{3.3}$$

Considering dimensionalized variables, the (linear) dispersive properties of $(\mathcal{G})$ (or equivalently of (1.36)-(1.37)) depend on the set of all $(\omega,\mathbf{k})\in\mathbb{R}\times\mathbb{R}^d$ such that there exists a plane wave solution $(\zeta_0,\psi_0)e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}$ to the linearization of (3.3) around the rest state $(h=h_0,\mathbf{v}=0)$, for flat bottoms. The frequency dispersion is consequently ruled by the following dispersion relation:

$$\omega_{\mathcal{G}}(\mathbf{k})^2 = \frac{|\mathbf{k}|^2 g h_0}{1+\dfrac{(|\mathbf{k}|h_0)^2}{3}}. \tag{3.4}$$

3.1.2 GN equations with improved dispersion relation

The frequency dispersion of (3.2) can be improved by adding some terms of order $O(\mu^2)$ to the momentum equation. Since this equation is already precise up to terms of order $O(\mu^2)$, this manipulation does not affect the accuracy of the model. The underlying idea, adapted from the BBM trick [9], is to realize that

$$\partial_t \mathbf{v} = -\nabla\zeta - \varepsilon(\mathbf{v}\cdot\nabla)\mathbf{v} + O(\mu),$$

and that consequently, for any parameter $\alpha \in [0,1]$, we have:

$$\partial_t \mathbf{v} = \alpha\partial_t \mathbf{v} - (1-\alpha)(\nabla\zeta + \varepsilon(\mathbf{v}\cdot\nabla)\mathbf{v}) + O(\mu).$$

Replacing $\partial_t \mathbf{v}$ by this expression in (3.2), dropping the $O(\mu^2)$ terms, and using $(h, h\mathbf{v})$ variables we obtain the equivalent GN equations with a free parameter

$$\left\{\begin{array}{l} \partial_t h + \varepsilon\nabla\cdot(h\mathbf{v}) = 0, \\ [I + \mu\alpha\mathbf{T}]\Big(\partial_t(h\mathbf{v}) + \varepsilon\nabla\cdot(h\mathbf{v}\otimes\mathbf{v})\Big) + (I - \mu(1-\alpha)\mathbf{T})h\nabla\zeta + \varepsilon\mu h\mathcal{Q}_1(\mathbf{v}) = 0. \end{array}\right. \tag{3.5}$$

If we gain some freedom for improving the dispersion properties of our model, we now have the second equation of (3.5) that requires the computation of third order derivatives of ζ. Fortunately, it is possible to show that these terms can be factorized by $I + \mu\alpha\mathbf{T}$. We obtain the following new family of 1-parameter $(\mathcal{G}_\alpha)$ models:

$$(\,\mathcal{G}_\alpha\,)\left\{\begin{array}{l} \partial_t h + \varepsilon\nabla\cdot(h\mathbf{v}) = 0, \\ [I + \alpha\mu\mathbf{T}]\Big(\partial_t(h\mathbf{v}) + \dfrac{\alpha-1}{\alpha}h\nabla\zeta + \varepsilon\nabla\cdot(h\mathbf{v}\otimes\mathbf{v})\Big) + \dfrac{1}{\alpha}h\nabla\zeta + \varepsilon\mu h\mathcal{Q}_1(\mathbf{v}) = 0, \end{array}\right. \tag{3.6}$$

which is the family of models used for numerical simulations in [A11]. The dispersion relation of $(\,\mathcal{G}_\alpha\,)$ is given by

$$\omega_{\mathcal{G}_\alpha}(\mathbf{k})^2 = |\mathbf{k}|^2 g h_0 \frac{1 + \dfrac{(\alpha-1)(|\mathbf{k}|h_0)^2}{3}}{1 + \dfrac{\alpha(|\mathbf{k}|h_0)^2}{3}}. \tag{3.7}$$

For the applications we have in mind, we are rather interested in using a model which has the widest possible range of validity and provide a correct description of the waves in intermediate or even deep water. Using (3.7), we deduce the expressions of the associated linear phase and group velocities and an optimal value of α is found by minimizing a weighted averaged error obtained by comparing these linear phase and group $(\,\mathcal{G}_\alpha\,)$ velocities with the linear phase and group velocities coming from the *water waves* equations (1.26)-(1.27), over the range $|\mathbf{k}|h_0 \in [0,4]$, see Fig. 3.1.3.

3.1.3 A family of 3-parameters models

In this subsection, we follow the approach of [A10] and only focus on 1d propagating waves. We'll come back to 2d models in §3.1.4. It was remarked quite early that it was possible to obtain *enhanced* Boussinesq systems by replacing the vertically averaged horizontal velocity $\mathbf{v}$ by another unknown: for instance in [77] $\mathbf{v}$ is replaced by the horizontal velocity at the bottom. In [80], it is showed that it was possible to improve the dispersive properties of BT models by working with the velocity at a

certain depth as dependent variable. This approach was generalized in [101] to the fully nonlinear case.

However, when the bottom is not flat, the GN equations written with the average velocity $\mathbf{v}$ do not belong to this class of fully nonlinear models. This is the reason why we use a slightly different approach, with the introduction of a new dependent variable v_θ that is not the velocity at a certain depth. The interest is that the computations are somehow simpler and, more importantly, that the average velocity v appears as a particular case ($\theta = 0$). Starting from the 1d version of (3.2), we define for all $\theta \geq 0$:

$$v_\theta = (1 + \mu\theta\mathcal{T})^{-1} v, \tag{3.8}$$

and inject it into (3.2). After some computations, introducing also the parameter α as previously described, we obtain the following family of *2 parameters* $(\mathcal{G}_{\alpha,\theta})$ models, which is equivalent, up to $O(\mu^2)$ terms to (3.6):

$$(\mathcal{G}_{\alpha,\theta}) \begin{cases} \partial_t h + \varepsilon\partial_x(hv_\theta) + \varepsilon\mu\theta\partial_x(h\mathcal{T}v_\theta) = 0, \\ [1 + \mu(\alpha+\theta)\mathbf{T}]\Big(\partial_t(hv_\theta) + \varepsilon\partial_x(hv_\theta^2) + \frac{\alpha-1}{\alpha+\theta}h\partial_x\zeta\Big) \\ \qquad\qquad + \frac{1+\theta}{\alpha+\theta}h\partial_x\zeta + \varepsilon\mu h\widetilde{\mathcal{Q}}(v_\theta) = 0. \end{cases} \tag{3.9}$$

We refer to [A10] for the definition of the new nonlinear operator $\widetilde{\mathcal{Q}}$, which still has the kindness of not involving any 3^{rd}-order derivatives.
At this step, the introduction of the parameter θ reveals to be useless as far as the minimization of the dispersion errors is concerned. Indeed, keeping in mind that $\alpha \geq 1$ and $\theta \geq 0$, computing the dispersion relation associated to these new equations and minimizing the dispersion errors over the range $kh \in [0, 4]$ yields the values $\alpha = 1.159$ and $\theta = 0$.

A workaround for this issue is to introduce a third parameter $\gamma \geq 0$, defining for all $\theta, \gamma \geq 0$

$$v_{\theta,\gamma} = (1 + \mu\theta\mathcal{T})^{-1}(1 + \mu\gamma\mathcal{T})v, \tag{3.10}$$

and inject it into (3.2) instead of (3.8). This is the approach suggested in [62]. In [A10], we use an equivalent approach, applying directly the operator $[1 + \mu\gamma\mathbf{T}]$ to the first equation of (3.9). On the whole, this final manipulation yields a new 3-parameters family of GN models, denoted $(\mathcal{G}_{\alpha,\theta,\gamma})$, that can be written in dimensional form as

$$(\mathcal{G}_{\alpha,\theta,\gamma}) \begin{cases} [1 + \gamma\mathbf{T}]\Big(\partial_t h + \partial_x(hv_\theta)\Big) + \theta\partial_x(h\mathcal{T}v_\theta) = 0, \\ [1 + (\alpha+\theta)\mathbf{T}]\Big(\partial_t(hv_\theta) + \partial_x(hv_\theta^2) + \dfrac{\alpha-1}{\alpha+\theta}gh\partial_x\zeta\Big) \\ \qquad\qquad + \dfrac{1+\theta}{\alpha+\theta}gh\partial_x\zeta + h\widetilde{\mathcal{Q}}(v_\theta) = 0, \end{cases} \tag{3.11}$$

We insist on the fact that all these models are equivalent at order $O(\mu^2)$ to the GN equations (3.3), which correspond to the particular case $\alpha = 1$, $\theta = \gamma = 0$. The dispersion relation associated to (3.11) is given by,

$$\omega^2_{\mathcal{G}_{\alpha,\theta,\gamma}}(k) = gh_0k^2\frac{(1 + \frac{\theta+\gamma}{3}(kh_0)^2)(1 + \frac{\alpha-1}{3}(kh_0)^2)}{(1 + \frac{\gamma}{3}(kh_0)^2)(1 + \frac{\alpha+\theta}{3}(kh_0)^2)}.$$

and again, minimization processes based on the linear phase and group velocities of (3.11), denoted respectively $C^p_{\alpha,\theta,\gamma}$ and $C^g_{\alpha,\theta,\gamma}$, allow to find optimized sets of parameters (α, θ, γ) that extend the range of validity of our models. We show on Fig. 3.1.3. a comparison between the values of $C^p_{\alpha,\theta,\gamma}$ and $C^g_{\alpha,\theta,\gamma}$ and the targeted velocities C^p_w, C^g_w coming from the *water waves* equations, for different values of the parameters. The numerical results obtained in [A10], for a simple case of periodic waves propagating over a submerged bar, confirm the improvements.

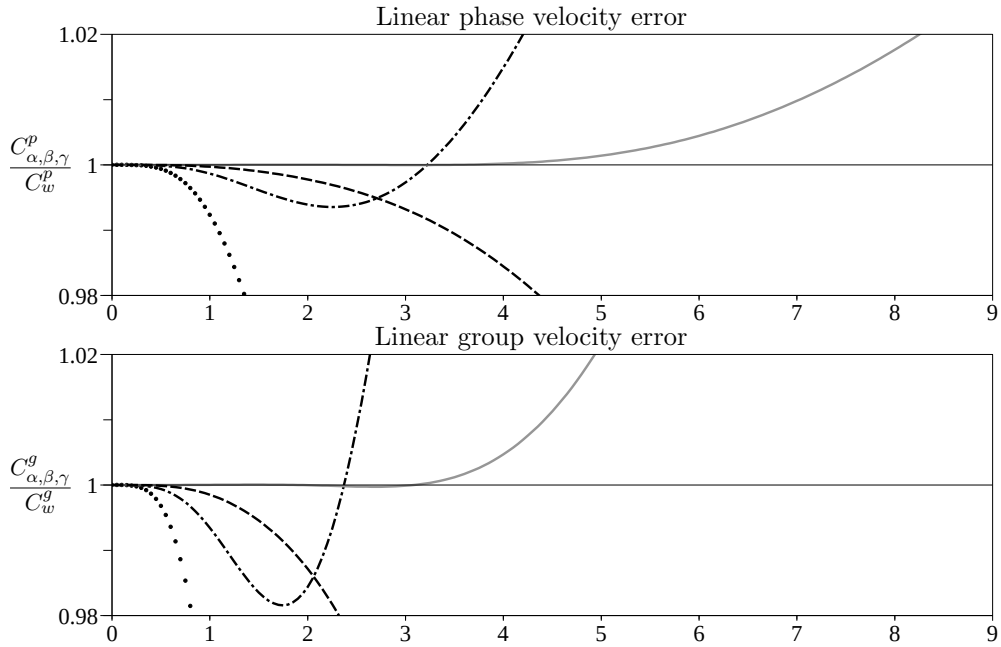

Figure 3.1: Errors on linear phase (top) and group velocities (bottom). The $(\mathcal{G})$ model is in dots, the optimized $(\mathcal{G}_\alpha)$ model in dash-dots, the optimized $(\mathcal{G}_{\alpha,\theta,\gamma})$ model in grey plain line, and the $(\mathcal{G}_{\alpha,\theta,\gamma})$ model with optimized coefficients for shoaling in dashes.

◇ ◇ ◇ Interlude ◇ ◇ ◇

Following the "historical" development of our works, we started to develop 1d numerical approximations of the $(\mathcal{G}_\alpha)$ and $(\mathcal{G}_{\alpha,\theta,\gamma})$ models respectively in [A11] and [A10]. When coming alongside with the first 2d simulations, it rapidly became clear that while the extensions of our numerical methods to 2d surface waves do not raise theoretical difficulties, most of the computational time was absorbed in the inversion of the discrete approximation of the operator $I + \mu\alpha\mathbf{T}$, and that it required a lot of computational resources, even for reasonably sized problems. Indeed, the discrete inversion of $I+\mu\alpha\mathbf{T}$ is computationally demanding in the $2d$ case for several reasons:

♯ 1 it is a *matricial* operator of size 2×2 which is moreover not in diagonal form; the operator $\mathbf{T}$ can be written in matricial form as

$$\mathbf{T} = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix},$$

where T_{ij} $(1 \leq i, j \leq 2)$ are second order *scalar* differential operators. The fact that the antidiagonal terms are nonzero complicates considerably the numerical computation of $(I + \mu\alpha\mathbf{T})^{-1}$. Indeed, once discretized, this leads to the resolution of an $2N_{dof} \times 2N_{dof}$ linear system, eventually with many nonzero diagonals, where N_{dof} is the number freedom degrees introduced in the domain.

♯ 2 it is a *time depending* operator through the water height h (remember that $\mathcal{T}$ is a simplified notation for $\mathcal{T}[h, b]$) and the corresponding matrix has therefore to be assembled at each time step, not to mention the subsequent factorization (*resp.* preconditioning/renumbering) step if a direct (*resp.* iterative) linear solver is used.

These computational difficulties are actually one of the reasons why most of the $2d$ numerical codes are based on *weakly nonlinear* Boussinesq models rather than the *fully nonlinear* GN equations. Of course, one could entirely treat the dispersive components as source terms and therefore avoid the inversion of the operator $I + \mu\alpha\mathbf{T}$, but one would loose the high robustness of the code, and indeed, our numerical investigations have highlighted strong instabilities, even for weakly nonlinear problems.

The aim of [A20] is to overcome these difficulties. In order to keep the stabilizing effects of the inversion step without loosing too much in terms of computational time, we derive some new families of physical models that are equivalent to the standard GN equations. It is possible to implement our numerical strategies on these new systems with the benefit of removing the two numerical obstructions mentioned above. This is done in two steps:

♯ 1 for the systems of §3.1.4, the matricial operator $\mathbf{T}$ is replaced by a new differential operator $\mathbf{T}_{diag}$, which is still matricial but has a diagonal structure:

$$\mathbf{T}_{diag} = \begin{pmatrix} \tilde{T}_{11} & 0 \\ 0 & \tilde{T}_{22} \end{pmatrix}. \tag{3.12}$$

Its inversion is therefore equivalent to the inversion of two *scalar* operators; numerically, this leads to the resolution of 2 sparse $N_{dof} \times N_{dof}$ linear systems.

♯ 2 for the systems of §3.1.5, the matricial operator $\mathbf{T}$ is replaced by an operator $\mathbf{T}^b_{diag}$ which, as $\mathbf{T}_{diag}$, has a diagonal structure, but for which the dependence on h has been removed; it is therefore *time independent*. A motivating idea came from the observation of the weakly nonlinear *Boussinesq-Peregrine* model (1.39)-(1.40) which is interesting from a computational point of view, since the 2^{nd}-order operator $I + \mu\mathcal{T}[h_b, b]$ only depend on the bottom variations.

The simple structure of these systems leads to considerable improvements in terms of computational time, without any loss of accuracy when compared to the standard model (3.1). In particular, they allow us to consider:

♯ 1 "reasonably" large scale $2d$ configurations (the parallelisation may allow us to remove the "reasonably" adverb), see the on-going works Chapter,

♯ 2 the construction of dispersion enhanced 2d models for configurations that involve deep water waves for instance, see §3.1.6,

♯ 3 the use of very high order methods (with the inherent increasing of freedom degrees), unavoidable for the study of long time integrations of small scale perturbations, see §3.2.2.

◇ ◇ ◇

3.1.4 The diagonal model

To obtain a diagonal model, we use the fact that

$$\partial_t \mathbf{v} + \varepsilon(\mathbf{v}\cdot\nabla)\mathbf{v} = -\nabla\zeta + O(\mu), \tag{3.13}$$

and rewrite the second equation of (3.2) under the form

$$[I + \mu\mathcal{T}_{diag}]\Big(\partial_t \mathbf{v} + \varepsilon(\mathbf{v}\cdot\nabla)\mathbf{v}\Big) + \nabla\zeta + \varepsilon\mu\mathcal{Q}_1(\mathbf{v}) + \mu\mathcal{Q}_2(\zeta) = O(\mu^2),$$

where $\mathcal{T}_{diag}$ has a "diagonal" structure (as discussed in the previous section):

$$\mathcal{T}_{diag} = -\frac{1}{3h}\partial_1(h^3\partial_1\cdot) - \frac{1}{3h}\partial_2(h^3\partial_2\cdot),$$

while $\mathcal{Q}_2(\zeta)$ only involves derivatives of ζ up to 2^{nd}-order

$$\mathcal{Q}_2(\zeta) = -h(\nabla^\perp h\cdot\nabla)\nabla^\perp\zeta - \frac{\beta}{2h}\nabla(h^2\nabla b\cdot\nabla\zeta) + \beta(\frac{h}{2}\Delta\zeta - \beta(\nabla b\cdot\nabla\zeta))\nabla b. \tag{3.14}$$

Replacing this new formulation in (3.2), neglecting $O(\mu^2)$ terms, switching to $(h, h\mathbf{v})$ variables and using again the BBM trick to introduce an additional enhancing parameter α, we obtain the following family of non-dimensionalized 1-parameter asymptotically equivalent systems

$$(\ \mathcal{G}_\alpha^{diag}\ )\quad \left\{\begin{array}{l} \partial_t h + \varepsilon\nabla\cdot(h\mathbf{v}) = 0, \\ [I + \mu\alpha\mathbf{T}_{diag}]\Big(\partial_t(h\mathbf{v}) + \varepsilon\nabla\cdot(h\mathbf{v}\otimes\mathbf{v}) + \dfrac{\alpha-1}{\alpha}h\nabla\zeta\Big) \\ \qquad\qquad + \frac{1}{\alpha}h\nabla\zeta + \varepsilon\mu h\mathcal{Q}_1(\mathbf{v}) + \mu h\mathcal{Q}_2(\zeta) = 0. \end{array}\right. \tag{3.15}$$

with

$$\mathbf{T}_{diag} = h\mathcal{T}_{diag}\frac{1}{h}.$$

The interest of working with (3.15) rather than (3.6) is that instead of inverting the operator $I+\mu\mathbf{T}$, we have to invert $I+\mu\mathbf{T}_{diag}$ that has the diagonal structure (3.12), with

$$\widetilde{T}_{11} = \widetilde{T}_{22} = -\frac{1}{3}\partial_1(h^3\partial_1\frac{1}{h}\cdot) - \frac{1}{3}\partial_2(h^3\partial_2\frac{1}{h}\cdot).$$

3.1.5 The diagonal-constant model

It is now possible to replace the inversion of $I+\mu\mathbf{T}_{diag}$ by the inversion of $I+\mu\mathbf{T}^b_{diag}$, where $\mathbf{T}^b_{diag}$ corresponds to $\mathbf{T}_{diag}$ when the fluid is at rest (i.e. $\zeta = 0$),

$$\mathbf{T}^b_{diag} = h_b\mathcal{T}^b_{diag}\frac{1}{h_b}, \quad \text{with} \quad \mathcal{T}^b_{diag} = -\frac{1}{3h_b}\partial_1(h_b^3\partial_1\cdot) - \frac{1}{3h_b}\partial_2(h_b^3\partial_2\cdot),$$

and where h_b is the water depth at rest,

$$h_b = 1 - \beta b = h - \varepsilon\zeta.$$

The ingredients to achieve such a formulation are first to remark that for all scalar function f, one has

$$h\mathcal{T}_{diag}\frac{1}{h}f = h_b\mathcal{T}^b_{diag}\frac{1}{h_b}f + \mathcal{R}(f)$$

with $\mathcal{R}$ a 2^{nd}-order operator, and then to use the identity

$$\partial_t(h\mathbf{v}) + \varepsilon\nabla\cdot(h\mathbf{v}\otimes\mathbf{v}) = -\left[I + \mu\mathbf{T}^b_{diag}\right]^{-1}(h\nabla\zeta) + O(\mu), \tag{3.16}$$

instead of (3.13). Indeed, we show in [A20], with the help of a dispersion properties analysis, that a straightforward use of (3.13) would lead to a family of models that are inherently linearly unstable. We obtain the new 1-parameter *diagonal constant* ($\mathcal{CG}^{diag}_{\alpha}$) family of models:

$$(\,\mathcal{CG}^{diag}_{\alpha}\,)\begin{cases} \partial_t h + \varepsilon\nabla\cdot(h\mathbf{v}) = 0, \\ \left[I + \mu\alpha\mathbf{T}^b_{diag}\right](\partial_t(h\mathbf{v}) + \varepsilon\nabla\cdot(h\mathbf{v}\otimes\mathbf{v}) + \dfrac{\alpha-1}{\alpha}h\nabla\zeta) + \dfrac{1}{\alpha}h\nabla\zeta \\ \qquad\qquad +\varepsilon\mu h\mathcal{Q}_1(\mathbf{v}) + \mu h\mathcal{Q}_2(\zeta) + \mu\mathcal{Q}_3(\zeta) = 0. \end{cases} \tag{3.17}$$

We refer to [A20] for the definition of the nonlinear differential operator $\mathcal{Q}_3$, and a discussion about the linear dispersive properties and the inherited linear stability of the model.

3.1.6 The three-parameters diagonal-constant models

The family of 3-parameters models ($\mathcal{G}_{\alpha,\theta,\gamma}$) derived in [A10] would easily be generalized to the $2d$ case, but the computational cost would be prohibitive. To obtain a family of 3-parameters family of *diagonal-constant* models we introduce, as in the 1d case, a new velocity $\mathbf{v}_\theta$ as follows

$$\forall\theta \geq 0, \qquad \mathbf{v} = (1 + \mu\theta\mathcal{T}_{diag})^{-1}\mathbf{v}_\theta.$$

After some tedious computations, and neglecting all the $O(\mu^2)$ terms, the family of 3-parameters constant-diagonal ($\mathcal{CG}^{diag}_{\alpha,\theta,\gamma}$) models is given by

$$(\,\mathcal{CG}^{diag}_{\alpha,\theta,\gamma}\,)\begin{cases} (1 + \mu\gamma\mathcal{T}^b_{diag})[\partial_t\zeta + \nabla\cdot(h\mathbf{v}_\theta)] + \mu\theta\nabla\cdot(h\mathcal{T}_{diag}\mathbf{v}_\theta) = 0 \\ (I + \mu\alpha(1+\theta)\mathbf{T}^b_{diag})[\partial_t(h\mathbf{v}_\theta) + \varepsilon\nabla\cdot(h\mathbf{v}_\theta\otimes\mathbf{v}_\theta) + \dfrac{\alpha-1}{\alpha}\nabla\zeta] + \dfrac{1}{\alpha}\nabla\zeta \\ \qquad +\varepsilon\mu h\mathcal{Q}_1(\mathbf{v}_\theta) + \mu h\mathcal{Q}_2(\zeta) + \mu(1+\theta)\mathcal{Q}_3(\zeta) + \mu\theta\mathcal{Q}_4(\mathbf{v}_\theta) = 0, \end{cases} \tag{3.18}$$

with $\mathcal{Q}_j$ $(j = 1,\ldots,4)$ defined in [A20]. Taking $\theta = \gamma = 0$, (3.18) coincides with the ($\mathcal{CG}^{diag}_{\alpha}$) family. Again, an optimization of the phase and group velocities leads to optimized sets of coefficients. The linearized system and associate dispersion relation are detailed in the communication [C13]. In some stiff configurations where higher harmonics are released, working with the three parameters optimized model (3.18) leads to considerable improvements.

3.2 Discretization

So far, even if mainly inspired by numerical considerations, we only focused on the models derivations. Let us now give some details regarding the numerical strategies investigated and applied in [A9, A10, A11, A12, A14, A19, A20]. The first approach relies on a splitting scheme and will be referred to as the *transport + dispersion correction* method in the following. Next, in §3.2.2, we describe a recent unsplitted DG discretization, which has the benefit of providing an arbitrary order of convergence and more geometric flexibility.

3.2.1 The *transport + dispersion correction* approach

We develop in [A11] a new strategy for the numerical approximation of the $(\mathcal{G}_\alpha)$ family. This strategy is subsequently extended to the $(\mathcal{G}_{\alpha,\theta,\gamma})$ family in [A10] and to the $(\mathcal{CG}^{diag}_{\alpha})$ and $(\mathcal{CG}^{diag}_{\alpha,\theta,\gamma})$ families in [A20]. The underlying idea is to consider the dispersive terms as source terms in the NSW equations, introduce a suitable splitting and discretize the hyperbolic part with the battery of efficient and highly validated schemes described in Chapter 2.

We approximate the solution operator $S(\cdot)$ associated to the dimensionalized version of (3.6) at each time step by the 2^{nd} order splitting scheme

$$S(\Delta t) \simeq S_1(\Delta t/2) S_2(\Delta t) S_1(\Delta t/2), \tag{3.19}$$

with

♯1 *a transport step* $S_1(t)$:

$$\begin{cases} \partial_t \zeta + \nabla \cdot (h\mathbf{v}) = 0, \\ \partial_t (h\mathbf{v}) + h\nabla\zeta + \nabla \cdot (h\mathbf{v} \otimes \mathbf{v}) = 0, \end{cases} \tag{3.20}$$

♯2 *a momentum dispersion correction step* $S_2(t)$:

$$\begin{cases} \partial_t \zeta = 0, \\ [I + \alpha \mathbf{T}]\Big(\partial_t (h\mathbf{v}) - \dfrac{1}{\alpha} gh\nabla\zeta\Big) + \dfrac{1}{\alpha} gh\nabla\zeta + h\mathcal{Q}_1(\mathbf{v}) = 0. \end{cases} \tag{3.21}$$

This splitting approach raises several remarks:

1. we choose to split the equations, in such a way that the optimization parameter α does not appear in the *transport step*, allowing this step to exactly coincide with the resolution of the NSW equations,

2. only the momentum is concerned with the *dispersion correction* step. When compared for instance with the widely used weakly nonlinear equations of *Nwogu* [80], or the fully non-linear equations of *Chen* [28], this is a real advantage of this formulation. Indeed, it allows to inherit the robustness and well-balancing properties from the discretization of the NSW equations.

Using a 4^{th}-order *Runge-Kutta* time marching algorithm for both steps of the splitting, we show in [A11] that the semi-discretized dispersion relation approaches the exact dispersion relation of the GN equations (3.6) at order 2 in Δt:

Proposition 5. *The dispersion relation associated to the semi-discretized scheme (3.19), (3.20), (3.21) is given by*

$$\omega_{sd,\pm}(k) = \omega_{\alpha,\pm}(k) + \frac{\Delta t^2}{24}\omega_{\alpha,\pm}(k)^3\Big(\frac{(kh_0)^2}{3+(\alpha-1)(kh_0)^2}\Big)^2 + O(\Delta t^3).$$

An additional information is that the $O(\Delta t^2)$ error made by the splitting scheme is always real. Therefore, the numerical errors are of *dispersive* type and there is no linear instability induced by the splitting. Note also that since the main error in the dispersive relation is of dispersive type, it is natural to try to remove it with techniques inspired by the classical *Lax-Wendroff* scheme. This is possible, but this does not yield better results. These investigations reassured us in the choice of a simple 2^{nd}-order splitting instead of a more elaborate strategy. Instead, we developed a Δt-optimized strategy for the choice of the parameter α.

A similar splitting approach is applied in [A10] for the 3-parameters family $(\mathcal{G}_{\alpha,\theta,\gamma})$.

The 1d case

In [A10, A11], we choose to use an hybrid FVM-FDM discretization of the 1d equations. As the simulation of highly nonlinear short-waves transformations requires high-order methods to avoid as much as possible damping and dispersive errors, we have implemented high-order schemes for the transport step, like the V6 scheme [22] and the WENO5 reconstruction with negative weight splitting [95, 96], while a 4^{th}-order FDM is used for the dispersive correction.

The global approach automatically preserves motionless equilibria, providing that a well-balanced scheme is implemented for the fully discrete *transport steps* $S_1(\frac{\Delta t}{2})$. Concerning the robustness property, while a local degeneracy to 2^{nd}-order conservative MUSCL schemes was initially implemented to achieve robustness within the WENO5 implementation in [A10], we have more recently adapted the recent accuracy preserving strategy of [111] for the WENO5 approach. This relies on the reconstruction of an *Hermite* polynomial per cell from the mean values and the WENO5 interface values, and the application of a suitable accuracy preserving limiter on pointwise values computed on the *Gauss-Lobatto* distribution of quadrature nodes.

The inversion of the unsymmetric matrix coming from the FD discretization of $I+\alpha\mathbf{T}$ is performed in the CSR format with a sparse direct solver (*UMFPACK*, see [35]).

The 2d case

In 2d, the *transport+dispersive correction* splitting approach results, for the $(\mathcal{CG}^{diag}_{\alpha})$ and $(\mathcal{CG}^{diag}_{\alpha,\theta,\gamma})$ models, in a high-order hybrid vertex-centered WENO-FVM-FDM fully discrete scheme on *cartesian* meshes. The robustness and preservation of motionless equilibrium are again inherited from the FVM discretization of the transport part.

1. Some technical difficulties are arising in the WENO-FVM 2d framework, as it is not straightforward to ensure an accuracy-preserving robustness property with WENO schemes. Indeed, the recent strategy of [111] can be successfully adapted but some additional technical difficulties come from the fact that the classical WENO strategy is inherently 1d, and only provides edges values (*i.e.* the local in-cell high-order 2d polynomials are not available as with DG method). In [A20], the method to ensure the robustness while preserving the high-order accuracy relies on the following ingredients:
• some suitable distributions of quadrature nodes on the element $[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] \times [y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}]$, obtained by tensor products of *Gauss* and *Gauss-Lobatto* rules in each direction,
• the construction of a bunch of WENO edge values per cell (at each *Gauss* nodes),
• the reconstruction of a bunch of 1d *Hermite* polynomials per cell, to get the nodal values at the relevant quadrature nodes,
• the limitation of these polynomials to deduce the revised edge values, injected into the numerical fluxes.

   The resulting scheme is shown to be robust under the following CFL condition:

$$\frac{\Delta t}{\Delta x}\|\,|v_1| + \sqrt{gh}\|_\infty + \frac{\Delta t}{\Delta y}\|\,|v_2| + \sqrt{gh}\|_\infty \leq \hat{\omega}_1, \tag{3.22}$$

   where $\|\,|v_1| + \sqrt{gh}\|_\infty$ and $\|\,|v_2| + \sqrt{gh}\|_\infty$ are some approximated maximum faces wave speeds, computed within the *Riemann* solver respectively in the first and second direction. We have $\hat{\omega}_1 = 1/6$ for the WENO3 method and $\hat{\omega}_1 = 1/12$ for the WENO5. Obviously, trying to ensure both high-accuracy and robustness lead to a stringent time-step restriction. An efficient implementation relies on a more relaxed CFL condition and, if a preliminary calculation to the next time step produces negative water height, we re-run the computation with the limitations and the more stringent restriction.

2. The computation of the terms $\left[I + \alpha \mathbf{T}^b_{diag}\right]^{-1}\left(\frac{1}{\alpha} gh\nabla\zeta + h(\mathcal{Q}_1(V) + g\mathcal{Q}_2(\zeta)) + g\mathcal{Q}_3(\zeta)\right)$ and $\left[I + \alpha \mathbf{T}^b_{diag}\right]^{-1}(gh\nabla\zeta)$ results in the resolution of sparse unsymmetric linear systems. For each term, the *diagonal* structure of the operator $\left[I + \alpha \mathbf{T}^b_{diag}\right]$ allows to consider 2 sparse linear systems of size N_{dof} instead of one system of size $2N_{dof}$.

   Additionally, as $I + \alpha \mathbf{T}^b_{diag}$ is a constant operator in time, the associated matrix is build and *LU*-factorized in the *compressed sparse row* format in a preprocessing step, and used through the whole computation at each time step or substep without being modified.
   To quantify the resulting improvements, we implemented the same whole numerical approach for the diagonal ($\mathcal{G}^{diag}_\alpha$) models. The resulting comparison is shown on Fig. 3.2. To be fair-play with the ($\mathcal{G}^{diag}_\alpha$) models, we use a *Bi-Conjugate Gradient Stabilized (BiCGStab)* iterative method, together with the ILUT preconditionner and a reverse *Cuthill-McKey* re-ordering algorithm.

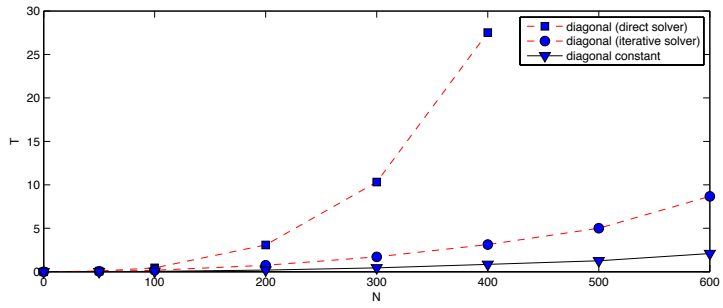

Figure 3.2: Propagation of a solitary wave over a flat bottom: comparison between the cpu-time per time step between the ($\mathcal{G}_{\alpha}^{diag}$) model and the ($\mathcal{CG}_{\alpha}^{diag}$) model for an increasing number of cells.

Applications

An extensive set of applications and validations has been performed for the 1d and 2d numerical codes, based on the *transport+dispersion correction* approach.
After the inclusion of wave breaking strategies, which will be detailed in §3.3, the resulting numerical models are able to account for a large variety of phenomena occurring in the nearshore area, as for instance the propagation of tsunami-like wave fronts, see [A12]. Such waves can evolve into a large range of bore types, from undular non-breaking bore to purely breaking bore. It is the complex competition between nonlinear effects, dispersive effects and energy dissipation which governs their transformations, making the prediction of their evolution a challenging task for numerical models. In [A12], we show that our numerical code is able to predict bore shapes, transformation and dynamics for a large range of *Froude* numbers. Note that to our knowledge, such encouraging results were not available in the specialized literature before.
Let us also mention the studies of the generation, propagation and breaking of infra gravity waves in the communication [C10] or of the effects of the bore transformations on wave run-up over sloping beaches in [C8] and the Ph.D. thesis of *M.Tissier* [T1]. For the 2d model, we also refer to the communication [C13] for applications of the ($\mathcal{CG}_{\alpha,\theta,\gamma}^{diag}$) models to overtopping problems. An example of wave shoaling over a varying semi-circular topography is shown on Fig. 3.3:

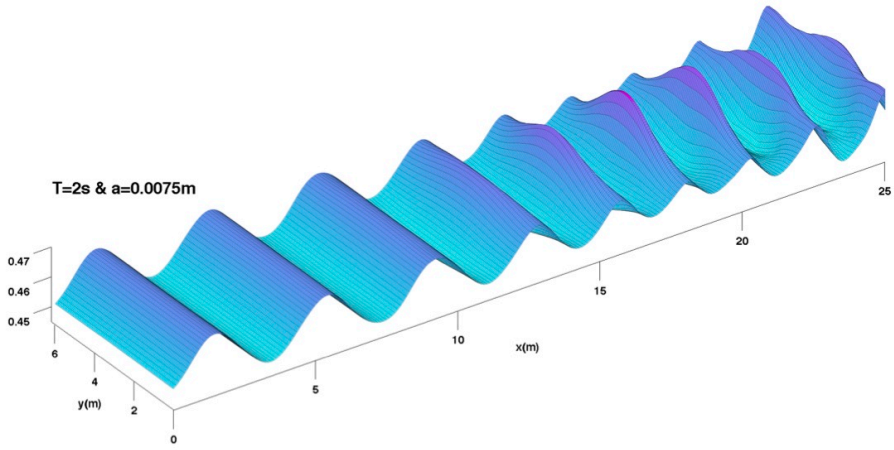

Figure 3.3: Propagation of periodic waves over a semi-circular shoal - ($T = 2\,s$ and $a = 0.0075m$): free surface elevation at $t = 100\,s$.

3.2.2 An unsplitted discontinuous-Galerkin discretization

In an effort to provide high order accuracy on unstructured meshes, we recently introduced in [A19] some arbitrary order DGMs for the $(\mathcal{G}_\alpha)$ and $(\mathcal{CG}_\alpha^{diag})$ families, relying on the non-conforming approach described in §1.2.6 for the elliptic and dispersive parts of the equations.
Surprisingly, DG discretizations of asymptotic shallow water dispersive models have been largely under-investigated and to our knowledge, [A19] is the first DG discretization available for the GN equations (we have to mention however the recent work [75], where 2^{nd} and 3^{rd}-order central DG methods are used for the hyperbolic part, while the dispersive part discretization is performed with a classical continuous FEM). For the time being, only 1d propagating waves have been investigated and the 2d extension is currently under study.

Semi-discrete weak formulation

Let us focus on the $(\mathcal{CG}_\alpha^{diag})$ models. Due to the number of nonlinear terms and high-order derivatives involved in the dispersive terms, we do not explicitly write equation (3.17) as a first-order system. Instead, we choose in [A19] to reformulate the dimensionalized 1d version of $(\mathcal{CG}_\alpha^{diag})$ as follows, using the PB-NSW formulation:

$$\partial_t \eta + \partial_x q = 0\,, \tag{3.23a}$$
$$\partial_t q + \partial_x \mathcal{H}^q(\mathbf{u}, b) + \mathcal{D} = \mathcal{S}_b^q(\mathbf{u}, b)\,, \tag{3.23b}$$

where $\mathcal{D}$ is an auxiliary scalar valued variable, obtained as the solution of the following auxiliary problem:

$$\begin{cases} [1 + \alpha \mathbf{T}^b](\mathcal{D} + \frac{1}{\alpha} g h \partial_x \eta) = \mathcal{K}, \\ \mathcal{K} = \frac{1}{\alpha} g h \partial_x \eta + h(\mathcal{Q}_1(v) + g\mathcal{Q}_2(\eta)) + g\mathcal{Q}_3(\mathcal{L}), \\ [1 + \alpha \mathbf{T}^b]\mathcal{L} = g h \partial_x \eta. \end{cases} \tag{3.24}$$

The weak formulation reads as follows: *find* $\mathbf{u}_\mathfrak{h} = (\eta_\mathfrak{h}, q_\mathfrak{h}) \in (\mathcal{V}_\mathfrak{h})^2$ *such that,* $\forall (\phi_\mathfrak{h}, \pi_\mathfrak{h}) \in (\mathcal{V}_\mathfrak{h})^2$, *and* $\forall \mathcal{C}_j \in \mathcal{P}_\mathfrak{h}$ *we have:*

$$\int_{\mathcal{C}_j} \partial_t \eta_\mathfrak{h} \phi_\mathfrak{h} dx - \int_{\mathcal{C}_j} q_\mathfrak{h} \phi_\mathfrak{h}' dx + \left[q_\mathfrak{h}\phi_\mathfrak{h}\right]_{x_l^j}^{x_r^j} = 0, \tag{3.25}$$

$$\int_{\mathcal{C}_j} \partial_t q_\mathfrak{h} \pi_\mathfrak{h} dx - \int_{\mathcal{C}_j} \mathcal{H}^q(\mathbf{u}_\mathfrak{h}, b_\mathfrak{h}) \pi_\mathfrak{h}' dx + \left[\mathcal{H}^q(\mathbf{u}_\mathfrak{h}, b_\mathfrak{h})\pi_\mathfrak{h}\right]_{x_l^j}^{x_r^j} \tag{3.26}$$

$$= \int_{\mathcal{C}_j} \mathcal{S}_b^q(\mathbf{u}_\mathfrak{h}, b_\mathfrak{h}) \pi_\mathfrak{h} dx - \int_{\mathcal{C}_j} \mathcal{D}_\mathfrak{h} \pi_\mathfrak{h} dx.$$

Seeking on each cell $\mathcal{C}_j$ for polynomial expansions, we write $\forall x \in \mathcal{C}_j, \forall t \in [0, t_{max}]$:

$$\eta_{\mathfrak{h}|\mathcal{C}_j}(x,t) = \sum_{i=1}^{N_p} \tilde{\eta}_i^j(t)\theta_i^j(x) \quad \text{and} \quad q_{\mathfrak{h}|\mathcal{C}_j}(x,t) = \sum_{i=1}^{N_p} \tilde{q}_i^j(t)\theta_i^j(x), \tag{3.27}$$

where $\{\theta_i^j\}_{i=1\ldots N_p}$ is the local *Lagrange* interpolant basis on the element $\mathcal{C}_j$, defined on the LGL set of N_p nodes and

$$\tilde{\eta}^j = {}^t(\tilde{\eta}_1^j(t), \ldots, \tilde{\eta}_{N_p}^j(t)), \qquad \tilde{q}^j = {}^t(\tilde{q}_1^j(t), \ldots, \tilde{q}_{N_p}^j(t)),$$

are the expansion coefficients vectors. We consider the local polynomial expansion $b_{\mathfrak{h}|\mathcal{C}_j}(x)$ for the topography parameterization and we assume that a local expansion is available for $\mathcal{D}_\mathfrak{h}$, given by:

$$\tilde{\mathcal{D}}^j = {}^t(\tilde{\mathcal{D}}_1^j(t), \ldots, \tilde{\mathcal{D}}_{N_p}^j(t)).$$

We obtain the semi-discrete formulation of our DGM:

$$\sum_{i=1}^{N_p} \frac{d}{dt}\tilde{\eta}_i^j(t)\mathbf{M}_{ik}^j - \sum_{i=1}^{N_p} \tilde{q}_i^j(t)\mathbf{S}_{ik}^j + \left[\widehat{q}\theta_k^j\right]_{x_l^j}^{x_r^j} = 0, \tag{3.28}$$

$$\sum_{i=1}^{N_p} \frac{d}{dt}\tilde{q}_i^j(t)\mathbf{M}_{ik}^j - \int_{\mathcal{C}_j} \mathcal{H}^q(\mathbf{u}_\mathfrak{h}, b_\mathfrak{h})\theta_k^{j'} dx + \left[\widehat{H}\theta_k^j\right]_{x_l^j}^{x_r^j} = \int_{\mathcal{C}_j} \mathcal{S}_b^q(\mathbf{u}_\mathfrak{h}, b_\mathfrak{h})\theta_k^j dx - \sum_{i=1}^{N_p} \tilde{\mathcal{D}}_i^j(t)\mathbf{M}_{ik}^j, \tag{3.29}$$

for $1 \leq k \leq N_p$, and $1 \leq j \leq N_e$, where

$$\mathbf{M}_{ik}^j = \int_{\mathcal{C}_j} \theta_i^j(x)\theta_k^j(x)dx \quad \text{and} \quad \mathbf{S}_{ik}^j = \int_{\mathcal{C}_j} \theta_i^j(x)\frac{d}{dx}\theta_k^j(x)dx$$

are respectively the local *mass* and *stiffness* matrix coefficients on $\mathcal{C}_j$.

High-order derivatives and dispersive terms computation

While we use the strategy described in §2.3.6 to discretize the hyperbolic part of (3.29) and compute the interface fluxes $(\widehat{q}, \widehat{H})$, we are left with the resolution of systems (3.24) to compute the expansion coefficients $\tilde{\mathcal{D}}$. Although the semi-discrete formulation (3.28)-(3.29) could appear as purely local, the computation of the dispersive term expansions $\tilde{\mathcal{D}}$ requires a global assembly process, for gathering the local discrete operators into a global one, in the computation of the inverse of the operator $1 + \alpha\mathbf{T}^b$. Additionally, this computation involves the discrete approximation of

space derivatives up to 3^{rd} order.

Applying the discretization strategy briefly recalled in §1.2.6, all the derivatives are computed in a DG sense (1.69)-(1.70). This leads us to introduce the following steps:

♯ Weak DG discretization of $1 + \alpha \mathbf{T}^b$

The differential operator $1 + \alpha \mathbf{T}^b$ involves 1^{st} and 2^{nd} order on h_b and, to compute these derivatives in a weak DG sense, we build some global weak derivative matrices. For the computation of 1^{st} and 2^{nd} order derivatives of an arbitrary quantity w, let us denote

$$-\vartheta = \partial_x w, \quad \chi = \partial_x^2 w, \tag{3.30}$$

and consider the weak formulation associated with the resulting coupled 1^{st} order equations:

$$\begin{aligned} \int_{\mathcal{C}_j} \vartheta_{\mathfrak{h}} \phi_{\mathfrak{h}} - \int_{\mathcal{C}_j} w_{\mathfrak{h}} \frac{d}{dx} \phi_{\mathfrak{h}} + [\widehat{w} \phi_{\mathfrak{h}}]_{x_l^j}^{x_r^j} = 0, \\ \int_{\mathcal{C}_j} \chi_{\mathfrak{h}} \pi_{\mathfrak{h}} - \int_{\mathcal{C}_j} \vartheta_{\mathfrak{h}} \frac{d}{dx} \pi_{\mathfrak{h}} + [\widehat{\vartheta} \pi_{\mathfrak{h}}]_{x_l^j}^{x_r^j} = 0. \end{aligned} \tag{3.31}$$

We aim at computing the expansion coefficients of the derivatives:

$$\tilde{\vartheta}^j = {}^t(\tilde{\vartheta}_1^j(t), \dots, \tilde{\vartheta}_{N_p}^j(t)), \qquad \tilde{\chi}^j = {}^t(\tilde{\chi}_1^j(t), \dots, \tilde{\chi}_{N_p}^j(t)).$$

We define the exchanging fluxes $\widehat{w}$ and $\widehat{\vartheta}$ with one of the fluxes (1.71)-(1.72)-(1.73)-(1.74). In [A19], we focus on the BR and LDG fluxes:

$$\widehat{w} = \{\{w_{\mathfrak{h}}\}\} - \boldsymbol{\beta} \cdot [\![w_{\mathfrak{h}}]\!], \quad \widehat{\vartheta} = \{\{\vartheta_{\mathfrak{h}}\}\} + \boldsymbol{\beta} \cdot [\![\vartheta_{\mathfrak{h}}]\!] - \frac{\nu}{\mathfrak{h}_j} [\![w_{\mathfrak{h}}]\!]. \tag{3.32}$$

Gathering the N_e expansion coefficients vectors $\{\tilde{w}^j\}_{i=1..N_e}$ in a $N_d \times N_e$ vector $\tilde{w} = {}^t(\tilde{w}^1, \dots, \tilde{w}^{N_e})$, the computations of the $N_d \times N_e$ vectors $\tilde{\vartheta} = {}^t(\tilde{\vartheta}^1, \dots, \tilde{\vartheta}^{N_e})$ and $\tilde{\chi} = {}^t(\tilde{\chi}^1, \dots, \tilde{\chi}^{N_e})$ respectively for the 1^{st} and 2^{nd} order derivatives is done globally. Injecting $\tilde{w}^j$ into (3.31) and replacing $\phi_{\mathfrak{h}}$ by the local basis functions for all elements, we obtain the global discrete formulation:

$$\begin{aligned} \mathbb{M}\tilde{\vartheta} &= \mathbb{S}\tilde{w} - (\mathbb{E} - \beta\mathbb{F})\, \tilde{w}\,, \\ \mathbb{M}\tilde{\chi} &= \mathbb{S}\tilde{\vartheta} - (\mathbb{E} + \beta\mathbb{F})\, \tilde{\vartheta} - \frac{\nu}{\mathfrak{h}} \mathbb{F}\tilde{w}\,, \end{aligned} \tag{3.33}$$

where the square $N_d \times N_e$ global mass and stiffness matrices $\mathbb{M}$ and $\mathbb{S}$ have a block-diagonal structure and the matrices $\mathbb{E}$ and $\mathbb{F}$ accounts for the inter-element exchanging fluxes. These matrices are detailed in [A19]. It leads to a global discrete weak DG formulation of the 1^{st} and 2^{nd} order derivative operators, with the numerical fluxes incorporated in the matrices:

$$\mathbb{D}_x = \mathbb{M}^{-1} (\mathbb{E} - \mathbb{S} - \beta\mathbb{F})\,, \tag{3.34}$$

$$\mathbb{D}_x^2 = \mathbb{M}^{-1} \Big((\mathbb{S} - \mathbb{E} - \beta\mathbb{F})\, \mathbb{D}_x - \frac{\nu}{\mathfrak{h}} \mathbb{F} \Big). \tag{3.35}$$

Even if we obtain a global matrix through this global assembly process, all the derivatives are still computed in an element-wise sense and (3.34) and (3.35) are then used to assemble the square $N_d \times N_e$ matrix of the weak DG discretization of $1 + \mathbf{T}^b$. For instance, considering the simplified flat bottom case ($h_b = h_0$), we approximate the corresponding operator as follows:

$$1 + \alpha \mathbf{T}^b = 1 - \alpha \frac{h_0^2}{3} \partial_x^2 \quad \Rightarrow \quad I - \alpha \frac{h_0^2}{3} \mathbb{D}_x^2 \,,$$

where I is the $N_p \times N_e$ identity matrix. The locality of the LDG approach results in a sparse block-structure matrix, which is stored in a sparse format and LU-factorized at the beginning of the computation, in a pre-processing step. Then, the factorization and the resolution of the resulting triangular linear systems can be performed with any good direct solver for sparse matrix.

♯ Discretization of $\mathcal{K}$

Still in the process of computing $\widetilde{\mathcal{D}}$, we need to compute expansion coefficients for the auxiliary variable $\mathcal{K}$. The differential operators $(\mathcal{Q}_i)_{1\leq i\leq 3}$ involve high-order derivatives, up to 3^{rd}-order (with respect to ζ). A formulation similar to (3.31) is obtained for the 3^{rd} order derivative, using again the BR or LGL fluxes

We choose to compute all the elemental derivatives at the N_p *Gauss-Lobatto* nodes, and then to assemble the weak DG discretization of $(\mathcal{Q}_i)_{1\leq i\leq 3}$ in a collocation-interpolation manner. Even if such a choice is known to possibly introduce some *aliasing*, it is mainly motivated by computational efficiency, as this has to be done at each time step. Indeed, the alternative would be to compute all the derivatives values on a suitable set of quadrature points (chosen accordingly to the polynomial order of the considered nonlinear term) and then compute the inner products consistently, possibly with over-integration strategies. But we recall that the $(\mathcal{Q}_i)_{1\leq i\leq 3}$ involve nonlinear terms which themselves are rational functions of the conservatives variables and that we are committed to introduce some aliasing error anyway.

We show in [A19] the following results:

Theorem 6. *We consider the DGM (3.28)-(3.29) for the approximation of* $(\mathcal{CG}_\alpha^{diag})$ *and we assume that*

1. *the interface fluxes* $\widehat{q}$ *and* $\widehat{\mathcal{H}}$ *defined accordingly to* §2.3.6*, with the LF flux (1.59),*
2. *the integrals* $\int_{\mathcal{C}_j} \mathcal{H}^q(\mathbf{u}_\mathfrak{h}, b_\mathfrak{h}) \frac{d}{dx}\theta_k^j dx$ *and* $\int_{\mathcal{C}_j} \mathcal{S}_b^q(\mathbf{u}_\mathfrak{h}, b_\mathfrak{h})\theta_k^j dx$ *are computed exactly when a motionless steady state is reached,*
3. $\overline{h}^n_{\mathfrak{h}|\mathcal{C}_j} \geq 0$ *and that* $\eta^n_{\mathfrak{h}|\mathcal{C}_j}(x) - b_{\mathfrak{h}|\mathcal{C}_j}(x) \geq 0$, $\forall x \in S^p_{\mathcal{C}_j}$*, the set of LGL points on the element* $\mathcal{C}_j$*, for a polynomial interpolation of order* p*,* $\forall j$*,*

then the scheme exactly preserves the motionless steady states and we have $\overline{h}^{n+1}_{\mathfrak{h}|\mathcal{C}_j} \geq 0$ *for all* j *under the CFL condition*

$$\lambda_o \frac{\Delta t}{|\mathcal{C}_j|} \leq \hat{w}_1,$$

with $\hat{w}_1 = \frac{1}{6}$ *for* $p = 2, 3$ *and* $\frac{1}{12}$ *for* $p = 4, 5$.

We show in [A19] that the expected convergence rates are obtained up to a 6^{th}-order scheme.
Here is an example of the simulations obtained with this approach:

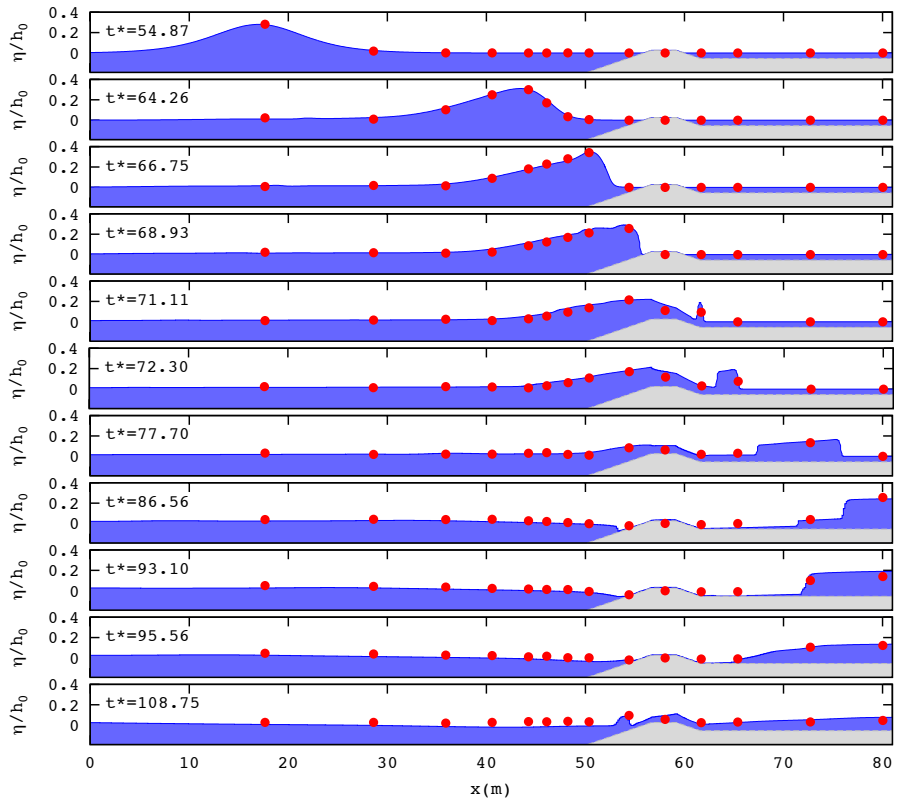

Figure 3.4: Overtopping over fringing reefs [88]: comparison between computed and experimental data for the free surface elevation, at several times during the propagation. Circles denote experimental data.

The same numerical approach has also been implemented, for comparison purpose, for the $(\mathcal{G}_\alpha)$ family. For this model, the matrix associated with $1 + \alpha\mathbf{T}$ has to be computed and assembled at each time step or sub-step. A quantification of the computational savings obtained with the use of the $(\mathcal{CG}_\alpha^{diag})$ family is performed in [A19].

3.3 Breaking waves: local dynamic switching

The notion of *wave breaking* in the shallow water asymptotics studied here is a difficult mathematical problem. A rough characterization of a solution that involves wave breaking (while remaining a single valued function) would be that at a given time, this solution remains bounded while its slope becomes unbounded. Physically, broken waves can be (not so roughly) regarded as waves of quasi-permanent form

whose shore-facing front is characterized by a recirculating mass of water, usually refers to as the *surface roller*.
We saw that BT and GN equations take into account, at different degrees of accuracy, both nonlinear and dispersive aspects of wave propagation in the nearshore area. The GN equations can accurately predict most phenomena exhibited by non-breaking waves in finite depth. However, these equations are derived upon the assumption of irrotational flow, and the effects of vorticity injection at the bottom and free-surface are neglected. As they do not include intrinsically energy dissipation, wave breaking cannot occur within these models and they become invalid to model the surf zone.

As a consequence, with BT and GN equations, the waves are allowed to steepen until the front face is almost vertical. Then, fundamental dissipative mechanisms have to be numerically introduced *a posteriori* through wave breaking dissipation (and also possibly parametrized bottom friction as numerically studied in §2.2). Several attempts have been made to introduce wave breaking in BT and GN models by the mean of *eddy viscosity* techniques [31, 29]. This approach generally requires:

1. the inclusion of an energy dissipation mechanism through the activation of additional dissipative terms in the governing equations when wave breaking is likely to occur,
2. explicit criteria to activate/deactivate these extra terms,
3. a method to follow the waves during their propagation since the breaking parameterizations depend on the age of the breaker. Moreover, the breaking model parameters need to be calibrated to ensure that the artificially induced energy dissipation is in agreement with the rate of energy dissipated in surf zone waves. It is an important point since the amount of energy dissipated will determine crucial phenomena such as the water level set-up and will impact wave-driven circulation.

As a consequence, each use of the model implies the tuning of several parameters. Additionally, it is difficult to predict simultaneously accurate wave height and asymmetry along the surf zone within this approach.

In §3.1 and §3.2 we saw that our GN models naturally degenerate to NSW equations when dispersive effects are negligible. Therefore, in order to handle wave breaking, an alternative approach is to *locally* switch in space and time from GN to NSW equations in the vicinity of the wave fronts, when the wave is ready to break, by suppressing the dispersive term. By switching to NSW equations, we decide to represent breaking wave fronts as shocks. Of course, the detailed process of breaking (spilling / overturning) is not modelled, but we conserve mass and momentum across the wave front, and allow energy to dissipate according to the shock theory (*i.e.* without any parameterization). We show in [A19, A14, A12] that this method allows for a natural treatment of wave breaking, which gives the correct amount of energy dissipation and a good description of broken-wave distortion and celerity.

3.3.1 *Energy dissipation* based sensor

In [A14], see also the communication [C8], we introduce a new *physical based* approach to switch between the 2 sets of equations, that takes benefits from the *trans-*

port + velocity dispersion correction splitting approach described in §3.2.1. With this local treatment, we obtain an approach able to account simultaneously and accurately for the effects of dispersion, non-linearities and wave breaking. This is of importance when considering for instance tsunami wave front transformations in the nearshore, which can evolve into a large range of bore types, including partially breaking undular bores. Moreover, the proposed "wave-by-wave" treatment allows for a precise description of the breaking events, from the initiation to the termination.

We recall that the GN equations (3.3) conserve the following energy

$$E(t) = \frac{1}{2}\int_{\mathbb{R}^d} \left[g\zeta(t)^2 + h(t)|\mathbf{v}(t)|^2 + h(t)\mathcal{T}[h,b]\mathbf{v}(t)\cdot\mathbf{v}(t)\right],$$

which reduces when $d = 1$ and at 1^{st}-order in μ, to $E(t) = \frac{1}{2}\int_{\mathbb{R}}(hv^2 + g\zeta^2)$, which is the conserved energy for the NSW equations, when smooth solutions are considered.

Therefore, we use the first *transport* step as a predictor, and detect the forming shocks (and distinguish broken fronts and fronts likely to break from others) through the study of the *energy dissipation*. Indeed, the energy dissipation forms peaks at the steepest parts of the wave fronts when shocks are forming, see Fig. 3.5. In practice, we compute, at each time step, an estimation of the local energy dissipation $\mathscr{D}(x,t)$ during the first *transport* step, as follows:

$$\mathscr{D}(x,t) = -(\partial_t \mathscr{E}(x,t) + \partial_x \mathscr{F}(x,t)),$$

where $\mathscr{E}$ and $\mathscr{F}$ are the local energy and energy flux densities, and localize the local extrema of $\mathscr{D}$.

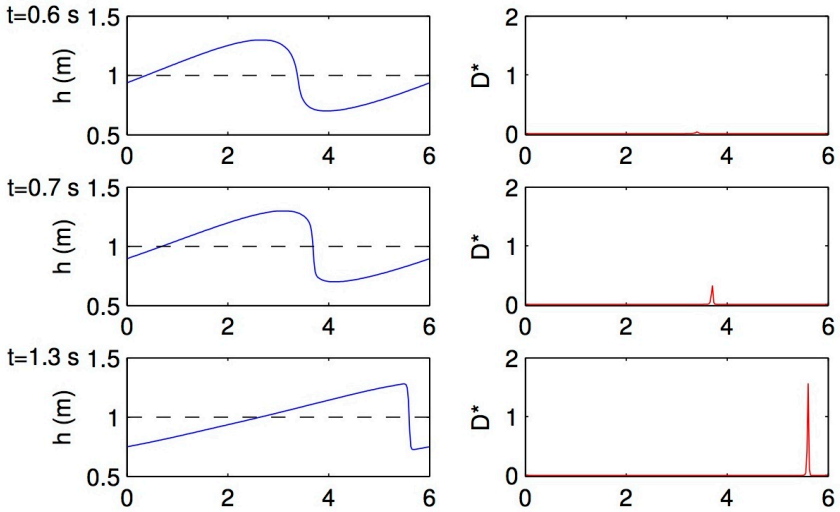

Figure 3.5: Generation of a shock and associated dissipation

Then, an efficient criterion for initiation and termination of breaking, based both on the critical front slope [91] and the *Froude* number across the fronts, is applied to dynamically determine the area where to locally switch in space and time between

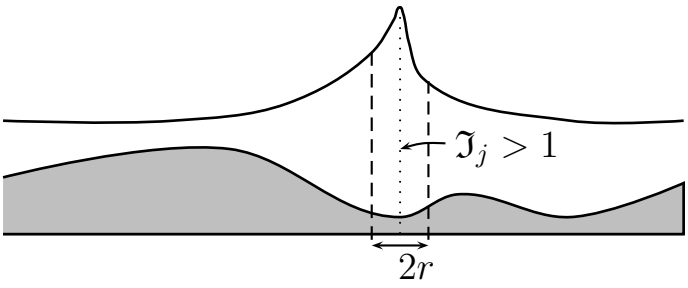

Figure 3.6: Example of breaking area

GN and NSW equations. After breaking, the wave fronts are handled as shocks by the NSW equations: we suppress the *velocity dispersive correction* step in the splitting approach.
The whole *dynamic switching* algorithm flowchart is documented in [A14], and further validated in [A12] for the study of *undular* tsunami-like waves.

3.3.2 A *DGM local super-convergence* based sensor

In [A19], we choose to take benefit from the strong super-convergence property of the DGM at the outflow boundary of each element in smooth regions of the flow [2], as used in the discontinuity detector of [61], as an alternative way of detecting the forming shocks. More precisely, for a given element $\mathcal{C}_j$, we compute the following indicator:

$$\mathfrak{I}_j = \frac{\Delta^- |\tilde{h}_1^j - \tilde{h}_{N_p}^{j-1}| + \Delta^+ |\tilde{h}_{N_p}^j - \tilde{h}_1^{j+1}|}{|\mathcal{C}_j|^{\frac{p+1}{2}} \|\tilde{h}^j\|_\infty}, \tag{3.36}$$

where

$$\Delta^- = \begin{cases} 1 & \text{if} \quad \tilde{q}_1^j \geq 0, \\ 0 & \text{otherwise.} \end{cases}, \qquad \Delta^+ = \begin{cases} 1 & \text{if} \quad \tilde{q}_{N_p}^j \leq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{3.37}$$

Waves about to break are identified in elements such that $\mathfrak{I}_j > 1$ and again, we locally switch to the NSW equations in such elements and the broken waves are handled as shocks by the NSW equations.
Practically, our numerical investigations have shown that the DGMs are more sensible to the local switching than the hybrid FVM-FDM, especially for higher order of approximations: for a given breaking wave, the *switching areas* need to be slightly enlarged to prevent the possible occurrence of spurious oscillations. Some neighboring cells are added to the switching area to include the steepening shore facing side of the wave, and a part of the offshore facing side. This can be simply done with the help of a mask, defining the switching area as a band of length $2r$ centered to the elements verifying $\mathfrak{I}_j > 1$, see Fig. 3.6. Additionally, the use of a local limitation procedure, applied as a post-process, is required in identified elements to suppress *Gibbs*-type oscillations.

An example of dissipation-based *local dynamic switching* area is shown on Fig. 3.7 for the experiment of *Cox* (obtained with the dissipation-based sensor).

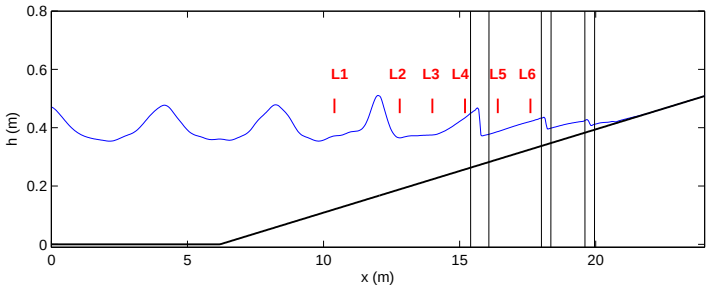

Figure 3.7: Spatial snapshot of cnoidal waves propagating over a 1:35 sloping beach. L1 to L6: locations of the wave gauges. Between 2 consecutive vertical lines, the flow is governed by NSW equations.

Author's publications

Research papers in peer-reviewed journals

[A1] F. Marche. Derivation of a new two-dimensional viscous shallow water model with varying topography, bottom friction and capillary effects. *Eur. J. Mech. B Fluids*, 26:49–63, 2007.

[A2] F. Marche, P. Bonneton, P. Fabrie, and N. Seguin. Evaluation of well-balanced bore-capturing schemes for 2d wetting and drying processes. *Internat. J. Numer. Methods Fluids*, 53(5):867–894, 2007.

[A3] C. Berthon and F. Marche. A positive preserving high order VFRoe scheme for shallow water equations: a class of relaxation schemes. *SIAM J. Sci.Comput.*, 30:2587–2612, 2008.

[A4] P. Fabrie and F. Marche. Another proof of stability for global weak solutions of 2d degenerated shallow water models. *J. Math. Fluid Mech.*, 11(4):536–551, 2009.

[A5] P. Bonneton, N. Bruneau, B. Castelle, and F. Marche. Large-scale vorticity generation due to dissipating waves in the surf zone. *Discrete Contin. Dyn. Syst. Ser. B.*, 13(4):729–738, 2010.

[A6] Q. Liang and F. Marche. Numerical resolution of well-balanced shallow water equations with complex source terms. *Advances in Water Resources*, 32(6):873 – 884, 2009.

[A7] A. Bouharguane, P. Azerad, F. Bouchette, F. Marche, and B. Mohammadi. Low complexity shape optimization and a posteriori high fidelity validation. *Discrete Contin. Dyn. Syst. Ser. B*, 13(4):759–772, 2010.

[A8] C. Berthon, F. Marche, and R. Turpault. An efficient scheme on wet/dry transitions for shallow water equations with friction. *Computers & Fluids*, 48(1):192 – 201, 2011.

[A9] P. Bonneton, E. Barthelemy, F. Chazel, R. Cienfuegos, D. Lannes, F. Marche, and M. Tissier. Recent advances in Serre-Green-Naghdi modelling for wave transformation, breaking and runup processes. *Eur. J. Mech. B Fluids*, 30(6):589–59, 2011.

[A10] F. Chazel, D. Lannes, and F. Marche. Numerical simulation of strongly nonlinear and dispersive waves using a Green-Naghdi model. *J. Sci. Comput.*, 48:105–116, 2011.

[A11] P. Bonneton, F. Chazel, D. Lannes, F. Marche, and M. Tissier. A splitting approach for the fully nonlinear and weakly dispersive Green-Naghdi model. *J. Comput. Phys.*, 230(4):1479 – 1498, 2011.

[A12] M. Tissier, P. Bonneton, F. Marche, F. Chazel, and D. Lannes. Nearshore dynamics of tsunami-like undular bores using a fully nonlinear Boussinesq model. *J. Coastal Res.*, SI 64:603–607, 2011.

[A13] O. Delestre and F. Marche. A numerical scheme for a viscous shallow water model with friction. *J. Sci. Comput.*, 48(1-3):41–51, 2011.

[A14] M. Tissier, P. Bonneton, F. Marche, F. Chazel, and D. Lannes. A new approach to handle wave breaking in fully non-linear Boussinesq models. *Coastal Engineering*, 67:54–66, 2012.

[A15] A. Duran, Q. Liang, and F. Marche. On the well-balanced numerical discretization of shallow water equations on unstructured meshes. *J. Comput. Phys.*, 235:565–586, 2013.

[A16] A. Duran and F. Marche. Recent advances on the discontinuous Galerkin method for shallow water equations with topography source terms. *Computers & Fluids*, 101:88–104, 2014.

[A17] M. Guerra, R. Cienfuegos, C. Escauriaza, F. Marche, and J. Galaz. Modeling rapid flood propagation over natural terrains using a well-balanced scheme. *J. Hydraulic Res.*, 140(7), 2014.

[A18] A. Duran, F. Marche, C. Berthon, and R. Turpault. Asymptotic preserving scheme for the shallow water equations with source terms on unstructured meshes, *J. Comput. Phys.*, 287:184-206, 2015.

[A19] A. Duran and F. Marche. Discontinuous-Galerkin discretization of a new class of Green-Naghdi equations, *Comm. Comput. Phys.*, 17(3):721-760, 2015.

[A20] D. Lannes and F. Marche. A new class of fully nonlinear and weakly dispersive Green-Naghdi models for efficient 2d simulations, *to appear in J. Comput. Phys.*, 282:238-268, 2015.

[A21] S. Le Roy, R. Pedreros, C. André, F. Paris, S. Lecacheux, F. Marche, and C. Vinchon. Coastal flooding of urban areas by overtopping: dynamic modelling application to the Johanna storm (2008) in Gavres (France), *Natural Hazards and Earth System Sciences Disc.*, 2(8), 4947-4985, 2014.

Communications and proceedings

[C1] F. Marche and P. Bonneton. A simple and efficient well-balanced scheme for 2d bore propagation and run-up over a sloping beach. In *Proceedings of the* 31^{th} *International Conference In Coastal Engineering*, pages 998–1010. World Scientific, 2006.

[C2] F. Marche and P. Bonneton. Un nouveau modèle pour la simulation du phénomène de découvrement/recouvrement sur des bathymétries complexes. In *IXie Journées Nationales Génie Côtier-Génie Civil*, volume 9, pages 81–88. Paralia, 2007.

[C3] F. Marche and P. Bonneton. Un modèle "équilibre" d'ordre élevé pour les équations de Saint-Venant et extensions au cas dispersif pour la propagation de la houle en milieu littoral. In 18^{ie} *Congrés Français de Mécanique*. AFM Maison de la Mécanique, 2007.

[C4] F. Marche. A simple well-balanced model for 2D coastal engineering applications. In *Proceedings of the* 11^{th} *International Conference On Hyperbolic Problems*, pages 271–283. Springer, 2008.

[C5] F. Marche. A simple hybrid well-balanced method for a 2d viscous shallow water model. In H. Deconinck and E. Dick, editors, *Proceeding of the* 11^{th} *International Conference in Computational Fluid Dynamics*, pages 261–266. Springer, 2009.

[C6] C. Berthon and F. Marche. A robust high order VFRoe scheme for shallow water equations. In *Hyperbolic problems: theory, numerics and applications*, volume 67 of *Proc. Sympos. Appl. Math.*, pages 785-794. AMS, 2009.

[C7] C. Berthon and F. Marche. A positive well-balanced VFroe-ncv scheme for non-homogeneous shallow-water equations. In *Proceedings of the* 2^{nd} *ISCM and* 12^{th} *EPMESC*, volume 1233, pages 1495-1500, AIP Conference Proceedings, 2010.

[C8] M. Tissier, P. Bonneton, F. Marche, F. Chazel, and D. Lannes. Serre-Green-Naghdi modelling of wave transformation, breaking and run-up using a high-order Finite-Volume Finite-Difference scheme. In *Proceedings of the 32th Internatioanl Conference on Coastal Engineering*, volume 1, pages 158-169. Curran Associates, Inc., 2010.

[C9] F. Marche, P. Bonneton, M. Tissier, F. Chazel, and D. Lannes. Optimized Green-Naghdi equations for the modelling of waves nearshore transformations. In *Proceedings of the* 21^{st} *International Offshore and Polar Engineering Conference*, pages 835–840. International Society of Offshore and Polar Engineers, 2011.

[C10] M. Tissier, P. Bonneton, R. Gerben, F. Marche, D. Lannes, and F. Chazel. Fully nonlinear Boussinesq-type modelling of infra-gravity wave transformation over a low-sloping beach. In J. Smith and P. Lynett, editors, *Proceedings of the* 33^{th} *International Conference on Coastal Engineering*, volume 33. Coastal Engineering Research Council, 2012.

[C11] F. Marche, F. Chazel, and D. Lannes. A numerical method for a shallow water model with dispersive source terms. In *Proceeding of the* 13^{th} *International Conference On Hyperbolic Problems*, volume 17 of *Series in Contemporary Applied Mathematics*, pages 357–364. World Scientific, 2012.

[C12] A. Duran, F. Marche, C. Berthon, and R. Turpault. Numerical discretizations for shallow water equations with source terms on unstructured meshes. In *Proceedings of the* 14^{th} *International Conference on Hyperbolic Problems: Theory Numerics and Applications, Padova*, volume 8, pages 541–549, AIMS Series on Applied Mathematics, 2014.

[C13] F. Marche and D. Lannes. Some new enhanced 2d Green-Naghdi equations for the modeling of waves nearshore transformations. In *Proceedings of the* 24^{th} *International Ocean and Polar Engineering Conference, 2014.*

Advised and co-advised Ph.D.

[T1] M. Tissier. *Etude numérique de la transformation des vagues en zone littorale, de la zone de levée aux zones de surf et de jet de rive.* Université Bordeaux 1, 2011.

[T2] M. Cathala. *Problématiques d'analyse numérique et de modélisation pour écoulements de fluides environnementaux.* Université Montpellier 2, 2013.

[T3] A. Duran. *Numerical simulation of depth-averaged flows models : Finite Volume and discontinuous Galerkin approaches.* Université Montpellier 2, 2014.

Numerical softwares

[S1] WaveBox: *a numerical platform for shallow water free surface flows simulations, based on FVM and DGM approximations of Saint-Venant and Green-Naghdi equations.* Université Montpellier 2 and Inria LEMON, 2014.

References

[1] R. Abgrall. On essentially non-oscillatory schemes on unstructured meshes. *J. Comput. Phys.*, 114:45–58, 1994.

[2] S. Adjerid, K. Devine, J. Flaherty, and L. Krivodonova. A posteriori error estimation for discontinuous galerkin solutions of hyperbolic problems. *Comput. Methods Appl. Mech. Engrg*, pages 1097–1112, 2002.

[3] B. Alvarez-Samaniego and D. Lannes. Large time existence for 3d water-waves and asymptotics. *Invent. math.*, 171(3):485–541, 2008.

[4] D.N. Arnold, F. Brezzi, B. Cockburn, and L.D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, 39(5):1749–1779, 2002.

[5] H.L. Atkins and C.-W. Shu. Quadrature-free implementation of discontinuous Galerkin method for hyperbolic equations. *ICASE - Report No. 96-51*, 1996.

[6] 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(6):2050–2065, 2004.

[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(1):311–333, 2005.

[8] F. Bassi and S. Rebay. A high-order accurate discontinuous finite element method for the numerical solution of the compressible navier–stokes equations. *J. Comput. Phys.*, 131:267–279, 1997.

[9] T.B. Benjamin, J.L Bona, and J.J. Mahony. Model equations for long waves in nonlinear dispersive systems. *Phil. Trans. Roy. Soc. London A*, 272:47–78, 1972.

[10] A. Bermudez and M.-E. Vazquez. Upwind methods for hyperbolic conservation laws with source terms. *Comput. & Fluids*, 23(8):1049–1071, 1994.

[11] C. Berthon. Numerical approximations of the 10-moment gaussian closure. *Math. Comp.*, 75(256):1809–1831, 2006.

[12] C. Berthon. Robustness of MUSCL schemes for 2D unstructured meshes. *J. Comput. Phys.*, 218(2):495–509, 2006.

[13] C. Berthon, P. Lefloch, and R. Turpault. Late-time/stiff relaxation asymptotic-preserving approximations of hyperbolic equations. *Math. Comp.*, 82(282):831–860, 2013.

[14] C. Berthon and R. Turpault. Asymptotic preserving hll schemes. *Num. Method Partial Diff. Equations*, 27(6):1396–1422, 2011.

[15] Christophe Berthon. Stability of the muscl schemes for the euler equations. *Commun. Math. Sci.*, 3(2):133–157, 2005.

[16] S. Bianchini, B. Hanouzet, and R. Natalini. Asymptotic behavior of smooth solutions for partially dissipative hyperbolic systems with a convex entropy. *Communications Pure Appl. Math.*, 60:1559–1622, 2007.

[17] P. Bonneton. Modelling of periodic wave transformation in the inner surf zone. *Ocean Engineering*, 34(10):1459–1471, 2007.

[18] F. Bouchut. *Nonlinear stability of finite volume methods for hyperbolic conservation laws, and well-balanced schemes for sources.* Frontiers in Mathematics. Birkhauser, 2004.

[19] F. Bouchut. A reduced stability condition for noninear relaxation to conservations laws. *J. Hyp. Diff. Eq.*, pages 149–170, 2004.

[20] C. Buet, B. Després, and E. Frank. Design of asymptotic preserving finite volume schemes for the hyperbolic heat equation on unstructured meshes. *Numer. Math.*, 122(2):227–278, 2012.

[21] T. Buffard, T. Gallouët, and J.-M. Hérard. A sequel to a rough godunov scheme: application to real gases. *Computers and Fluids*, 2000.

[22] S. Camarri, M.V. Salvetti, B. Koobus, and A. Dervieux. A low-diffusion muscl scheme for les on unstructured grids. *Computers and Fluids*, 33(9):1101–1129, 2004.

[23] M. Castro, J. Gallardo, J. López-GarcÍa, and C. Parés. Well-balanced high order extensions of godunov's method for semilinear balance laws. *SIAM Journal on Numerical Analysis*, 2008.

[24] M. Cathala. *Problématiques d'analyse numérique et de modélisation pour écoulements de fluides environnementaux.* PhD thesis, Université Montpellier 2, 2012.

[25] C. Chalons, F. Coquel, E. Godlewski, P.A. Raviart, and N. Seguin. Godunov-type schemes for hyperbolic systems with parameter dependent source. The case of Euler system with friction. *Math. Models and Methods in Appl. Sci.*, 20(11):2109–2166, 2010.

[26] G.Q. Chen, C.D. Levermore, and T. P. Liu. Hyperbolic conservation laws with stiff relaxation terms and entropy. *Comm. Pure Appl. Math.*, 47(6):787–830, 1994.

[27] M. Chen. Numerical investigation of a two-dimensional Boussinesq system. *Discrete and Continuous Dynamical Systems - S*, 23(4):1169—1190, 2009.

[28] Q. Chen. Fully nonlinear Boussinesq-type equations for waves and currents over porous beds. *Journal of Engineering Mechanics*, 2006.

[29] Q. Chen, J.T. Kirby, R. Dalrymple, and A. Kennedy. Boussinesq modeling of wave transformation, breaking, and runup. II: 2 D. *Journal of Waterway*, 2000.

[30] R. Cienfuegos, E. Barthélemy, and P. Bonneton. A fourth-order compact finite volume scheme for fully nonlinear and weakly dispersive Boussinesq-type equations. I: Model development and analysis. *Internat. J. Numer. Methods Fluids*, 51(11):1217–1253, 2006.

[31] R. Cienfuegos, E. Barthelemy, and P. Bonneton. A wave-breaking model for Boussinesq-type equations including mass-induced effects. *J. Wtrwy. Port Coast. and Oc. Engrg.*, 136:10–26, 2010.

[32] B. Cockburn and C.-W. Shu. The local discontinuous galerkin method for time-dependent convection-diffusion systems. *SIAM J. Numer. Anal.*, 141:2440–2463, 1998.

[33] R. Cools. Monomial cubature rules since "Stroud": a compilation - part 2. *J. Comput. Applied. Math.*, 112:21–27, 1999.

[34] F. Coquel and B. Perthame. Relaxation of energy and approximate Riemann solvers for general pressure laws in fluid dynamics. *SIAM J. Numer. Anal.*, 35:2223—2249, 1998.

[35] T.A. Davis and I.S. Duff. An unsymmetric-pattern multifrontal method for sparse LU factorization. *SIAM J. on Matrix Anal. and Applications*, 18:140–158, 1997.

[36] O. Delestre. Simulation du ruissellement d'eau de pluie sur des surfaces agricoles. *Thèse de Doctorat, Université d'Orléans*, 2010.

[37] D. Di Pietro and A. Ern. *Mathematical Aspects of Discontinuous Galerkin Methods*, volume 69 of *MATHEMATIQUES ET APPLICATIONS*. Springer, 2012.

[38] D. Dutykh, T. Katsaounis, and D. Mitsotakis. Finite volume schemes for dispersive wave propagation and runup. *J. Comput. Phys.*, 230:3035 – 3061, 2011.

[39] K. Erduran, S. Ilic, and V. Kutija. Hybrid finite-volume finite-difference scheme for the solution of Boussinesq equations. *Int. J. Numer. Meth. Fluids*, 49:1213–1232, 2005.

[40] A. Ern and J.-L. Guermond. Discontinuous Galerkin methods for Friedrichs' systems. i. general theory. *SIAM J. Numer. Anal.*, 44(2):753–778, 2006.

[41] A. Ern and J.-L. Guermond. Discontinuous Galerkin methods for Friedrichs' systems. ii. second-order elliptic pdes. *SIAM J. Numer. Anal.*, 44(6):2363–2388, 2006.

[42] A. Ern, S. Piperno, and K. Djadel. A well-balanced runge-kutta discontinuous galerkin method for the shallow-water equations with flooding and drying. *Internat. J. Numer. Methods Fluids*, 58(1):1–25, 2008.

[43] C. Eskilsson and S. J. Sherwin. An hp/spectral element model for efficient long-time integration of Boussinesq-type equations. *Coastal Engineering*, 45:295 – 320, 2003.

[44] C. Eskilsson and S.J.Sherwin. Spectral/hp discontinuous galerkin methods for modelling 2d Boussinesq equations. *J. Comput. Phys.*, 212:566–589, 2006.

[45] R. Eymard, T. Gallouet, and R. Herbin. *Finite Volume Methods*. Number 7 in Handbook of Numerical Analysis,. J.L.Lions ed., 2006.

[46] T. Gallouët, J.M. Hérard, and N. Seguin. Some approximate godunov schemes to compute shallow-water equations with topography. *Computers and Fluids*, 2003.

[47] E. Godlewski and P.-A. Raviart. *Numerical approximation of hyperbolic systems of conservation laws*. Springer, 1996.

[48] L. Gosse and G. Toscani. Asymptotic-preserving well-balanced scheme for the hyperbolic heat equations. *C. R. Math. Acad. Sci. Paris*, 334:337–342, 2002.

[49] A. E Green and P. M Naghdi. A derivation of equations for wave propagation in water of variable depth. *Journal of Fluid Mechanics Digital Archive*, 78(02):237–246, 1976.

[50] J.M. Greenberg and A.Y. Leroux. A well-balanced scheme for the numerical processing of source terms in hyperbolic equations. *SIAM Journal on Numerical Analysis*, 33(1):1–16, 1996.

[51] B. Hanouzet and R. Natalini. Global existence of smooth solutions for partially dissipative hyperbolic systems with a convex entropy. *Arch. Ration. Mech. Anal.*, 169(2):89–117, 2003.

[52] A. Harten, P.D. Lax, and B. Van Leer. On upstream differencing and Godunov-type schemes for hyperbolic conservation laws. *SIAM Review*, 25(1):35–61, 1983.

[53] J Hérard, N Seguin, and T Gallouet. On the use of symmetrizing variables for vacuums. *Calcolo*, 40(3):163–194, 2003.

[54] J.S. Hesthaven. From electrostatics to almost optimal nodal sets for polynomial interpolation in a simplex. *SIAM*, 35(2):655–676, 1998.

[55] S. Jin and Z. Xin. The relaxation scheme for systems of conservation laws in arbitrary space dimension. *Comm. Pure Appl. Math.*, 45(235–276), 1995.

[56] N. Kaewbanjak, S. Grilli, J. Kirby, and P. Watts. Modeling the 26 december 2004 indian ocean tsunami: Case study of impact in thailand. *Journal of Geophysical Research*, 2007.

[57] M. Kazolea, A.I. Delis, I.K. Nikolos, and C.E. Synolakis. An unstructured finite volume numerical scheme for extended 2d Boussinesq-type equations. *Coastal Engineering*, 69:42–66, 2012.

[58] G. Kesserwani and Q. Liang. A discontinuous Galerkin algorithm for the two-dimensional shallow water equations. *Comput. Methods Appl. Mech. Engrg.*, 199(49-52):3356–3368, 2010.

[59] G. Kesserwani and Q. Liang. Locally limited and fully conserved rkdg2 shallow water solutions with wetting and drying. *J. Sci. Comput.*, 50(120–144), 2012.

[60] Georges Kesserwani and Qiuhua Liang. A conservative high-order discontinuous Galerkin method for the shallow water equations with arbitrary topography. *Internat. J. Numer. Methods Engrg.*, 86(1):47–69, 2011.

[61] L. Krivodonova, J. Xin, J.-F. Remacle, N. Chevaugeon, and J.E. Flaherty. Shock detection and limiting with discontinuous galerkin methods for hyperbolic conservation laws. *Applied Numerical Mathematics*, 48:323 – 338, 2004.

[62] D. Lannes. *The water waves problem: mathematical analysis and asymptotics*. Number 188 in Mathematical Surveys and Monographs. American Mathematical Society, 2013.

[63] O. Le Métayer, S. Gavrilyuk, and S. Hank. A numerical scheme for the Green-Naghdi model. *J. Comput. Phys.*, 34(229):2034–2045, 2010.

[64] B. Van Leer. Towards the ultimate conservative difference scheme. V. A second-order sequel to Godunov's method. *J. Comput. Phys.*, 135(2):227–248, 1997.

[65] R.J. Leveque. Geoclaw – a variant of clawpack for geophysical flows - http://depts.washington.edu/clawpack/geoclaw/.

[66] R.J. LeVeque. Balancing source terms and flux gradients in high-resolution godunov methods: the quasi-steady wave-propagation algorithm. *J. Comput. Phys.*, 146(1):346–365, 1998.

[67] R.J. LeVeque. *Finite Volume methods for hyperbolic problems*. Cambridge Texts in Applied Mathematics. Cambridge University Press, 2002.

[68] R.J. Leveque and M. Pelanti. A class of approximate riemann solvers and their relation to relaxation schemes. *J. Comput. Phys.*, 172(572–591), 2001.

[69] D. Levy, C.-W. Shu, and J. Yan. Local Discontinuous Galerkin Methods for nonlinear dispersive equations. *J. Comput. Phys.*, 751-772(196), 2004.

[70] Y. Li, S. Liu, Y. Yu, and G. Lai. Numerical modeling of Boussinesq equations by finite element method. *Coastal Engineering*, 1999.

[71] T.P. Liu. Hyperbolic conservation laws with relaxation. *Comm. Math. Phys.*, 108(153–175), 1988.

[72] M.S. Longuet-Higgins and R.W. Stewart. Radiation stresses in water waves: a physical discussion, with applications. *Deep-Sea Research*, 11:529–562, 1964.

[73] P. Lynett, T. Wu, and P. Liu. Modeling wave runup with depth-integrated equations. *Coastal Engineering*, 46:89—107, 2002.

[74] P.A. Madsen and O.R. Sorensen. A new form of the Boussinesq equations with improved linear dispersion characteristics: Part 2. a slowly-varying bathymetry. *Coastal Engineering*, 16:183–204, 1996.

[75] L. Maojin, P. Guyenne, F. Li, and L. Xu. High order well-balanced cdg-fe methods for shallow water waves by a Green-Naghdi model. *J. Comput. Phys.*, 257, Part A(0):169 – 192, 2014.

[76] J.M. Masella, I. Faille, and T. Gallouet. On an approximate godunov scheme. *Int. J. of Computational Fluid Dynamics*, 12(2):133–149, 1999.

[77] C.C. Mei and B. Le Méhauté. Note on the equations of long waves over an uneven bottom. *J. Geophys. Res.*, 71:393–400, 1966.

[78] J. Miles and R. Salmon. Weakly dispersive nonlinear gravity waves. *J. Fluid Mech.*, 157:519–531, 1985.

[79] R. Natalini. Convergence to equilibrium for the relaxation approximation of conservation laws. *Comm. Pure Appl. Math.*, 49:1–30, 1996.

[80] O. Nwogu. Alternative form of Boussinesq equations for nearshore wave propagation. *Journal of Waterway*, 119:618–638, 1993.

[81] J.D. Pearce and J.G. Esler. A pseudo-spectral algorithm and test cases for the numerical solution of the two-dimensional rotating Green-Naghdi shallow water equations. *J. Comput. Phys.*, 229:7594–7608, 2010.

[82] R. Pedreros, Delvallée E., S. Lecacheux, Y. Balouin, M. Garcin, Y. Krien, G. Le Cozannet, B. Poisson, and J. Thiebot. Mise en place d'une approche multi-modèles pour éval- uer l'exposition du littoral languedocien aux submersions marines dans un contexte de changement climatique. In *Journées "Impacts du Changement Climatique sur les Risques Côtiers", Orléans, France*, 2010.

[83] M. Pelanti, F. Bouchut, and A. Mangeney. A riemann solver for single-phase and two-phase shallow flow models based on relaxation. relations with roe and vfroe solvers. *J. Comput. Phys.*, 230:515–550, 2011.

[84] B. Perthame and C.-W. Shu. On positivity preserving finite volume schemes for euler equations. *Numer. Math.*, 73:119–130, 1996.

[85] Benoît Perthame and Youchun Qiu. A variant of Van Leer's method for multidimensional systems of conservation laws. *J. Comput. Phys.*, 112(2):370–381, 1994.

[86] M. Ricchiuto and A.G. Filippini. Upwind Residual discretization of enhanced Boussinesq equations for wave propagation over complex bathymetries. *J. Comput. Phys.*, in press, 2014.

[87] Volker Roeber and Kwok Fai Cheung. Boussinesq-type model for energetic breaking waves in fringing reef environments. *Coastal Engineering*, 70(0):1 – 20, 2012.

[88] Volker Roeber, Kwok Fai Cheung, and Marcelo H. Kobayashi. Shock-capturing Boussinesq-type model for nearshore wave processes. *Coastal Engineering*, 57(4):407 – 423, 2010.

[89] B. Rogers, A. Borthwick, and P. Taylor. Mathematical balancing of flux gradient and source terms prior to using Roe's approximate riemann solver. *J. Comput. Phys.*, 192:422–451, 2003.

[90] B. Rogers, M. Fujihara, and A. Borthwick. Adaptive Q-tree Godunov-type scheme for shallow water equations. *Int. J. Numer. Methods Fluids*, 35:247–280, 2001.

[91] H. Schaeffer. A Boussinesq model for waves breaking in shallow water. *Coastal Engineering*, 20:185–202, 1993.

[92] F.J. Seabra-Santos, D.P. Renouard, and A.M. Temperville. Numerical and experimental study of the transformation of a solitary wave over a shelf or isolated obstacle. *J. Fluid Mech.*, 176:117–134, 1987.

[93] F Serre. Contribution à l'étude des écoulements permanents et variables dans les canaux. *Houille Blanche*, 8:374–388, 1953.

[94] F. Shi, J.T. Kirby, J.C. Harris, J.D. Geiman, and S.T. Grilli. A high-order adaptive time-stepping TVD solver for Boussinesq modeling of breaking waves and coastal inundation. *Ocean Modelling*, 43-44:36–51, 2012.

[95] Jing Shi, Changqing Hu, and Chi-Wang Shu. A technique of treating negative weights in weno schemes. *J. Comput. Phys.*, 175(1):108–127, 2002.

[96] C.-W. Shu. High order weighted essentially non-oscillatory schemes for convection dominated problems. *SIAM Review*, 51:82–126, 2009.

[97] C.H. Su and C.S. Gardner. Korteweg–de vries equation and generalizations. iii. derivation of the korteweg–de vries equation and burgers equation. *J. Math. Phys.*, 10(3):536–539, 1969.

[98] J. Tao. Numerical modeling of wave runup and breaking on the beach. *Acta Oceanol. Sin.*, 6(5):692–700, 1984.

[99] M. Tonelli and M. Petti. Hybrid finite volume-finite difference scheme for 2dh improved Boussinesq equations. *Coastal Engineering*, 56(5-6):609–620, 2009.

[100] E.F. Toro, M. Spruce, and W. Speare. Restoration of the contact surface in the hll riemann solver. *Shock waves*, 4(25-34), 1994.

[101] G. Wei and J. Kirby. Time-dependent numerical code for extended Boussinesq equations. *Journal of Waterway*, 121:251–261, 1995.

[102] Y. Xing and C.-W. Shu. A new approach of high orderwell-balanced finite volume weno schemes and discontinuous galerkin methods for a class of hyperbolic systems with source terms. *Commun. Comput. Phys.*, 1:100–134, 2006.

[103] Y. Xing and S.-W. Shu. High order well-balanced finite volume weno schemes and discontinuous galerkin methods for a class of hyperbolic systems with source terms. *J. Comput. Phys.*, 214:567–598, 2006.

[104] Y. Xing and X. Zhang. Positivity-preserving well-balanced discontinuous galerkin methods for the shallow water equations on unstructured triangular meshes. *J. Sci. Comput.*, 57:19–41, 2013.

[105] Y. Xu and C.-W. Shu. Local Discontinuous Galerkin Methods for three classes of nonlinear wave equations. *J. Comput. Math.*, 22:250–174, 2004.

[106] Y. Xu and C.-W. Shu. Local Discontinuous Galerkin methods for high-order time-dependent partial differential equations. *Commun. Comput. Phys.*, 7(1):1–46, 2010.

[107] J. Yan and C.-W. Shu. A Local Discontinuous Galerkin method for KdV-type equations. *SIAM J. Numer. Anal.*, 40:769–791, 2002.

[108] J. Yan and C.-W. Shu. Local Discontinuous Galerkin methods for partial differential equations with higher order derivatives. *J. Sci. Comput.*, 17:27–47, 2002.

[109] V.E. Zakharov. Stability of periodic waves of finite amplitude on the surface of a deep fluid. *J. Appl. Mech. Tech. Phys*, 2:190–194, 1968.

[110] J.A. Zelt. The runup of nonbreaking and breaking solitary waves. *Coastal Engineering*, 15:205–246, 1991.

[111] X. Zhang and C.-W. Shu. On maximum-principle-satisfying high order schemes for scalar conservation laws. *J. Comput. Phys.*, 229(9):3091 – 3120, 2010.

[112] X. Zhang, Y. Xia, and C.-W. Shu. Maximum-principle-satisfying and positivity-preserving high order discontinuous Galerkin schemes for conservation laws on triangular meshes. *J. Sci. Comput.*, 50:29–62, 2012.